ROUTLEDGE

The Routledge Companion to Social and Political Philosophy

Edited by Gerald Gaus and
Fred D'Agostino

THE ROUTLEDGE COMPANION TO SOCIAL AND POLITICAL PHILOSOPHY

Routledge Philosophy Companions

Routledge Philosophy Companions offer thorough, high quality surveys and assessments of the major topics and periods in philosophy. Covering key problems, themes and thinkers, all entries are specially commissioned for each volume and written by leading scholars in the field. Clear, accessible and carefully edited and organised, *Routledge Philosophy Companions* are indispensable for anyone coming to a major topic or period in philosophy, as well as for the more advanced reader.

The Routledge Companion to Aesthetics, Second Edition
Edited by Berys Gaut and Dominic Lopes

The Routledge Companion to Philosophy of Religion
Edited by Chad Meister and Paul Copan

The Routledge Companion to the Philosophy of Science
Edited by Stathis Psillos and Martin Curd

The Routledge Companion to Twentieth Century Philosophy
Edited by Dermot Moran

The Routledge Companion to Philosophy and Film
Edited by Paisley Livingston and Carl Plantinga

The Routledge Companion to Philosophy of Psychology
Edited by John Symons and Paco Calvo

The Routledge Companion to Metaphysics
Edited by Robin Le Poidevin, Peter Simons, Andrew McGonigal, and Ross Cameron

The Routledge Companion to Nineteenth Century Philosophy
Edited by Dean Moyar

The Routledge Companion to Ethics
Edited by John Skorupski

The Routledge Companion to Epistemology
Edited by Sven Bernecker and Duncan Pritchard

The Routledge Companion to Philosophy and Music
Edited by Theodore Gracyk and Andrew Kania

The Routledge Companion to Phenomenology
Edited by Søren Overgaard and Sebastian Luft

The Routledge Companion to Philosophy of Language
Edited by Gillian Russell and Delia Graff Fara

The Routledge Companion to Philosophy of Law
Edited by Andrei Marmor

The Routledge Companion to Social and Political Philosophy
Edited by Gerald Gaus and Fred D'Agostino

Forthcoming:

The Routledge Companion to Seventeenth Century Philosophy
Edited by Dan Kaufman

The Routledge Companion to Eighteenth Century Philosophy
Edited by Aaron Garrett

The Routledge Companion to Philosophy of Mental Disorder
Edited by Jakob Hohwy and Philip Gerrans

The Routledge Companion to Islamic Philosophy
Edited by Richard C. Taylor and Luis Xavier López-Farjeat

The Routledge Companion to Philosophy of Literature
Edited by Noël Carroll and John Gibson

The Routledge Companion to Philosophy of Religion, Second Edition
Edited by Chad Meister and Paul Copan

The Routledge Companion to Aesthetics, Third Edition
Edited by Berys Gaut and Dominic Lopes

The Routledge Companion to Philosophy of Science, Second Edition
Edited by Edited by Stathis Psillos and Martin Curd

PRAISE FOR THE SERIES

The Routledge Companion to Aesthetics

"This is an immensely useful book that belongs in every college library and on the bookshelves of all serious students of aesthetics."—***Journal of Aesthetics and Art Criticism***

"The succinctness and clarity of the essays will make this a source that individuals not familiar with aesthetics will find extremely helpful."—***The Philosophical Quarterly***

"An outstanding resource in aesthetics … this text will not only serve as a handy reference source for students and faculty alike, but it could also be used as a text for a course in the philosophy of art."—***Australasian Journal of Philosophy***

"Attests to the richness of modern aesthetics … the essays in central topics—many of which are written by well-known figures—succeed in being informative, balanced and intelligent without being too difficult."—***British Journal of Aesthetics***

"This handsome reference volume … belongs in every library."—***Choice***

"The *Routledge Companion*s to Philosophy have proved to be a useful series of high quality surveys of major philosophical topics and this volume is worthy enough to sit with the others on a reference library shelf."—***Philosophy and Religion***

The Routledge Companion to Philosophy of Religion

"… a very valuable resource for libraries and serious scholars."—***Choice***

"The work is sure to be an academic standard for years to come … I shall heartily recommend *The Routledge Companion to Philosophy of Religion* to my students and colleagues and hope that libraries around the country add it to their collections."—***Philosophia Christi***

The Routledge Companion to Philosophy of Science

A ***Choice*** Outstanding Academic Title 2008

"With a distinguished list of internationally renowned contributors, an excellent choice of topics in the field, and well-written, well-edited essays throughout, this compendium is an excellent resource. Highly recommended."—***Choice***

"Highly recommended for history of science and philosophy collections."—***Library Journal***

"This well conceived companion, which brings together an impressive collection of distinguished authors, will be invaluable to novices and experience readers alike."—***Metascience***

The Routledge Companion to Twentieth Century Philosophy

"To describe this volume as ambitious would be a serious understatement. … full of scholarly rigor, including detailed notes and bibliographies of interest to professional philosophers. … Summing Up: Essential."—***Choice***

The Routledge Companion to Philosophy and Film

"A fascinating, rich volume offering dazzling insights and incisive commentary on every page … Every serious student of film will want this book … Summing Up: Highly recommended."—***Choice***

The Routledge Companion to Philosophy of Psychology

"This work should serve as the standard reference for those interested in gaining a reliable overview of the burgeoning field of philosophical psychology. Summing Up: Essential."—***Choice***

The Routledge Companion to Metaphysics

"The *Routledge Philosophy Companions* series has a deserved reputation for impressive scope and scholarly value. This volume is no exception … Summing Up: Highly recommended."—***Choice***

The Routledge Companion to Nineteenth Century Philosophy

A ***Choice*** Outstanding Academic Title 2010

"This is a crucial resource for advanced undergraduates and faculty of any discipline who are interested in the 19th-century roots of contemporary philosophical problems. Summing Up: Essential."—***Choice***

The Routledge Companion to Ethics

"This fine collection merits a place in every university, college, and high school library for its invaluable articles covering a very broad range of topics in ethics[.] … With its remarkable clarity of writing and its very highly qualified contributors, this volume is must reading for anyone interested in the latest developments in these important areas of thought and practice. Summing Up: Highly recommended."—***Choice***

THE ROUTLEDGE COMPANION TO SOCIAL AND POLITICAL PHILOSOPHY

Edited by
Gerald Gaus and Fred D'Agostino

Routledge
Taylor & Francis Group
NEW YORK AND LONDON

First published 2013
by Routledge
711 Third Avenue, New York, NY 10017

Simultaneously published in the UK
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

Routledge is an imprint of the Taylor & Francis Group, an informa business

© 2013 Taylor & Francis

The right of the editors to be identified as the authors of the editorial material, and of the authors for their individual chapters, has been asserted in accordance with sections 77 and 78 of the Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this book may be reprinted or reproduced or utilized in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

Trademark notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Library of Congress Cataloging in Publication Data
The Routledge companion to social and political philosophy /
edited by Gerald F. Gaus and Fred D'Agostino
p. cm.–(Routledge philosophy companions)
Includes bibliographical references and index.
1. Political Science–Philosophy–History.
2. Social sciences–Philosophy–History.
3. Political sociology–Philosophy–History.
1. Gaus, Gerald F. II. D'Agostino, Fred.
JA71.R686 2012
300.1–dc23
2012013795

ISBN: 978–0–415–87456–4 (hbk)
ISBN: 978–0–203–09223–1 (ebk)

Typeset in Goudy by Swales & Willis Ltd, Exeter, Devon

CONTENTS

Notes on Contributors xiv
Preface xxii

PART I
The History of Social and Political Theory 1

1 Plato's Political Philosophy 3
GEORGE KLOSKO

2 Aristotle's Social and Political Philosophy 14
RACHANA KAMTEKAR

3 Aquinas 25
PAUL SIGMUND

4 Medieval Political Thought 36
CARY J. NEDERMAN

5 Machiavelli 47
VICKIE B. SULLIVAN

6 Hobbes 59
S.A. LLOYD

7 Locke 71
ERIC MACK

8 Rousseau 82
CHRISTOPHER BERTRAM

9 Hume and Smith on Justice 92
STEPHEN BUCKLE

10 Kant 103
OLIVER SENSEN

11 Hegel 114
DAVID EDWARD ROSE

12 Mill 124
C.L. TEN

13 Marx 135
DAVID LEOPOLD

14 Late Nineteenth- and Early Twentieth-Century British Thought 146
MARIA DIMOVA-COOKSON

15 Continental Political Philosophy 158
JAMES BOHMAN

16 French Political Thought in the Twentieth Century 169
JEREMY JENNINGS

17 The Political Philosophy of China 181
TONGDONG BAI

18 Indian Political Theory 192
A. RAGHURAMARAJU

19 Islamic Political Thought 204
ANDREW F. MARCH

PART II
Political Theories and Ideologies 215

20 Anarchism 217
RODERICK T. LONG

21 Liberalism 231
MICHAEL FREEDEN

22 Conservatism 243
JOHN KEKES

23 Republicanism 254
CHRISTIAN NADEAU

24 Marxism and Contemporary Political Thought 266
ALEX CALLINICOS

25 Feminism and the History of Political Philosophy 278
PENELOPE DEUTSCHER

26 Environmentalism 291
MATHEW HUMPHREY

PART III
Normative Foundations 303

27 Contractarianism 305
CLAIRE FINKELSTEIN

28 Contractualism and Political Liberalism 317
AARON JAMES

29 Utilitarianism and Consequentialism 329
DALE E. MILLER

30 Perfectionism 342
STEVEN WALL

31 Pluralism 353
GEORGE CROWDER

32 Virtue Ethics and Political Philosophy 364
DANIEL C. RUSSELL

33 Natural Law and Rights Theory 375
DAVID S. ODERBERG

PART IV
Distributive Justice 387

34 Luck Egalitarianism 389
ZOFIA STEMPLOWSKA

35 The Difference Principle 401
REX MARTIN

36 Left Libertarianism 412
HILLEL STEINER

37 Libertarianism 421
JOHN MEADOWCROFT

38 Desert 433
DAVID SCHMIDTZ

39 Needs and Distributive Justice 444
GILLIAN BROCK

40 The Capability Approach (and Social Justice) 456
INGRID ROBEYNS

41 Intergenerational Distributive Justice 467
CLARK WOLF

PART V
The National State and Beyond 481

42 Nationalism
MARGARET MOORE 483

43 Human Rights and Cosmopolitanism
DAVID A. REIDY 494

44 Multiculturalism
CHANDRAN KUKATHAS 505

45 Global Justice and Politics
THOM BROOKS 517

46 Justice and Borders
DAVID MILLER 526

47 War
FERNANDO R. TESÓN 537

PART VI
Political Concepts 549

48 Equality
THOMAS CHRISTIANO 551

49 Freedom
KATRIN FLIKSCHUH 562

50 Autonomy
HORACIO SPECTOR 573

51 Power
PETER MORRISS 585

52 Authority and Legitimacy
FABIENNE PETER 596

53 Democracy
ROBERT B. TALISSE 608

54 Rights
JONATHAN QUONG 618

55 Toleration
PETER JONES 629

PART VII
Approaches 641

56 Social Evolution 643
GERALD GAUS AND JOHN THRASHER

57 The Pragmatist Project in Political Philosophy 656
CHERYL MISAK

58 Postmodernism and Politics 667
TODD MAY

59 Social Choice Theory 679
JOHN A. WEYMARK

60 Rational Choice Theory 691
PETER VANDERSCHRAAF

61 Discourse Theory 706
WILLIAM REHG

PART VIII
Issues in Social and Political Philosophy 719

62 Education 721
HARRY BRIGHOUSE

63 Health 732
NORMAN DANIELS

64 Marriage, Sex, and the Family 744
DAVID ARCHARD

65 Work 755
NIEN-HÊ HSIEH

66 Punishment 765
MARK R. REIFF

67 Terrorism 777
C. A. J. COADY

68 Paternalism, Moralism, and Markets 788
MARK D. WHITE

69 Religion in Public Life 800
KEVIN VALLIER AND CHRISTOPHER EBERLE

Index 812

CONTRIBUTORS

David Archard is Professor of Philosophy and Public Policy at Lancaster University. He is the author of several books, and numerous chapters and articles in social, political, legal and applied moral philosophy.

Tongdong Bai is a Professor of Philosophy at Fudan University in China. The English and revised version of his book, *A New Mission of an Old State: Classical Confucian Political Philosophy in a Contemporary and Comparative Context* is being reviewed for publication, and his short introduction to Chinese political philosophy, *China: The Middle Way of the Middle Kingdom* is forthcoming.

Christopher Bertram is Professor of Social and Political Philosophy at the University of Bristol, and author of *Rousseau and The Social Contract* (2004) together with numerous articles in political philosophy.

James Bohman is the Danforth Professor of Philosophy and International Studies at Saint Louis University. He is author of *New Philosophy of Social Science*, *Public Deliberation* and *Democracy across Borders*, all with MIT Press.

Harry Brighouse is Professor of Philosophy and Affiliate Professor of Educational Policy Studies at the University of Wisconsin, Madison. He is Co-Director, with Michael McPherson, of the Spencer Foundation's Initiative on Philosophy in Educational Policy and Practice.

Gillian Brock is Associate Professor in Philosophy at the University of Auckland, New Zealand. She is the author of *Global Justice: A Cosmopolitan Account* and editor or co-editor of *Current Debates in Global Justice*, *The Political Philosophy of Cosmopolitanism*, *Necessary Goods: Our Responsibilities to Meet Others' Needs*, and *Global Heath and Global Health Ethics*.

Thom Brooks is Reader in Political and Legal Philosophy at Newcastle University. He is editor and founder of the *Journal of Moral Philosophy*. His books include *The Global Justice Reader* (2008) and *Global Justice and International Affairs* (2011).

Stephen Buckle is Senior Lecturer in Philosophy at the Australian Catholic University, Sydney. He is the author of *Natural Law and the Theory of Property* and *Hume's Enlightenment Tract*, and editor of Hume's *An Enquiry Concerning Human Understanding and Other Writings* and (with Craig Taylor) *Hume and the Enlightenment*.

Alex Callinicos was born in Harare, Zimbabwe. For many years, he taught political philosophy at the University of York. He is now Professor of European Studies at King's College London. His most recent books, all published by Polity Press, are *The Resources of Critique* (2006), *Imperialism and Global Political Economy* (2009), and *Bonfire of Illusions* (2010).

Thomas Christiano is Professor of Philosophy and Law at the University of Arizona. He has published *The Rule of the Many* (1996) and *The Constitution of Equality* (2008) and articles in democratic theory, distributive justice and international political philosophy. He is currently working on the foundations of equality, the legitimacy of international institutions and the basis of human rights. He is an editor of *Politics, Philosophy and Economics*.

C.A.J. Coady is Vice Chancellor's Fellow and Professorial Fellow in the Centre for Applied Philosophy and Public Ethics at the University of Melbourne. In 2012 he was Leverhulme Visiting Professor at the University of Oxford. His most recent books are *Morality and Political Violence* (2008) and *Messy Morality: the Challenge of Politics* (2008).

George Crowder is Professor in the Discipline of Politics and Public Policy, Flinders University, Adelaide, Australia. His books include *Liberalism and Value Pluralism* (2002), *Isaiah Berlin: Liberty and Pluralism* (2004), and *The One and the Many: Reading Isaiah Berlin* (co-edited with Henry Hardy, 2007).

Norman Daniels is Mary B. Saltonstall Professor and Professor of Ethics and Population Health at Harvard School of Public Health. His most recent books include *Just Health* (2008) and *Setting Limits Fairly* (with Jim Sabin, 2008).

Penelope Deutscher is Professor of Philosophy at Northwestern University. Her publications include *The Philosophy of Simone de Beauvoir: Ambiguity, Conversion, Resistance* (2008), *Yielding Gender: Feminism, Deconstruction and the History of Philosophy* (1997), and (co-edited with F. Collin) *Repenser le politique: l'apport du féminisme* (2004).

Maria Dimova-Cookson is a Lecturer in Politics at Durham University and Director of the Centre for the History of Political Thought. Until July 2010 she was one of the editors of the journal *Contemporary Political Theory*. Her latest books are *T.H.Green: Ethics, Metaphysics and Political Philosophy* (co-edited with W.J. Mander) (Oxford University Press, 2006), and *Multiculturalism and Moral Conflict* (co-edited with P.M.R. Stirk) (Routledge 2009). She is currently writing a book *Positive and Negative Liberty in Past and Present Liberal Thought*, to be published by Routledge.

Christopher Eberle is an Associate Professor of Philosophy at the United States Naval Academy. The author of *Religious Conviction in Liberal Politics* (Cambridge, 2002) and various journal articles, he is currently writing a book on religion, just war and military service in a liberal democracy.

Claire Finkelstein is the Algernon Biddle Professor of Law and Professor of Philosophy at the University of Pennsylvania, and Director of the Center for Ethics and the Rule of Law at the University of Pennsylvania Law School. She has published in the areas

of criminal law theory, moral and political philosophy, philosophy of law, rational choice theory, and more recently, military ethics. She is currently completing a book entitled *Contractarian Legal Theory*.

Katrin Flikschuh is a Reader in Modern Political Theory at the London School of Economics. She works mainly on Kant's political philosophy and its relation to contemporary liberalism. She is currently working on a new book, *Kant contra Cosmopolitanism: Assessing the Global Justice Debate*.

Michael Freeden is Professor of Politics, Director of the Centre for Political Ideologies at the University of Oxford, and Professorial Fellow of Mansfield College, Oxford. His books include *The New Liberalism* (1978), *Liberalism Divided* (1986), *Ideologies and Political Theory* (1996), and *Liberal Languages* (2005). He is founder-editor of the *Journal of Political Ideologies*.

Gerald Gaus is the James E. Rogers Professor of Philosophy at the University of Arizona, where he directs the program in Philosophy, Politics, Economics, and Law. Among his books are *On Philosophy, Politics and Economics* (2008), *Justificatory Liberalism* (1996), and *Value and Justification* (1990). His most recent book is *The Order of Public Reason: A Theory of Freedom and Morality in a Diverse and Bounded World* (2011).

Nien-hê Hsieh is Associate Professor of Legal Studies and Business Ethics at the Wharton School, University of Pennsylvania. He writes about ethical issues in business and economics, incommensurable values, and ways to organize economic production to meet the demands of justice.

Mathew Humphrey is Reader in Political Philosophy at the University of Nottingham. He is the author of *Preservation versus the People?* (2002) and *Ecological Politics and Democratic Theory* (2007).

Aaron James is Associate Professor of Philosophy at the University of California, Irvine. He was an ACLS Burkhardt fellow, a fellow at CASBS, Stanford University, and is author of *Fairness in Practice: A Social Contract for a Global Economy* (2012).

Jeremy Jennings is Director of the Centre for the Study of the History of Political Thought at Queen Mary University of London. He is the author of *Revolution and the Republic: A History of Political Thought in France Since the Eighteenth Century* (2011).

Peter Jones is Professor of Political Philosophy at the University of Newcastle, UK. He is the author and editor of books on rights and has written on many different aspects of rights, including human rights, group rights, democratic rights, welfare rights, and rights of free expression. His published work has also ranged over a number of other subjects, including toleration, identity, recognition, cultural diversity, democracy, neutrality, international justice, and the nature of liberalism.

Rachana Kamtekar is Associate Professor of Philosophy at the University of Arizona. She works primarily in ancient philosophy and has written extensively on Plato's political philosophy. She is currently writing a book at on Plato's psychology.

John Kekes is Professor Emeritus at the University at Albany. He has published many books, some of which are *Against Liberalism* (1997), *A Case for Conservatism* (1998), *The Illusions of Egalitarianism* (2003), and *The Roots of Evil* (2005).

George Klosko is Henry L. and Grace Doherty Professor of Politics at the University of Virginia. He works in both normative political theory and the history of political thought. His books include *Political Obligations* (2005) and *The Oxford Handbook of the History of Political Philosophy* (2011), which he edited.

Chandran Kukathas holds the Chair in Political Theory in the Department of Government at the London School of Economics. He is the author of *The Liberal Archipelago* (2003), and of *Hayek and Modern Liberalism* (1989).

David Leopold is University Lecturer in Political Theory in the Department of Politics and International Relations, University of Oxford, and a Fellow of Mansfield College, Oxford. His publications include an edition of William Morris's *News from Nowhere*, and a monograph entitled *The Young Karl Marx. German Philosophy, Modern Politics, and Human Flourishing.*

S.A. Lloyd is Professor of Philosophy, Law, and Political Science at the University of Southern California. She writes on topics in Hobbes, Rawls, and contemporary feminist philosophy.

Roderick T. Long is Professor of Philosophy at Auburn University, President of the Molinari Institute, author of *Wittgenstein, Austrian Economics, and the Logic of Action*, and co-editor of *Anarchism/Minarchism* (2008).

Eric Mack is a Professor of Philosophy and Faculty Member of the Murphy Institute of Political Economy at Tulane University. He is the author of *John Locke* and of numerous articles on natural rights, property rights, distributive justice, the scope of legitimate state authority, and many related topics.

Andrew F. March is Associate Professor of Political Science at Yale University and is the author of *Islam and Liberal Citizenship* (2009). He received the Award for Excellence in the Study of Religion from the American Academy of Religion.

Rex Martin is Professor of Philosophy, Emeritus, at the University of Kansas and Honorary Professor at Cardiff University. His fields of major interest include political and legal philosophy and the history of political thought. Among his books are *A System of Rights* (1997) and *Rawls's Law of Peoples: A Realistic Utopia?*, co-editor (2006).

Todd May is Class of 1941 Memorial Professor of the Humanities at Clemson University. He is the author of ten books of philosophy, mostly engaged with the work of Michel Foucault, Gilles Deleuze, and Jacques Rancière. He is currently working on a book on friendship in the neoliberal period.

John Meadowcroft is Senior Lecturer in Public Policy in the Department of Political Economy at King's College London. He is the author of *The Ethics of the Market* (2005) and *James M. Buchanan* (2011), co-editor of *The Cambridge Companion to*

Nozick's Anarchy, State, and Utopia (2011), and editor of the *Major Conservative and Libertarian Thinkers* series.

Dale E. Miller is Professor of Philosophy at Old Dominion University. His research mainly concerns different aspects of utilitarianism, with J.S. Mill as a particular interest. He is the author of *J. S. Mill: Moral, Social and Political Thought* (2010) and a co-editor of *Morality, Rules, and Consequences* (2000) and *John Stuart Mill and the Art of Life* (2011).

David Miller is Professor of Political Theory at the University of Oxford, and an Official Fellow of Nuffield College. Among his books are *On Nationality* (Clarendon Press, 1995), *Principles of Social Justice* (Harvard University Press, 1999) and *National Responsibility and Global Justice* (Oxford University Press, 2007).

Cheryl Misak is Professor of Philosophy at the University of Toronto, where she also serves as Vice-President and Provost. She is the author of *Truth and the End of Inquiry: A Peircean Account of Truth*, *Verificationism: Its History and Prospects*, and *Truth, Politics, Morality: Pragmatism and Deliberation*. Her new book *The American Pragmatists* will come out in 2012.

Margaret Moore is Sir Edward Peacock Professor of Political Theory at Queen's University (Kingston, Canada). She has published books and articles on distributive justice, democratic theory, nationalism and multiculturalism, including: *Foundations of Liberalism* and *Ethics of Nationalism*.

Peter Morriss is the author of *Power: A Philosophical Analysis* (second edition, 2002; originally 1987). He has also written many articles on power, on applying political philosophy to various issues, and on the politics of South Korea. He has taught in a number of universities, including Liverpool University, Manchester University, and the National University of Ireland, Galway.

Christian Nadeau is a Professor at the Université de Montréal where he teaches political philosophy. He has published numerous essays, chapters, and books, including (with D. Weinstock) *Republicanism: History, Theory and Practice* (2004), and (with B. Dubreuil) *Elster. Passions, raisons et délibération* (2011).

Cary J. Nederman is Professor of Political Science at Texas A&M University. He is the author or editor of approximately twenty books and has contributed more than 100 articles to academic journals and edited collections.

David S. Oderberg is Professor of Philosophy at the University of Reading, and has published many books and articles on moral philosophy, including *Moral Theory* (2000), *Applied Ethics* (2000), and (co-edited with Timothy Chappell) *Human Values: New Essays on Ethics and Natural Law* (2004).

Fabienne Peter is an Associate Professor of Philosophy at the University of Warwick. She specializes in political philosophy and in issues at the intersection of economics and philosophy. She is the author of *Democratic Legitimacy* (2008).

Jonathan Quong is Senior Lecturer in Political Philosophy at the University of Manchester. He is the author of *Liberalism Without Perfection* (2011), as well as articles on liberalism, public reason, democracy, distributive justice, and the morality of defensive harm.

A. Raghuramaraju teaches philosophy at the University of Hyderabad. His books include: *Debates in Indian Philosophy* which got *Choice* Outstanding Academic Title in 2007, *Debating Gandhi: A Reader*, *Enduring Colonialism: Classical Presences and Modern Absences in Indian Philosophy*, and *Modernity in Indian Social Theory*. His essay "Rethinking West" is included in *Documenta Magazine 12*.

William Rehg is Professor of Philosophy at Saint Louis University. He is the author of *Cogent Science in Context: The Science Wars, Argumentation Theory, and Habermas* (2009) and *Insight and Solidarity: The Discourse Ethics of Jürgen Habermas* (1994), and the translator of Habermas's *Between Facts and Norms* (1996).

David A. Reidy is Professor of Philosophy at the University of Tennessee. He has published widely in political and legal philosophy and in particular on John Rawls and on human rights.

Mark R. Reiff is a Senior Lecturer in Legal and Political Philosophy at the University of Manchester School of Law. He is the author of *Punishment, Compensation, and Law: A Theory of Enforceability* (2005), as well as various papers on topics within legal, political, and moral philosophy.

Ingrid Robeyns is a Professor of Practical Philosophy at the Erasmus University, Rotterdam. Her main area of research is normative political philosophy, especially theories of justice and normative questions regarding the economy, the family, and development. She has also published widely on the capability approach.

David Edward Rose is a Lecturer in Philosophy at Newcastle University. He is mostly interested in applying Hegel's social ethics to contemporary issues and has published widely on Hegel, Vico, and the history of European political and philosophical thought.

Daniel C. Russell is Professor of Philosophy in the University of Arizona's Center for the Philosophy of Freedom. He is the author of *Plato on Pleasure and the Good Life* (2005), *Practical Intelligence and the Virtues* (2009), *Happiness for Humans* (forthcoming), and editor of the *Cambridge Companion to Virtue Ethics* (forthcoming).

David Schmidtz is Kendrick Professor of Philosophy, joint Professor of Economics, and founding Director of the Center for Philosophy of Freedom at the University of Arizona. He recently published "Nonideal Theory: What It Is and What It Needs To Be" in *Ethics*.

Oliver Sensen is Associate Professor in Philosophy at Tulane University. He is the author of *Kant on Human Dignity* (2011), editor of *Kant on Moral Autonomy* (2012), and the author of several articles on Kant's moral philosophy.

Paul Sigmund is Professor Emeritus of Politics at Princeton University. He has published nineteen books on political theory and comparative politics, including *Nicholas of Cusa and Medieval Political Thought* (1963), *St. Thomas Aquinas On Politics and Ethics* (1988), the English translation of the fifteenth-century conciliarist text by Nicholas of Cusa, *The Catholic Concordance* (1995), and *The Selected Political Writings of John Locke* (2005).

Horacio Spector is Dean and Professor of Law and Philosophy at the School of Law of Universidad Torcuato Di Tella (Buenos Aires). He is the author of *Autonomy and Rights, The Moral Foundations of Liberalism* (1992, 2008).

Hillel Steiner is a Fellow of the British Academy and Professor of Political Philosophy at Manchester (Emeritus) and Salford Universities. He is the author of *An Essay on Rights* and co-author (with Matthew Kramer and Nigel Simmonds) of *A Debate Over Rights: Philosophical Enquiries*.

Zofia Stemplowska is University Lecturer in Political Theory at Oxford University and a Fellow of Worcester College. Her research focuses on domestic and global distributive justice. She is the co-editor (with Carl Knight) of *Responsibility and Distributive Justice* (2011).

Vickie B. Sullivan is a Professor of Political Science at Tufts University in Massachusetts. Her two books examine Machiavelli's thought and its influence. She is also the editor of *The Comedy and Tragedy of Machiavelli: Essays on the Literary Works* (2000). Her articles have appeared in *The American Political Science Review*, *History of European Ideas*, *History of Political Thought*, *Political Theory*, and *Polity*.

Robert B. Talisse is Professor of Philosophy at Vanderbilt University. He works primarily on issues in liberalism and democratic theory. His most recent book is *Democracy and Moral Conflict* (2009).

C.L. Ten is Professor of Philosophy at the National University of Singapore. He was previously Professor of Philosophy (Personal Chair) at Monash University, Australia. His books include *Mill on Liberty* and *Crime, Guilt, and Punishment*. He has also edited two volumes of essays on Mill, *Mill's Moral, Political, and Legal Philosophy* and *Mill's* On Liberty: *A Critical Guide*.

Fernando R. Tesón, a native of Buenos Aires, is the Tobias Simon Eminent Scholar at Florida State University. He is the author of *Humanitarian Intervention: An Inquiry into Law and Morality* (2005), *Rational Choice and Democratic Deliberation* (2006) (with Guido Pincione), *A Philosophy of International Law* (1998), and numerous articles in journals.

John Thrasher is a PhD candidate in philosophy at the University of Arizona. He is writing a dissertation on the evolutionary foundations of contractarian moral and political theory.

Kevin Vallier is an Assistant Professor of Philosophy at Bowling Green State University. His current research focuses on the proper place of religious commitment in liberal politics and various facets of the public reason liberal political tradition.

Peter Vanderschraaf is Associate Professor of Philosophy at the University of California Merced. His research focuses on analyzing social conventions and investigating their proper role in moral and political philosophy.

Steven Wall is Professor of Philosophy at the University of Arizona where he is also a member of the Center for the Philosophy of Freedom. He is the author of *Liberalism, Perfectionism and Restraint* (1998) and an editor (with David Sobel) of *Reasons for Action* (2009).

John A. Weymark is the Gertrude Conaway Vanderbilt Professor of Economics at Vanderbilt University. His research focuses on social choice theory and normative economics. He is a former President of the Society for Social Choice and Welfare and Managing Editor of *Social Choice and Welfare*.

Mark D. White is Professor in the Department of Political Science, Economics, and Philosophy at the College of Staten Island/CUNY, and is the author of *Kantian Ethics and Economics: Autonomy, Dignity, and Character* (2011).

Clark Wolf is Professor of Philosophy and Political Science at Iowa State University, where he also directs the Bioethics Program. His published research includes papers on intergenerational justice, jurisprudence, intellectual property, social choice, and ethical theory.

PREFACE

In compiling this *Companion to Social and Political Philosophy*, our aim was to combine comprehensive coverage and in-depth analysis. Typically, reference works must choose one or the other, or compromise between the two. The typical encyclopedia provides comprehensive coverage but sacrifices on depth; some recent handbooks go deep into fifteen or twenty issues, but ignore most of social and political philosophy. Routledge and, in particular, our wonderful editor Andrew Beck, has enabled us to avoid this dilemma of handbooks. Our sixty-nine 6,000-word chapters carefully examine almost all facets of social and political philosophy.

Because current political philosophy is so closely tied to the great historical figures, Part I provides studies of the giants of western political philosophy, as well as analyses of nonwestern political thought, something missing in many handbooks. There follows a comprehensive examination of political theories and ideologies, from anarchism through liberalism, republicanism and conservatism to Marxism, environmentalism and feminism. Part III is devoted to a central concern of recent analytic political philosophy—the normative grounding of political thought (contract-centered, consequentialist, perfectionist, pluralist, virtue-based and natural law). The *Companion* next analyzes what is, perhaps, the dominant debate in political philosophy over the last forty years, the nature of distributive justice. Rather than, as in some reference works, having a single chapter on distributive justice, the *Routledge Companion to Social and Political Philosophy* has eight chapters exploring the main proposals discussed today. We then move to another substantive issue that is at the forefront of contemporary political philosophy, the nature of the nation-state and the global order. Again, rather than one summary chapter, the *Companion*'s six chapters explore borders and secession, cosmopolitanism, multiculturalism, global justice, and war.

From Plato onwards political philosophy has been an inquiry into the nature of the concepts that structure our thinking of the social and political world. Having considered justice in such depth in the fourth part, Part VI provides examinations of other core concepts: equality, freedom, autonomy, power, authority/legitimacy, democracy, rights, and toleration. We then move on to studies of different methods of analyzing social and political life, from rational and social choice theories and evolutionary models to pragmatism, discourse theory, and postmodernism. The final part of the *Companion* is devoted to substantive issues that are fundamental to contemporary social and political philosophy, but are often lost in the focus on distributive justice and the state. We are delighted with the range of issues that are examined: education, health, the family, work, punishment, terrorism, paternalism in markets, and religion in public life.

The *Companion*, we believe, is the definitive reference work, for those beginning their study of social and political philosophy, those who are expanding their interest

to new areas and concerns, and those who wish to review current thought in almost any area. We are deeply grateful to Taylor & Francis for their generous support of this unique undertaking and, most of all, to our contributors who have made this such an excellent work.

Gerald Gaus
Fred D'Agostino

Part I

THE HISTORY OF SOCIAL AND POLITICAL THEORY

1
PLATO'S POLITICAL PHILOSOPHY

George Klosko

Plato is generally viewed as one of the greatest and most influential philosophers in the Western tradition. His political philosophy is held in similarly high regard, and is the earliest comprehensive political view we possess. As with other great figures, Plato's political theory was not only part of his overall philosophical system but profoundly shaped by his social and political circumstances. Plato's political experience was one of general decay and decline. In his view, the traditional polis, the main function of which was inculcating its moral values, was under assault from forces of democracy, individualism, and imperialism. Throughout his works Plato defended the traditional polis and argued for the need to pursue values of justice and the intellectual virtues rather than worldly goods such as wealth and reputation.

Plato was born in Athens during the Peloponnesian War, between Athens and Sparta (431–403 BC). He was profoundly affected by political turmoil in Athens after the war ended, which caused him to distrust existing political forms. In his autobiographical Seventh Epistle—the authenticity of which is disputed, though it is accepted by a majority of scholars (see Morrow 1962; Guthrie 1962–81: Vol. V, 401)—he describes his loss of faith in politics and his conclusion that a just regime will come about only when "those who are sincerely and truly lovers of wisdom come into political power, or the rulers of our cities, by the grace of God, learn true philosophy" (Ep. 7, 326a–b). The implications of the union of philosophy and political power are pursued in his most celebrated work, the Republic.

There is great controversy about the interpretation of Plato's political philosophy and numerous unresolved issues. In large part this is because he wrote dialogues rather than treatises, and it is not clear how closely we can identify the views of the main speakers—generally Socrates—as Plato's own. This problem is compounded by disagreements about the dating of different dialogues, the authenticity of others, and the reliability of certain historical evidence, including the *Seventh Epistle*. An additional factor is the radicalism of Plato's proposals and disagreements about his motives in proposing them. While different interpretations cannot be explored in this essay, the reader should be aware that they abound in the voluminous literature. (For discussion of central problems of interpretation, with references, see Klosko 2006a, ch. 2.)

1. Socratic Dialogues

The series of dialogues generally believed to be Plato's earliest vividly depict the character and teaching of Socrates. There are irresolvable problems in establishing the

relationship between the Socrates of these works and the historical Socrates—as Plato understood him—and Plato's own philosophical views. While it is possible that, like a number of other ancient authors, Plato was deeply concerned with depicting the Socrates he knew, it is also possible that he used Socrates as a mouthpiece for his own ideas (at the times he wrote different works). The solution to this "Socratic problem" may well lie somewhere between these poles (see Ross 1933; Guthrie 1962–81: Vol. III; Vlastos 1991: ch. 2). In Plato's later works, the influence of Socrates clearly decreases. In the dialogues that are generally viewed as Plato's last, Socrates assumes only a background role, while he is absent entirely from the Laws, generally viewed as Plato's last work, left unfinished at his death.

Wherever one comes down on the Socratic problem, the early dialogues explore a range of topics that remained central to Plato's political theory. To begin with, Socrates is deeply concerned with the importance of justice. In a series of dialogues, he argues that justice is not only necessary but sufficient for happiness. As he says in the *Apology*, no harm can come to a good man (*Ap.* 41d). Although he could be killed, that is not true harm, as only injustice, which damages the soul, can harm one. In the *Gorgias*, Socrates develops this view, arguing that it is worse to commit injustice (*adikein*) than to have injustice committed against one (*adikeisthai*). But nowhere in this series of dialogues is there a clear account of exactly what justice is and how it contributes to happiness. A series of dialogues feature inquiries into the nature of other moral virtues, several of which end in general *aporia*, or puzzlement. It is not until the *Republic* that questions raised in many of the early dialogues receive comprehensive answers (see Shorey 1903; 1933: 62–73).

Along with justice, Socrates evinces deep concern with a value akin to moral autonomy. As classically recounted in the *Apology*, although Socrates is ignorant about fundamental moral issues, he is still wiser than other people, because he recognizes his own ignorance. Knowing he is ignorant, Socrates searches for moral knowledge, and, through relentless questioning, attempts to awaken his complacent fellow citizens. Socrates compares himself to a gadfly, attempting to rouse a lazy horse by stinging it (*Ap.* 30e–31a). He claims that to talk about virtue is "the highest good for man" and that "the unexamined life is not worth living" (38a). But it is difficult to reconcile Socrates' professed ignorance with his strongly held convictions concerning the importance of justice (Vlastos 1994: ch. 2; Brickhouse and Smith 1996: ch. 2).

More directly political themes are also featured in these works. A prominent theme is Socrates' belief in government by expertise, as opposed to the Athenian system of appointing magistrates by lot. According to both Aristotle (*Rhetoric* 1395b5–8) and Xenophon (*Mem.* I, ii, 9, III, ix, 10–12), the historical Socrates was critical of the Athenian lottery system, which placed crucial matters of governing the state in the hands of people chosen by chance. To this practice, Socrates opposed the ideal of government as a matter of expertise. In the *Protagoras*, Socrates notes the Athenian policy of turning to experts for advice on technical matters, such as naval fortifications, but allowing anyone at all to speak on political questions. If government is a matter of expertise, then possession of the "political art" should be a basis for political authority.

In addition to its lack of qualified leadership, Socrates criticizes the Athenian democracy for being morally corrupt. In his defense speech in the *Apology*, one thing for which he actually apologizes (*apologia* is the Greek word for "defense") is his avoidance of Athenian politics. Socrates pursued his lifelong mission of attempting to turn his fellow citizens towards virtue outside the political system. In his role as gadfly, he takes

his fellow citizens aside individually, "like a father or an elder brother" (*Ap.* 31b). His reason for withdrawing from politics is that, because of the volatility of Athenian politics, someone who attempted to pursue justice within the system could not long survive (32e–3a). In the *Gorgias*, Socrates presents a searing indictment of existing politicians as panderers to the Athenian people. They tell the people what they want to hear rather than what is good for them. They are successful, as pastry cooks would be successful arguing against physicians before a jury of children. As a result, the citizens' appetites have been inflamed, resulting in a city that is swollen and festering with walls and harbors and similar trash (518e–19a). In the *Gorgias*, Socrates begins to describe a different kind of political leader, one who would restrain the citizens' harmful appetites instead of indulging them.

In spite of his criticisms of Athenian democracy, in the *Crito*, Socrates argues for strict compliance with the laws. Having been sentenced to death on charges of impiety and corrupting the young, Socrates refuses to escape from prison, because it would be unjust to do so. The arguments on which he bases this decision constitute the only comprehensive discussion of the question of political obligation in ancient Greek philosophy. Socrates' conclusion is extremely strong, a requirement to obey the law "whatever it commands" (*ha en keleuê*) (51b). However, not only does this conclusion appear to be objectionably authoritarian, but it is difficult to reconcile with Socrates' views expressed in other dialogues, especially the *Apology*, in which he declares that he will obey the command of the gods and continue his mission, more or less regardless of what the Athenian court demands (*Ap.* 29d–30a; see esp. Kraut 1984).

2. The Republic: Defense of Justice

The Republic centers upon an inquiry into the nature of justice. (Dikaiosunê, the Greek term, has broader connotations than "justice"; it is closer to "righteousness" or virtue in general.) In Book I, Socrates asks a series of interlocutors to define justice. He refutes all their efforts, the most notable of which is a famous definition of justice as "the interest of the stronger," put forth by the Sophist, Thrasymachus (see Kerferd 1947; Maguire 1971). Socrates resorts to a series of argumentative tricks to dismiss Thrasymachus. But at the beginning of Book II, Glaucon and Adeimantus (Plato's actual brothers) claim that Thrasymachus gave up too easily and wish to hear more.

In the role of devil's advocate, Glaucon challenges Socrates to identify the nature of justice and explain how it pays. Glaucon develops the first "social contract" argument in the Western tradition. He argues that justice is born of weakness; people set up rules to protect themselves from other people, but would violate the rules if they could. Doing so would allow them to take advantage of others—to steal from them, kill them, etc.—but, unable to violate the rules with impunity, people unwillingly adhere to them. Adeimantus supports Glaucon's claims with appeal to popular opinion, according to which the only reason to be just is fear of getting caught.

Accordingly, Socrates' challenge is not only to define justice but to demonstrate the value of justice itself, without regard to reputation, regardless, in other words, of whether one's justice is known by men or gods. To construct his response, Socrates says that, since the justice of an individual soul is small and so hard to make out, he will observe justice in a larger entity, a hypothetical just city. Having identified justice in the latter, Socrates will demonstrate the superiority of justice to injustice by demonstrating the superiority of the just city to four hypothetical unjust cities. Because justice

is analogous in city and soul, this procedure will allow him to identify justice in the soul and to demonstrate the superiority of justice, on the basis of a comparison between the just soul and four unjust souls that correspond to the four unjust cities.

Although formally subordinate to an inquiry into the nature of justice, Plato's account of the just city is clearly intended as much more, as a vehicle through which Plato can present his political philosophy (cf. Annas 1999; Klosko 2006b). The just city is composed of three classes: rulers and auxiliaries or soldiers who, together, are the city's "guardians," and an unnamed class of craftsmen and farmers. The city's virtues turn on relationships between the classes, with justice defined as an overall principle that each class should stay in its proper place and do its own job.

The soul is analogous to the city in being composed of three parts or aspects: reason, which corresponds to the rulers; the spirited part (*thumos* or *thumoeidês*) corresponding to the auxiliaries, and appetite, which corresponds to the workers. In arguing for parts of the soul, Plato may appear to be relying somewhat arbitrarily on the analogy with the city, but he has strong arguments for his position, based on the phenomenon of psychological conflict. The fact that one can feel urges both to do something and not to do it—e.g., to eat a sweet, and not to eat it, because one is on a diet—indicates separate faculties. Appetite centers on physiologically rooted drives, e.g., for food, drink, and sex. Reason, in contrast, is concerned with the soul's overall good. One reason the analogy between city and soul is particularly apt is that Plato believes that struggle between conflicting urges is settled politically. Thus if reason is able to control the urges of desire, it "rules" in the soul (Kraut 1973). Spirit is a kind of anger that is primarily directed back at oneself. As interpreted by scholars, it centers on anger at oneself for not living up to one's desired conception of oneself (Gosling 1973: ch. 3). In the virtuous soul, spirit is allied with reason, lending reason emotional force to help it keep the appetites in check. The virtues of the soul are analogous to those of the city. As in the just city, justice in the soul is the overall principle that each element stays in its place and does its own job.

The analogy between city and soul is used to demonstrate the superiority of justice to injustice. Only if reason rules in a soul will all elements receive satisfaction. Unlike the other parts, reason recognizes the legitimate needs of the other elements and pursues a balanced plan of life that works to the benefit of the soul as a whole. This is analogous to how the rulers function in the just city, ruling in the interests of all classes, not their own. Looking at the city as a whole, we can see the superiority of the balanced, harmonious just city to the four unjust cities—timarchy, oligarchy, democracy, and tyranny—in each of which the rulers rule by force. The fruits of injustice are most clear in the tyrannic city, in which the ruler oppresses and impoverishes his fellow citizens solely for his own interest. A similar comparison holds in regard to the just soul and four corresponding unjust souls. Like the just city, the just soul is balanced and harmonious; its health contrasts with the anguish of disharmony and conflict between elements that characterize the unjust souls. Once again, the contrast is most apparent in regard to the tyrannic soul. Possessed by an overriding lust, this soul sacrifices all other concerns in a vain attempt to satisfy this base desire.

However, even if we accept Plato's argument for the superiority of justice, it is not clear that this is an adequate response to Glaucon and Adeimantus. In making their case for injustice, Plato's brothers clearly describe injustice as concerned with people's conduct. As indicated above, injustice is epitomized by the tyrant, who kills whomever he pleases, steals from whomever he pleases, and commits other actions that are gener-

ally forbidden. In responding, Plato transforms justice into a quality of the soul. It does not concern a person's conduct, but internal psychic conditions. Accordingly, a notorious difficulty in the *Republic* concerns the need to connect up these internal and external aspects of justice, to establish, in other words, that the performance of unjust actions will disturb the balance and harmony of the soul that are necessary for happiness (Sachs 1971). However, although Plato does not address this problem directly, it is likely that a suitable response can be constructed from the materials he provides. Especially important is his basic assumption that conduct strongly affects psychic qualities. As we will see directly, the fact that, especially in early childhood, one's soul is shaped by how one behaves is central to his political theory.

3. The Republic: The Just City

The essence of Plato's political theory is straightforward. The purpose of the just city is to make its inhabitants as virtuous as possible, and so everything possible is done to achieve this. The result is a completely controlled environment in which all resources are devoted to the inculcation of virtue. The city's primary institution is the education system. Its other institutions, including Plato's famous system of communism, are intended to provide an environment in which the task of education can be successfully carried on.

As noted above, the just city is composed of the three classes: the rulers, auxiliaries, and farmers. Underlying the structure of the city and its educational function are two basic psychological assumptions. First, Plato believes that people are largely malleable. A person is strongly affected by the environment in which he or she grows up, and so can be made virtuous if brought up in a properly governed city. The second assumption limits malleability. Plato believes there are fundamental, innate differences between people. The three different kinds of people—those with gold, silver, or bronze in their souls (Rep. 415a–d)—have vastly different capacities to achieve virtue. Although the just city is designed to raise people to the highest levels of virtue they are capable of, the result must be different classes with different levels of virtue.

Plato is unusual in the history of political philosophy in the amount of emphasis he gives to the psychological effects of art, which is a central theme in the *Republic*. In ancient Greece, poetry was the primary artistic medium, and so poetry, especially Homer, receives detailed criticism, in regard to proper rhythms and meter, as well as content. Similar concerns with rhythm and harmony apply to music and the visual arts, though Plato does not discuss the last subject in detail. He believes the arts are most effective with the very young, instilling balance and harmony, readying children for true moral principles when they are old enough to understand them. As a result of lifelong, intensive education, all inhabitants are raised to the highest level of virtue possible for them. But because of the inherent incapacities of most people, this requires that members of the lower classes be completely subordinate—Plato describes them as "slaves" (590d)—to the rulers.

The just city's major institutions are intended to facilitate education. As Plato famously argues, the city should be ruled by philosophers. A just city is not possible until political power and philosophic wisdom are joined in the same hands. The philosophers' claim to rule rests on their apprehension of the Forms—perfect moral standards that exist outside the phenomenal world and are the only things that are truly real. This divine knowledge orients the philosophers' desires away from worldly things, so they

can be trusted to rule completely unchecked, although Plato does not provide a detailed account of how their knowledge actually helps in the business of ruling. The nature of the philosophers' knowledge and of the Forms themselves is illustrated in three famous images: the Form of the Good as the sun; the "divided line," which elaborates on different cognitive states and corresponding entities at different levels of existence; and the cave. The details of these images, and of the theory of Forms more generally, cannot be discussed in this essay, although, as is especially clear in the image of the cave, Plato's metaphysical and epistemological views are the essential underpinning of his political philosophy (see below).

The other distinctive institutions in the just city are the treatment of women and communism. In regard to the former, Plato argues that gender differences are not relevant to ruling. He is a pioneer in the history of women's equality in arguing that women who are qualified to rule should be allowed to do so. Accordingly, the common term "philosopher-kings" is actually misleading in not taking account of philosopher-queens. However, equal treatment is only for exceptional women. Outside the class of guardians, conventional family structures are apparently maintained.

Even more radical is the system of living arrangements for guardians, men and women. For them, the nuclear family is abolished. Children are raised communally, and with reproduction controlled by the state, the stock of guardians can be improved by selective breeding. Abolition of the family also makes the guardians, in effect, one big family, thereby eliminating causes of faction and promoting stability in the state.

Community of family receives much more detailed discussion than community of property. The guardians live a permanent barracks existence, as soldiers in camp, with common mess. They are allowed to possess only their weapons, are forbidden to own gold and silver, and are forbidden houses or storerooms which anyone who wishes is not free to enter. The city's land is owned by members of the lowest class, while the guardians are maintained by an annual tax upon this class, which they receive as a salary for protecting them.

Although it is often said that Plato's just city is communistic, this description is inaccurate. Communism (or socialism) is generally described as public ownership of the means of production. In Plato's city, the primary means of production, land, is owned by the lowest class. There is an element of communism in the city, as the guardians live communally without individual possessions. But this is community of poverty rather than community of property.

Through these arrangements, the just city is cleverly constructed in order to give each class what it wants. This is essential for the virtue of temperance, that each class willingly stays in its place, thereby allowing the city to remain just. The farmers, who are oriented towards the satisfaction of their appetites, control the state's wealth. Since moving up to the ruling class would require them to renounce property, they should prefer to remain where they are. Something similar is true of the auxiliaries, who are honored for their military service, thereby receiving what they desire. The philosophers, on the other hand, are to some extent forced to set aside their own desires. Although they would prefer to focus on their studies, they must serve as rulers, not for their own benefit but for the sake of their fellow citizens. Plato holds that a city will be able to prosper only if its rulers rule unwillingly, because there is no way they can profit from ruling. But in the final analysis, Plato argues that, living in a secure environment, freed from the cares of ordinary life, and honored by their fellow citizens, the rulers are as happy as Olympic victors (465b–e).

It is difficult to know how seriously Plato takes his sketch of an ideal city. Challenged by his interlocutors, Socrates devotes considerable effort to describing how the city could be realized. Though this would be difficult, it is not impossible (*Rep.* 501a–502c; Klosko 1981; Burnyeat 1992). But the city is later described as existing only "in theory," and Plato suggests that it does not matter whether it is ever realized (592a–b). Prominent scholars argue that the city is so extreme that it is an ingenious satire, intended to demonstrate the impossibility of political reform (Strauss 1964; Bloom 1968; criticized in Burnyeat 1985; Klosko 1986).

Whether or not the just city is intended to be realized, the *Republic* develops Plato's critique of Athenian democracy (cf. Monoson 2000). In the famous parable of the ship of state (488a–89d), the populace show no respect for those who truly understand navigation, whom they view as odd, turning over control of the ship to charlatans who only pretend to knowledge. Plato also argues that the mob is emotional and easily misled. Sophists are likened to animal trainers, who have mastered the art of managing the unruly beast of the mob (493a–c). To these somewhat familiar themes, Plato adds a wonderful satirical sketch of democracy in Book VIII, which criticizes democracy for treating unequals alike. This holds for the old and the young, men and women, masters and slaves, even animals and human beings.

Central to Plato's critique of democracy is belief in objective values, which only a few have the capacity to know. To follow other values, is simply to be wrong, and so the expertise of the few must take precedence over the wishes of the many. Accordingly, in addition to educative institutions that make the many as virtuous as possible, a system of higher education is intended to create a class of political experts. As there are relatively well-established means to train experts in the different crafts that Socrates frequently discusses in the early dialogues—shoemaking, carpentry, etc.—Plato describes a similar system to train future philosophers. Their curriculum centers on mathematics, which raises their minds from the material world to the abstract intellectual world, and finally to the Forms, which they study for five years. Then the philosophers spend fifteen years helping to govern the city, and so gain practical experience, before they are raised to knowledge of the Form of the Good.

The central orientation of Plato's political philosophy is illustrated by the cave image (514a–17a). Prisoners in the cave are bound so they can see only shadows that pass on a wall in front of them and contend for preeminence in this shadow world. If they could escape and see the world outside the cave, they would despise their former situation. However, when a prisoner who had been freed from the cave, goes back down to free his former fellows, he is disoriented by returning to darkness. He appears foolish to his former fellows, who resist his efforts and would kill him, if they could.

Because the prisoners inhabit a world that is not fully real, to attain knowledge and true values, they must move beyond appearances and out of the cave. Notable in this image is complete denial of the legitimacy of ordinary people's desires—and of the appetitive life, which Plato views with disdain. To be freed from the chains of appetite people must be educated in a properly run state. Once again, this requires complete subordination of the lower classes to the philosophers.

In the aftermath of twentieth-century fascist and totalitarian regimes, Plato received harsh criticism, because of resemblances between the just city and these systems. The most celebrated attack is by Karl Popper, in *The Open Society and Its Enemies* (Popper 1966). Although many of Popper's specific charges are wide of the mark and his work as a whole is wildly polemical, there is a clear element of truth to the general charge

(cf. Levinson 1953; Klosko 1996). Because of his overall lack of faith in most people's ability to achieve virtue on their own, Plato believes they must live in permanent subjection to the philosophers. At one point, Plato says that after people have been properly educated, they will become free (591a). But it is not clear what this freedom consists of in the totally controlled, totally censored just city. Popper's critique, along with similar criticisms from other scholars (Crossman 1939; Winspear 1940), raised additional related issues, including whether Plato is a racist (Klosko 1991), and possibly statist implications of the principles of distributive justice advocated in the *Republic* (Vlastos 1977).

4. Plato's Later Political Theory

In his later dialogues, Plato's political theory moves in directions of increased realism and pessimism. This may be connected to his incursions into Syracuse and his unsuccessful attempt to win over Dionysius II, the city's tyrant, to philosophy. Though Plato was not optimistic, he decided to go to Syracuse, lest he be "a pure theorist, unwilling to touch any practical task" (Ep. 7, 328c). Dionysius' commitment to philosophy proved to be weak, while Plato and his followers were implicated in the subsequent disastrous invasion of Sicily undertaken by Dion, Plato's close associate (see Morrow 1962).

The most important late political dialogues are the Statesman and the Laws. In both works, dramatic action is severely reduced. Though the dialogue form is retained, it is largely a shell for authoritative disquisitions by the works' main characters, who in neither work include Socrates (cf. Gill and McCabe 1996). In the Statesman, Socrates remains in the background while an unnamed Eleatic Stranger discusses the nature of the statesman—as he does in regard to the nature of the sophist, in a sister dialogue. In the Laws, from which Socrates is absent, the main character is an unnamed Athenian Stranger. These two works maintain Plato's concerns with government by expertise and inculcation of virtue. But in both works, Plato turns from the intense idealism of the Republic toward greater appreciation of laws and existing institutions.

The *Statesman*'s concern with laws shows up in the Eleatic Stranger's argument that the statesman's art should not be constrained by law. If a physician had to be away from his patients, he would leave instructions for them. But when he returns, his expertise takes precedence over his previous instructions, which may be set aside. Something similar holds in regard to laws. Because subjects of laws are numerous and differ in many respects, proper instructions for them must be general, not tailored to specific characteristics of individuals. As with the physician, the true statesman may disregard the generality of laws in order directly to apply his expertise to each case. However, if a true statesman is not available, laws should be adhered to. Although the text is compressed and at times obscure, it also appears that laws should remain unchanged. Laws are necessary to restrain the populace, while Plato suggests that attempts to change them will invariably be corrupted. Although law is no substitute for political expertise, it is a "second best" (297d, 297e, 300c), and Plato presents a classification of constitutions with good states distinguished from bad on the basis of whether or not the rulers adhere to the laws.

Plato's interest in laws is developed in the *Laws*. The participants in this dialogue are three old men, an unnamed Athenian Stranger, a Spartan, Megillus, and Clenias, a Cretan. As the three walk to a shrine on the island of Crete, they discuss the construction of a city, Magnesia, which is going to be founded as a colony of Crete. The *Laws* displays deep knowledge of the laws and institutions of many Greek cities, which

is obviously the fruits of years of research conducted by Plato and his school (Morrow 1960). Unlike the just city in the *Republic*, which is designed more or less freely from the ground up, the "second best" city in the *Laws* (739a–e, 875c–d) proceeds from existing political and social institutions, which Plato improves (see the invaluable discussion in Morrow 1960). Plato's basic plan combines a system of political institutions similar to those of democratic Athens and a social and economic system like that of Sparta. Overall control is vested in the laws, which are described in enormous detail. They are drawn up by a philosophical lawgiver, and, after an initial try-out period, will be virtually impossible to change. As the Athenian says in regard to laws concerning education, his aim is to emulate Egypt, where laws concerning music had not changed in thousands of years (656d–57a).

A decisive element in the movement from the *Republic* to the *Laws* is Plato's loss of faith in the possibility of philosopher-kings. Even if someone could understand what is necessary to do, "human nature" would impel him towards self-aggrandizement, thereby bringing ruin to the city (*Laws* 875a-d). Because no human being can be trusted with unaccountable power, only the rule of law allows human beings to rise above the level of savage beasts (874e–75a). Along with his abandonment of the philosopher-king, Plato appears to place more reliance in the abilities of more ordinary people (Bobonich 2002). All citizens are enfranchised politically, and all live in traditional families. In Magnesia, there are no guardians living apart from and over society as a whole. However, although Plato's attitude towards democracy in the *Laws* is undoubtedly more favorable than in the *Republic*, this still falls short of modern views. Virtually the only prerogative of citizens is to help choose their magistrates, while magistrates too do not determine the state's course. This is prescribed for all time by the laws.

The political system of the *Laws* is purged of the excesses of Athenian democracy. As in Athens, important institutions are an Assembly and popular courts. But the former is assigned fewer functions than in Athens, while the power of popular courts is curbed by a system for appeals. Magnesia departs notably from Athens in its large array of magistrates with important functions. While the Athenians generally appointed magistrates through the lottery, in Magnesia, most are appointed through a combination of election and the lot. Significant officials include "guardians of the laws," who are assigned a host of duties, a board of examiners, and a "nocturnal council."

This last body is composed of high office holders throughout the state. Its ominous sounding name comes from its meeting in the hours before dawn, a convenient time for busy officials. It is described as a kind of ongoing seminar to study the laws of different cities, with an eye to preserving those of Magnesia. Having described this council, the Athenian concludes the *Laws* by saying that the state must be "turned over to it" (969b). Although this may sound like a return to philosopher-kings—which is apparently Aristotle's interpretation (*Pol.* 1265a3–4)—this would contradict many elements of the *Laws*. Glenn Morrow, the most important scholar of the *Laws*, argues that the council exercise only informal authority, providing advice when it is necessary to change laws. However, on this construal, the apparent importance of the council does not rest well with the rigidity of the laws and the extreme difficulty of changing them. In regard to the nocturnal council, as other aspects of the *Laws*, it is possible that Plato had not fully worked out the details at the time of his death (see Morrow 1960: ch. 9; Klosko 2008).

In regard to economic arrangements, the Athenian argues that while community of property is best, it is possible only for gods or children of gods, and so the second

best state has private property (739a–e). But as in Sparta, land held by each family is inalienable, and citizens dine communally. As in Sparta, ownership of gold and silver is forbidden, while excessive wealth is confiscated by the state. While the traditional family structure is retained, women are to be educated along with men, to provide military service, and to be eligible for public office—although Plato provides few details.

Because their farms are worked by slaves, most citizens have the leisure necessary for pursuit of virtue. This is their central occupation, and the day is not long enough for all the activities this requires (806d–08c). As in the *Republic*, the need to inculcate virtue is the state's main business. Poetry and other arts are strictly censored. Magnesia has an intensive system of education, the main emphasis of which is on conditioning, education of pleasures and pains. An indication of how extreme Plato's views have become is that the process of conditioning must begin before birth, as pregnant women are to do rhythmic exercises, to make the young balanced and harmonious (789a–e).

Religious elements of Magnesia receive considerable emphasis, arguably taking over the role of philosophy in the just city of the *Republic*. In Book X, the Athenian develops elaborate—though philosophically weak—arguments against common heresies. He proves the existence of the gods, that the gods involve themselves in human affairs, and cannot be bought off with bribes. Although—or perhaps, because—the arguments are so weak, adherence to them is to be enforced through a kind of inquisition. People who do not believe in them are sentenced to death, even if their conduct is otherwise exemplary. Thus Plato has the dubious honor of being one of the first thinkers in the Western tradition to advocate death for thought-crime.

The distance Plato's political philosophy has traveled from its Socratic roots is apparent in the *Laws* (esp. Gould 1972). Changes that appear in the *Republic* are more pronounced in the *Laws*. In general, the movement of Plato's thought is from the Socratic ideal of moral autonomy to society-wide, intensive conditioning, beginning before birth. Amidst this overall development throughout Plato's career, there are also elements of continuity, especially continued emphasis on virtue and government by expertise. To realize this ideal, the *Republic* argues that philosophy and political power must be combined in the hands of the philosopher-king. At the end of his life, having lost faith in the possibility of philosophic rulers, in the *Laws*, Plato focuses on a second-best form of philosophic politics: adherence to unchanging laws constructed by a philosophic lawgiver.

Related Topics

Aristotle's Social and Political Philosophy, Democracy, Virtue Ethics and Political Philosophy

References

Quotations of Plato are from J. M. Cooper, ed., *Plato: Complete Works*. (Indianapolis: Hackett, 1997); translations occasionally modified slightly. A full list of the standard definitions used in this essay can be found in various reference works, including Liddell-Scott-Jones' Greek Lexicon.

Annas, J. (1999). *Platonic Ethics, Old and New*. Ithaca, NY: Cornell University Press.
Bloom, A. (ed. and trans.) (1968). *The Republic of Plato*. New York: Basic Books.
Bobonich, C. (2002). *Plato's Utopia Recast*. Oxford: Oxford University Press.
Brickhouse, T. and N. Smith. (1996). *Plato's Socrates*. Oxford: Oxford University Press.
Burnyeat, M. (1985). "Sphinx Without a Secret," *New York Review of Books* 32, #9 (May 30, 1985).

——. (1992). "Utopia and Fantasy: The Practicability of Plato's Ideally Just City," in J. Hopkins and A. Savile (eds.), *Psychoanalysis, Mind, and Art*. Oxford: Oxford University Press.
Crossman, R. S. (1939). *Plato Today*. New York: Oxford University Press.
Gill, C. and M. McCabe, eds. (1996). *Form and Argument in Late Plato*. Oxford: Oxford University Press.
Gosling, J. C. (1973). *Plato*. London: Routledge.
Gould, J. (1972). *The Development of Plato's Ethics*. 1955; rpt. New York: Russell and Russell.
Guthrie, W. K. C.(1962–81). *A History of Greek Philosophy*. 6 vols. Cambridge: Cambridge University Press (vols 4–5).
Kerferd, G. B. (1947). "The Doctrine of Thrasymachus in Plato's *Republic*," *Durham University Journal* 9: 19–27.
Klosko, G. (1981). "Implementing the Ideal State," *Journal of Politics* 43: 368–89.
——. (1986). "The 'Straussian' Interpretation of Plato's *Republic*." *History of Political Thought* 7: 275–93.
——. (1991). "'Racism' in Plato's *Republic*," *History of Political Thought* 12: 1–13.
——. (1996). "Popper's Plato: An Assessment," *Philosophy of Social Science* 26: 509–27.
——. (2006a). *The Development of Plato's Political Theory*. 2nd. ed. Oxford: Oxford University Press.
——. (2006b). "Politics and Method in Plato's Political Theory," *Polis* 23: 328–49.
——. (2008). "Knowledge and Law in Plato's *Laws*," *Political Studies* 56: 456–74.
Kraut, R. (1973). "Reason and Justice in Plato's *Republic*," in E. N. Lee *et al.* (eds.), *Exegesis and Argument*. Assen: Van Gorcum, pp. 207–24.
——. (1984). *Socrates and the State*. Princeton, NJ: Princeton University Press.
Levinson, R. B. (1953). *In Defense of Plato*. Cambridge, MA: Harvard University Press.
Maguire, J. (1971). "Thrasymachus … or Plato," *Phronesis* 16: 142–63.
Monoson, S. (2000). *Plato's Democratic Entanglements: Athenian Politics and the Practice of Philosophy*. Princeton, NJ: Princeton University Press.
Morrow, G. (1960). *Plato's Cretan City*. Princeton, NJ: Princeton University Press.
——, ed. and trans. (1962). *Plato's Epistles*. Revised ed. Indianapolis: Library of Liberal Arts.
Popper, K. (1966). *The Open Society and Its Enemies*. Vol. I: *The Spell of Plato*. 5th ed. Princeton, NJ: Princeton University Press.
Ross, W. D. (1933). "The Problem of Socrates," *Proceedings of the Classical Association* 30.
Sachs, D. (1971). "A Fallacy in Plato's *Republic*," in G. Vlastos (ed.), *Plato II*. Garden City, NY: Anchor Books, pp. 35–51.
Shorey, P. (1903). *The Unity of Plato's Thought*. Chicago, IL: University of Chicago Press.
——. (1933). *What Plato Said*. Chicago, IL: University of Chicago Press.
Strauss, L. (1964). *The City and Man*. Chicago, IL: University of Chicago Press.
Vlastos, G. (1977). "The Theory of Social Justice in the *Polis* in Plato's *Republic*," in H. North (ed.), *Interpretations of Plato*. Leiden: Brill.
——. (1991). *Socrates: Ironist and Moral Philosopher*. Ithaca, NY: Cornell University Press.
——. (1994). *Socratic Studies*. M. Burnyeat (ed.), Cambridge: Cambridge University Press.
Winspear, A. D. (1940). *The Genesis of Plato's Thought*. New York: Dryden Press.

Further Reading

For further reading, see: J. Adam (ed.), *The Republic of Plato*, 2 vols. (Cambridge: Cambridge University Press, 1902) is a classic Greek edition, with valuable commentary. Volumes IV and V of W. K. C. Guthrie, *A History of Greek Philosophy*, 6 vols. (Cambridge: Cambridge University Press, 1962–81), provide judicious commentary, dialogue by dialogue. Vols. 2 and 3 of P. Friedlander, *Plato*, 3 vols., H. Meyerhoff (trans.) (Princeton: Princeton University Press, 1958–69), provide a dialogue-by-dialogue commentary, with more of a literary focus, while Vol. 1 is an introductory volume on central themes. T. Irwin, *Plato's Ethics* (Oxford: Oxford University Press, 1995), is a valuable account of Plato's moral philosophy. M. Schofield, *Plato: Political Philosophy* (Oxford: Oxford University Press, 2006), is a recent thematic discussion, by an eminent classicist. As its title indicates, G. Klosko, *The Development of Plato's Political Theory*. 2nd ed. (Oxford: Oxford University Press, 2006), examines the movement of Plato's political theory throughout his career. M. Lane, *Method and Politics in Plato's Statesman* (Cambridge: Cambridge University Press, 1998), is the best recent book on the Statesman. G. Morrow's *Plato's Cretan City* (Princeton: Princeton University Press, 1960), is an indispensable historical analysis of the Laws. C. Bobonich, *Plato's Utopia Recast* (Oxford: Oxford University Press, 2002), is a brilliant, though iconoclastic analysis of central philosophical themes in the Laws.

2
ARISTOTLE'S SOCIAL AND POLITICAL PHILOSOPHY

Rachana Kamtekar

Aristotle (384–322 BCE) came to Athens as a young man to study in Plato's Academy. Upon Plato's death nearly twenty years later, Aristotle left Athens to spend time in Asia Minor and in Lesbos, returning in 343 BCE to his home in Macedonia. In 335 BCE he went back to Athens to set up his own school in the Lyceum, where he remained until the death of Alexander (323 BCE) unleashed anti-Macedonian sentiments in Athens and he was charged with impiety. He fled to nearby Chalcis where he died about a year later. Most of the works Aristotle wrote for publication are lost; what survive are later compilations of works written for use within his school. Aristotle's most influential political ideas connect human nature and its flourishing with political activity, ideally under a constitution in which virtuous citizens take turns at ruling and being ruled.

This essay falls into three parts. Section 1 guides readers through the first book of the *Politics* to acquaint them with the major concerns of Aristotle's political philosophy. Section 2 focuses on Aristotle's famous claims that the human being is, by nature, a political animal, that the *polis* (city-state) is natural, and that the *polis* is naturally prior to the individual, who belongs to the *polis*. (Throughout, I will use the Greek words relating to *polis* (e.g., *politeia*, *politikos*) rather than their English translations as I think this will bring out the connections in Aristotle's political thought that are obscured by the rather different associations of their English-language counterparts (constitution, statesman).) Finally, Section 3 examines Aristotle's ideal *politeia* (constitution) in the light of his criticisms of other ideal and existing *politeiai*.

1. Circumscribing and Valuing the Political

Near the beginning of the Nicomachean Ethics Aristotle says that the capacity that aims at the good, that for the sake of which we do everything that we do, is political science, and that political science is the "most authoritative" of the capacities and uses all the others (1.2 1094a18–b12). So when at the end of the *Nicomachean Ethics* Aristotle directs the student of politics to his writings on legislation (10.9 1180b29–81b22), the logical place to go is Aristotle's *Politics*, which discusses politeiai and their laws from Books 2 through 8. Unfortunately, the structure of the *Politics* is so unclear that it has

been described as "a book with great defects" (Robinson 1962: viii), and "a loosely connected set of essays on various topics in political philosophy ... [perhaps] put together ... by an editor ... [which] Aristotle may never have intended to form a single treatise" (Keyt and Miller 1991: 3). Indeed, the first book of the *Politics* does not even mention legislation, the topic which sent the reader of the *Nicomachean Ethics* to it. Further, in *Politics* 1 Aristotle devotes a number of chapters to defending his notorious theory of natural slavery: that some human beings, namely non-Greeks, are natural slaves and ought to be ruled by natural rulers as the body ought to be ruled by the soul (1252a30–b9, cf. 1254a14–1255a3). Not only is the view odious, it also seems unconnected to the topic of legislation. All this makes for a discouraging beginning.

Now a student of philosophy who is interested in ideas rather than texts might decide to ignore the false or philosophically uninteresting parts of any body of work. But we can also take a more historical approach and try to determine how the ideas, promising as well as objectionable, fit together into a larger project (are they central philosophical commitments, or do they follow from such commitments? Or are they incidental to, or even in conflict with, the main ideas?). Taking such an approach can help to explain not only why Aristotle is so concerned with slavery in *Politics* 1, but also why he might have thought *Politics* 1 to be a suitable introduction to his work on legislation.

The second paragraph of *Politics* 1 (1252a9–17, cf. 1253b15–23) describes a view against which Aristotle will spend the rest of the book arguing, namely, that the expertises of the *politikos* (statesman), king, household-manager and master of a slave are one and the same expertise, *politikê* (the art of politics). This is a claim argued for in Plato's *Statesman* (258d–59e), and Aristotle argues against it to show that there is something distinctive, and distinctively valuable, about practicing *politikê*, so that the political life is a good life for human beings. By contrast, Plato's *Republic* treats political expertise as a byproduct of the theoretical wisdom of philosophers, and political rule as a burden that falls on philosophers as a consequence (346e–47d, 519d–21b).

Aristotle argues that there is a difference between the expertises of the *politikos*, the king, the householder and the master, on the grounds that in each case the character of the one who is, by nature, ruled is different (1260a3), and different enough that ruling such a character requires a different expertise. So he emphasizes the way in which a slave is by nature different from a free person: the slave "by nature belongs to the master" (1254a10–16); the slave is to his master as the body is to the soul (1254b15–19); the slave can comprehend, but not originate, orders (1254b22); the bodies of natural slaves are strong for servile labor, whereas the bodies of free men are upright (1254b26). All this makes the appropriate way to rule over a slave different from the appropriate way to rule over a free person (1255b16–20). Arguing for this difference opens the door to showing how political rule, rule over a free person, is something finer than the rule over a slave, which Aristotle likens to the use of a tool (for a slave is a living tool). Aristotle's overall aim in *Politics* 1 is captured in his slogan, "the rule is always better when the ruled are better" (1254a25).

Plato's *Republic* had addressed the worry that actual forms of rule are all practiced to benefit the rulers, by stipulating a *politeia* whose laws aimed at benefiting the ruled and then exploring what that would be like. Highlighting the distinctive value of political rule requires Aristotle to contrast it not only with exploitative rule, in which the ruler benefits at the expense of the subjects, but also with the best instances of other kinds of rule. The case of natural slavery, in which a natural superior rules a natural inferior in such a way as to benefit both master and slave, is useful to bring this out. This might be

why Aristotle forces an opposition between the view that *politikê* is the same skill as that of the master of slaves and the view that slavery is contrary to nature (1253a3–12). The views need not be opposed; for example, a skeptic about *politikê* might maintain both, but Aristotle's discussion allows him to show that the superiority of political rule is not only a superiority to the kind of slavery that is forced on someone who could live freely, but also to the best kind of slavery, which is actually better for the slave than freedom. The contrast between political rule and the head of household's rule over his family makes Aristotle's case even more pointedly since, there, not only do the ruled benefit in fact, but their benefit is the aim of the rule.

Today we are likely to object that few people are better off being told what to do throughout their lives, and that there is something wrong with using such people as living tools. And we are likely to discount as the product of society the supposedly natural differences Aristotle finds between men and women. Modern readers will find least off-putting the way Aristotle makes his point by the contrast with the rule over domestic animals (1254b10–12). Certainly a good dog-owner makes the life of his dog much better than it would be if the dog had to fend for itself, and this is compatible with the owner using the dog for security (for example); however, the relationship between a person and his dog is not thought to be as valuable as the relationship between one person and another.

Aristotle's case for the distinctness and superiority of political rule to even other forms of mutually beneficial rule acquires more positive content from two other claims central to *Politics* 1: that as the household is subordinate to the *polis*, economics is subordinate to politics, and that humans are by nature political and the *polis* "is by nature." Let us examine each of these claims.

First, in *Politics* 1, Aristotle claims that the institutions ruled by the non-political kinds of rule (household, village) are parts of the *polis* (1.2 1252a24–1.3 1253b3). Now, when Aristotle says that the household is "part of" the *polis*, he is saying something much more substantive than that households are found in *poleis*. He is saying that the *polis* accomplishes, to the fullest extent, the purpose for the sake of which the household is established: the *polis* makes possible not only living, but living well (1252b29–30). His account of the evolution of the *polis* out of the household and village (1.2) is meant to demonstrate this. (Aristotle's claim also marks another disagreement with Plato's *Republic*, which abolishes the household among the guardians, and treats either individuals or the classes defined by political function as the proper parts of the city.) A consequence of the household's belonging to the *polis* seems to be that the expertise of ruling a household or clan is subordinated to *politikê*. And this recalls the *Nicomachean Ethics*' identification of *politikê* as the most authoritative expertise, the one that makes use of the general's expertise, economics, and so on (1.2 1094a26–b7).

Economics is the expertise of managing the household (*oikos*). The household, comprising the family and its property, is the institution that aims at obtaining and using the food, shelter, etc., required for living. Even though economics is not his subject, Aristotle pauses to note that the art of obtaining is part of the art of household management, and that its limit is set by the use that household management and politics make of the wealth it creates (1256b–39). He goes on to distinguish the use value of a thing (the use one makes of it in virtue of its nature) from the exchange value (the use one makes of it in exchange for other goods) and to offer a speculative history of money and trade. Meikle (1995) argues that here and in *Nicomachean Ethics* 5.5 where Aristotle argues that exchange must be based on proportional equality, Aristotle is only able to formu-

late the problem of relating use value to exchange value but not to give an account of their relationship. Meikle explains this inability as due to Aristotle's metaphysical view that there are many irreducible senses of being and of good. On the other hand, it seems that the notion of use value has already reduced the many senses of good to one kind of value, and that money then facilitates the conversion of use into exchange value. But in either case, Aristotle does not think a mechanism such as money or price could be a basis for just exchange.

In what sense is politics for the sake of living well (as opposed to merely living, like economics)? This brings us to the second of our remaining arguments for giving content to the superiority of *politikê*. Aristotle says that human beings are by nature political. Evidence of this is that we desire to live together even when we have no need of help from one another (cf. *Politics* 3.6 1278b20–22). Because we are by nature political, friendship belongs to the happiest life (*Nicomachean Ethics* 9.9 1169b18). Here again, we see a departure from Plato, who claims that a *polis* comes to be because we are not self-sufficient but need each other (*Republic* 369b). Aristotle agrees that this is why *poleis* originate or "come to be," but thinks that they now exist also for the sake of living well. In *Nicomachean Ethics* (1.7 1097b25–1098a17) Aristotle argues that human happiness consists in the excellent exercise of our distinctive capacity for reasoning, i.e., in activities expressing the virtues of justice, moderation, courage, and wisdom. Insofar as political activity expresses these virtues, it enables us to live well in two ways: we make our communities better, and we ourselves engage in an activity that fulfills our nature.

What activity counts as political activity? According to Aristotle's account of the *politês* (citizen), who is distinguished by having a nonrestricted share in judgment and office (1275a22–24, 1275b18–20), and whose function it is to preserve the *politeia* (1276b28–30), it would seem to be participating in government (rather than, say, opposing it, trying to change it, or discussing individual and social virtue, as did Socrates). Only *politeiai* that aim at the common good and apportion privilege in accordance with merit are correct, and as a result the virtue of a citizen—preserving the *politeia*—and the virtue of a human being (the excellence of the soul's rational capacities) sometimes come apart (1276b34–36). So political activity will not in all circumstances be conducive to human virtue and happiness. Rather, it is participating in judgment and office, under a correct *politeia*, that is a particularly worthy activity and that, when engaged in virtuously, makes for a happy life.

Finally, what is the connection between our natural desire for society and the particular form our association is supposed to take for living well: why is it judging and holding office, rather than, say, playing sports and going to the theater together? The *Politics* answers by giving a specification of the *Nicomachean Ethics*' idea that happiness consists in the excellent exercise of our rational capacities. Aristotle points to our sense of justice and our capacity for speech: human beings have a sense of good and bad, just and unjust, and we can communicate what is beneficial or harmful and just or unjust, in speech (1.2 1253a8–18). Thus it is our rationality in the sense of our capacities to value and to communicate value that are our distinctive capacities, and we achieve our own distinctive good by improving and then exercising them, which judging and holding office give us the opportunity to do. In his (1991) "The Connection between Aristotle's *Ethics* and *Politics*," Arthur Adkins argues that the *Nicomachean Ethics*' highly formal account of virtue in terms of the human function of reasoning is only given content by the *Politics*' account of the best *politês* in the best *politeia*—so that human virtue turns out to be an excellent condition for a leisured male. But although Aristotle thinks that judging and

holding office belong to the leisured male "by nature," it is not because of the content of judging and holding office that this is so, but because of Aristotle's false assumptions about which kinds of people have the ability to engage in judging and holding office.

Let us now take a closer look at Aristotle's ideas about human nature and the nature of the *polis*.

2. Nature and the *Polis*

Aristotle says that the *polis* exists by nature (1253a1–2), that human beings are, by nature, political animals (1253a2–3), and that the *polis* is prior to the individual (1253a18–19, 25–26). These claims are important because of their normative implications: for Aristotle, to say that something is "by nature" is not only to say that it is usual, but also that it is connected with some good end. Consequently, students of Aristotle have tried to state precisely what Aristotle means when he makes these claims.

In his classic paper, "Three Fundamental Theorems in Aristotle's *Politics*," Keyt (1987/1991) draws on Aristotle's sense of "by nature" from *Physics* 2.1, where it contrasts with "by reason" or "by craft" on the one hand, and "by chance" or "by luck" on the other, to argue that Aristotle fails to establish his naturalness claims. Aristotle describes the *politikos* as the craftsman of *polis* (*Politics* 7.4 1325b40–1236a5), but of course, as a natural thing, the *polis* should not be what it is as a result of craft, but of an internal principle of motion in virtue of the thing that it is. Natural things come to be either by a parent reproducing its form as do plants and animals, or by chance as do certain parasites, or are eternal, like the heavenly bodies. The *polis*, however, evolved from the household and village. This leads one to suspect that the *polis* comes to be by reason and art, rather than by nature. Keyt identifies and criticizes four arguments in the *Politics* involving the claim that the *polis* exists by nature in *Politics* 1.2: (1) At 1252b27–34, Aristotle argues that the *polis*, which is an association of villages, is by nature, for the reason that it has achieved the limit of self-sufficiency, the aim of the earlier associations. Keyt identifies as the basis of this argument the principle of transitivity of naturalness: since the associations out of which the *polis* grew (from the necessary pairings of man–woman and master–slave through household and village) are natural, so is their product. But this principle is false, since many artificial things, such as poems, are the products of natural impulses. (2) At 1253a1–4, Aristotle argues that since the *polis* exists for the sake of self-sufficiency, which is best, humans are, by nature, political. Keyt says this argument depends on the un-Aristotelian principle that what exists for the sake of the best is natural; however, many things that exist for the sake of the best are artificial. Further, even granting this principle, it is wrong to conclude from the naturalness of the *polis* that human beings are political—some might be asocial. (3) At 1253a9–18, Aristotle argues that nature does not make anything pointlessly, but nature has made us capable of speech, which is an instrument for communicating good and bad and just and unjust. This capacity is the same as the capacity to form political communities, from which it follows that human beings are, by nature, political. But Keyt argues that the capacity to form political communities is not the mere "natural justice" with which we are born, but that capacity developed, by habituation, into a virtue. And this shows that both the *polis*, and humans being political, are artificial rather than natural. (4) Finally, at 1253a19–29 Aristotle argues that the *polis* must be prior in nature to the individual, because the individual when separated from the *polis* is not self-sufficient. Keyt understands this as the claim that the *polis* can exist without the individual but

the individual cannot exist without the *polis*; but, he argues, an individual *can* survive without his *polis* (think of Philoctetes stranded on an uninhabited island). Further, Keyt argues that since parts of artificial wholes can survive the destruction of the wholes of which they are parts, Aristotle must be assuming that the *polis* is natural when he asserts that its parts can't survive the destruction of the whole *polis*. Before we attempt a wholesale defense of Aristotle's "by nature" claims, it's worth noting on point (4) that Aristotle does not say that an individual can't survive without the *polis*, only that he lacks self-sufficiency, so the priority by nature claim must be a weaker one, namely that a self-sufficient individual cannot exist without a *polis*.

Miller (1995: 30–61) rejects Keyt's "internal cause" criterion for existing by nature. Miller argues that human beings are "by nature political" in that we have an innate potentiality to form cooperative associations to bring about a common end (like ants and bees) and specifically to form *poleis*. This potentiality is seen in (1) our capacities to perceive what is just and to reveal what is advantageous and just in speech, since nature does nothing in vain, and (2) our impulse to live in communities—for we desire to live in communities, and living in communities enables the common advantage and life itself. The presence of these potentialities in all human beings is best explained as for the sake of political cooperation. Miller interprets "the *polis* exists by nature" to mean, the *polis* exists for the sake of the fulfillment of *human* nature, which is political. Miller adds that not all things said by Aristotle to exist by nature have an internal cause of their change and rest, for example, spider webs and birds' nests are natural (*Physics* 2.8 199a7–8, 29–30—although strictly speaking, what Aristotle says is that it is natural for spiders to make webs and birds to make nests). Miller's is a fairly weak sense of existing by nature; notice that the virtues, which Keyt points out are said by Aristotle *not* to exist by nature (*Nicomachean Ethics* 1.1 1103a19–5) also exist by nature in Miller's sense. But this is acceptable, Miller argues, if we remember that we should study each subject with the degree of precision appropriate to it (*Nicomachean Ethics* 1.3 1094b12–14): by "nature" in the *Physics* may have the strong sense of "having an internal cause of change and rest" but it does not in the *Politics*.

Even if we take claims about the naturalness of the *polis* to be shorthand for claims about the tendencies in human nature, as Miller does, Keyt's (1)–(3) are devastating criticisms if we agree that the force of "by nature" is to be understood, as in the *Physics*, by contrast with "by reason/art" as well as with "by chance/luck." However, we have at least two reasons not to understand nature by contrast with reason and art in the *Politics*. First, Aristotle defines human nature in terms of reason, and the natural human function is to use reason (*Politics* 7.15 1334b15: reason and intelligence are the endpoint of our nature). Second, if we think about Aristotle's naturalness claims in the context of claims made by his contemporaries and predecessors in political thought, we can see that the naturalness claims are, in the first instance, opposed to the idea that the *polis* is the result of a contract between parties seeking their own advantage to refrain from doing injustice in order to avoid suffering it. Aristotle is making the point that it is not just to avoid suffering injustice that human beings form *poleis* (nor only, as Plato argued against the social contractarians, to meet our needs) but because human nature is fulfilled and human virtue expressed in political life. By contrast, in the *Physics* Aristotle is not only arguing against physicists who think that the world is the result of chance and that many properties thought natural are actually conventional (Democritus) but also against Plato, who thinks that an orderly world such as ours must be the product of a craftsmanlike God—against Plato, Aristotle needs to argue against assimilating nature

to art. Aristotle's own view that human beings have capacities that are fulfilled by the political life, but that this fulfillment requires the (for human beings, natural) contributions of reason and art, is quite similar to that of the sophist Protagoras, according to whom humans have god-given capacities for justice and respect which are cultivated by habituation and perfected by the political art (Plato, *Protagoras* 322c–23a, 325c–28b).

3. The Best *Politeia*

By the time Aristotle wrote his *Politics*, political thinkers of several generations had been writing accounts of best *politeiai* (constitutions). Examples of these are Plato's *Republic* (from *res publica*, the Latin for "*politeia*") and *Laws*, and Xenophon's *Constitution of the Lacadaemonians*. These works describe the way of life of an ideal *polis*, and sketch, in more or less detail, the laws, that is to say the institutional arrangements and customs, that would be causes and effects of this way of life. They typically begin with the physical situation, move on to the production of healthy offspring, and then to the physical and moral education of the young. Aristotle follows his predecessors in thinking of a *politeia* as the way of life of a *polis* (4.11 1295a40–b1) and in the order of presentation of this way of life. However, Aristotle's own description of the best *politeia* (Books 7–8) is unfortunately incomplete, breaking off in the middle of a discussion of citizens' education in music. Aristotle has discussed the questions of extent of the territory and the size of the population of the ideal city, as well as the character and occupations of the citizens. He answers these and other questions by reference to the twin goals of self-sufficiency and virtue. Thus the best *politeia* requires a population just large enough for self-sufficiency but small enough that the citizens know one another; a territory large enough for self-sufficiency but small enough for the purposes of defense and commerce (7.4–5); whether it engages in trade and has a naval empire or not depends on how these measures contribute to or detract from the goals of self-sufficiency and virtue (7.6). Similar considerations apply to the provision of common meals and the distribution of land (7.10), and the layout of the city for the health and safety of its inhabitants (7.11–12). The goal of virtue requires that citizens be both intelligent and spirited (namely, Greek) in nature (7.7) and bars from citizenship farmers, craftsmen, laborers, and in general money-makers (7.8–9). This restriction of citizenship puts Aristotle not only at odds with democracy, but also with Plato's *Republic*, which, although anti-democratic, considered members of the economic class ("producers") to be citizens. We will look at why Aristotle restricts citizenship below. First, however, let us consider the account of education with which most of Aristotle's description of the best politeia is concerned.

Immediately prior to the account of education, Aristotle pauses to reiterate the goal of the *polis*: happy, and so actively virtuous, citizens (7.13). The account of citizens' physical and moral education that takes up the rest of the work is directly concerned with producing virtue. Although Aristotle discusses the sort of education that is necessary and useful—instrumental towards performing one's social function, for example, and physical education—his focus is on music, because that, he says, has to do with how we use our leisure. Virtue pertains not only to the socially useful things we do but also to that for the sake of which we have done this work, those activities that are not useful for the sake of anything else, but good in themselves. Music in fact serves three purposes: education, relaxation, and right use of leisure. First, Aristotle believes that music represents emotion most exactly and so listening to and playing the right kind of

music trains us, from childhood on, to have the right kinds of emotions for virtue. An important political idea in this discussion is that participation in an activity makes one a better judge of that activity (8.6 1340b24–25): hence, we may infer, participation in ruling makes one a better judge of the rulers when one is being ruled. This would not only make one appreciative of good decisions, but perhaps also sympathetic of decisions that involved some loss for oneself. Second, music as relaxation refreshes us for work. In *Aristotle: Political Philosophy*, Richard Kraut (2002: 200–2) elaborates on music's third purpose, suggesting that traditional poetry expresses wisdom and listening to it is the ordinary citizen's approximation of the best human activity, philosophizing or contemplating God and the way that divine activity results in an orderly natural world.

Because of its incompleteness, Aristotle's sketch of the best *politeia* will have to be filled out and evaluated in terms of what he says about the best *politeia* elsewhere in the *Politics*. Fortunately, Book 2 discusses actual and theoretical *politeiai* that have been put forward as particularly good by others because, Aristotle says, this discussion allows us to learn what is right and useful in them, and from what is wrong, to see the need for a *politeia* different from all those that have so far been considered (1260b29–36). Studying the views of others on a given topic is, in fact, Aristotle's standard procedure; he uses these views to formulate problems (most commonly, conflicting views on the topic) and measures his progress by his ability to solve these problems (*Nicomachean Ethics* 7.1 1145b1–8; *Metaphysics* 3.1 995a24–995b4). We can adopt Aristotle's procedure to follow his account of the best *politeia*, using the criteria by which he faults the *politeiai* surveyed in Book 2 to understand what motivates the institutions of the best *politeia* he describes, and to supplement his account where he falls silent. Relevant also is the idea that the virtue of citizens lies in ruling and being ruled, which, in the best *politeia*, would coincide with the virtue of a human being.

Politics 2 criticizes, among others, the best *politeia* of Plato's *Republic* (chapters 2–5) and the second-best *politeia* of Plato's *Laws* (chapter 6), Phaleas of Chalcedon's property-egalitarian best *politeia* (chapter 7), and the actual *politeia* of Sparta (chapter 9). Two themes recur in these discussions. First, Aristotle points out the lack of fit between the virtue these *politeiai* aim at and the institutional means by which they seek to achieve it. For instance, Plato's best *polis* aims at the greatest possible unity among citizens (*Republic* 462a–d), but Aristotle argues that this goal is undermined by the guardians holding all possessions in common, for that leads to a diminution in each individual's sense of responsibility for the care of those possessions, whether land or offspring (1261b32–40), and deprives individuals of the opportunity to exercise generosity towards fellow-citizens (1263a40–b14). (Aristotle claims that Plato's "greatest possible unity" is a degree of unity inappropriate to the *polis*, which is a naturally plural entity (1261a16–22 *passim*). This sounds sensible enough, but it is not easy to see how the kinds of plurality Aristotle points to interfere with the kind of unity aimed at by the *Republic*'s best *politeia*. For this *polis* is unified in the sense that each citizen feels that his or her fortunes and misfortunes are tied to those of the city as a whole (462a–d). Aristotle says that a *polis* is composed of people of different types, who perform different social functions. But Plato agrees with this; he just thinks that the citizens can, while performing different social functions (and in part because of it), also see that their fortunes are tied to one another.

Phaleas' constitution aims at economic equality among citizens, but unless the initial equal distribution of property is matched by a controlled population policy, there will be poor citizens (1266b8–14). Sparta's constitution aims at military virtue, but its inheritance laws are such that not enough men meet the property qualification to make a sizeable

defense, and its policy of encouraging births increases poverty (1270a19–33140b5). The most general misfit is between the *politeiai*'s aim, virtuous citizens, and their reliance on institutional arrangements rather than education to bring about citizens' virtue.

The second major theme of these discussions is related to this last point. Aristotle faults the authors of the *politeiai* for misdiagnosing the cause of social ills, taking them to come from, for example, economic inequality rather than vice. Thus Phaleas of Chalcedon's egalitarian *politeia* is built on the assumption that the source of political conflict lies in inequality of resources, but according to Aristotle, equality in honors among men who are not equal is as significant a source of political conflict (1266b37–1267a2); further, injustice is committed not only by those who lack necessities, but also by those whose desires are excessive (1267a2–17). Similarly, Aristotle thinks that Plato has misdiagnosed the source of disputes over property, supposing them to come from the institution of private property when they in fact come from vice. Evidence of this is the presence of property disputes within families (1263b19–29). Finally, Aristotle notes that quite apart from whether the distribution of resources is equal or unequal, it matters how much citizens have: too much and they will live too luxuriously; too little and they will be poverty-stricken (1266b24–31). But to avoid luxury it is necessary that desires be moderated, and so a constitution must give thought to education (1266b28–34).

Aristotle's criteria for the correctness of a *politeia*—legislating for the common advantage (1283b35–1284a2) and distributing political functions according to merit—are not foregrounded in these discussions, presumably because he is only discussing *politeiai* that have a claim to be correct. (The criteria themselves are closely linked to his account of justice. *Nicomachean Ethics* Book 5 distinguishes two senses of justice. (1) General justice, that is, obedience to the law, aims at the common good; this comprises the whole of virtue insofar as other people are concerned (5.1 1129b11–26). (2) Particular justice, insofar as it is distributive, is concerned with equality (5.2 1130b10), apportioning goods such as honor and wealth to merit (5.3 1131a22–25); insofar as it is corrective, it restores to the offended party in an unjust transaction what he has lost to the offender, irrespective of their merit as persons (5.4 1132a2–3, 25–29).) However, he does raise the worry about apportioning offices to merit that the *Republic*'s appointment of permanent rulers would give rise to faction from the military (*Politics* 1264b6–10). This is a reminder that Aristotle regards (and expects others to regard) ruling as a privilege to be shared among the deserving, and not merely as a job to be done that is shared among the qualified.

Yet this idea that citizens should share in ruling and being ruled, because there is a value to political participation for whoever cultivates sufficient virtue to engage in it, is in tension with another idea we see in Aristotle: that wherever there is a ruler sufficiently superior to others, the superior should be the one to rule (7.14 1332b12–40, cf. 3.17 1288a15–29). Aristotle says such conditions don't obtain, but this just papers over the tension. Suppose that we could be ruled by gods. Insofar as they would do a better job of ruling, they should. But politics was supposed to be an arena for the expression of practical wisdom for us, and the goodness of political activity for us doesn't disappear just because there are others who can do it better.

Aristotle's best *politeia* reveals another tension in the ideal of ruling and being ruled, for it excludes from citizenship farmers, craftsmen, and in general money-makers (7.9 1328b33–1329a1). Aristotle has already (in Book 1) discussed the mental incapacity of women and natural slaves that would exclude them from citizenship, but what is the reason for excluding the economic producers? Aristotle's main idea is that economic activity impairs the mind of its practitioner and leaves him too little leisure time to

develop virtue and engage in political activity. He seems to think there is something about money-making that distorts one's outlook. Perhaps it is difficult for the businessman to subordinate profit to virtue when he switches to running a *polis*? Perhaps it is difficult to adjust to a consultative political procedure? Kraut (2002: 216–17) suggests that the occupations Aristotle thinks disqualify their practitioners from political participation involve a narrowly restricted use of reason, namely, the determination of means to a pre-determined end, and an instrumental relation to others. While Kraut thinks Aristotle's concerns about the effects of such an outlook on politics are legitimate, in general he attributes Aristotle's "prejudice" to his sense that there is something base about using one's mind or body principally as a means of survival.

In *Politics* 7.8–9, Aristotle engages in some tortured reasoning to justify the exclusion of the economic classes from citizenship. He applies to the *polis* a distinction between what is necessary for something and what is part of it (as, for instance, the housebuilding art is necessary for but not part of the house it produces) and claims that there is nothing in common (shared) between the merely necessary and the part (as there is nothing in common between the housebuilder and the house). Then, enumerating all the functions necessary for the *polis*—providing food, defense, judgment, etc., he argues that only those who provide defense, judgment, and religious services can live a life of virtue, which is the common good shared by all the parts of the *polis*. But the economic producers' inability to live a life of virtue is the result of institutional arrangements, rather than innate capacities. An alternative to assigning to some the life of the farmer and to others the life of a priest would be to have people alternate between economic and political activity. At one point, Aristotle expresses the hope that the farmers will be slaves (1330a25), presumably natural slaves, suggesting that it would be preferable for the necessity of providing for the *polis* to coincide with the availability of those who can provide for it. But if nature does not deliver up this perfect fit, we might ask, isn't it an unjust arrangement that assigns the economic life to some who have the natural capacity to lead a political life, and then excludes them from political life on the grounds that their economic life doesn't allow them to develop and exercise the virtue required for political life (that they could have had, given different opportunities)? In aiming at happiness, Aristotle's ideal *politeia* takes on not just the task of enabling virtuous citizens to be active, but also of making citizens who are capable of it virtuous.

Our examination of Aristotle's political philosophy has taken the approach of thinking about the role particular (good or bad) ideas play in his overall project—in the case of natural slavery, to highlight the value of the distinctively political; in the case of the naturalness of the *polis*, as an alternative to the account of the *polis* as the result of a compromise social contract. Thinking about his extreme restriction of citizenship from the perspective of the *polis*'s goal of virtue and happiness for its citizens suggests that Aristotle is caught in a bind. If someone is a citizen, the *politeia* is supposed to aim at his virtue and happiness. But if his role in the *politeia* blocks him from developing virtue and happiness (as Aristotle thinks is the case with economic activity), the *politeia* can hardly be said to aim at his happiness. And yet the *politeia* cannot do without his economic activity; on Aristotle's view, virtuous and happy citizens require that someone else produce.

Related Topics

Plato's Political Philosophy, Aquinas, Virtue Ethics and Political Philosophy, Natural Law and Rights Theory

References

Adkins, A.W. (1991) "The Connection between Aristotle's *Ethics* and *Politics*," in Keyt, D. and Miller, F.D., Jr (1991) *A Companion to Aristotle's "Politics"*, Oxford: Blackwell.

Keyt, D. (1987) "Three Fundamental Theorems in Aristotle's Politics," *Phronesis* 32: 54–79; reprinted as "Three Basic Theorems in Aristotle's Politics," in Keyt, D. and Miller, F.D., Jr (1991) *A Companion to Aristotle's "Politics"*, Oxford: Blackwell.

—— (1999) *Aristotle: Politics Books V and VI*, Oxford: Clarendon Press.

—— and Miller, F.D., Jr (1991) *A Companion to Aristotle's "Politics"*, Oxford: Blackwell.

Kraut, R. (1995) *Aristotle: Politics Books VII and VIII*, Oxford: Clarendon Press.

—— (2002) *Aristotle: Political Philosophy*, Oxford: Oxford University Press.

Meikle, S. (1995) *Aristotle's Economic Theory*, Oxford: Oxford University Press.

Miller, F.D., Jr (1995) *Nature, Justice and Rights in Aristotle's Politics*, Oxford: Oxford University Press.

Robinson, R. (tr. and comm.) (1962/1995) *Aristotle: Politics Books III and IV*, Oxford: Clarendon Press.

Saunders, T.J. (tr. and comm.) (1995) *Aristotle: Politics Books I and II*, Oxford: Clarendon Press.

3
AQUINAS

Paul Sigmund

In the past many college courses in the history of political theory moved directly from Cicero to Machiavelli, a space of more than 1,500 years, regarding the writings of St. Augustine and St. Thomas Aquinas and the Middle Ages as works of theology rather than political philosophy. More recently Augustine has received attention for his pessimistic view of the possibilities of politics and human justice. In the case of Aquinas it is now recognized that his theory of natural law is an original synthesis of earlier approaches that can be applied to many contemporary moral problems, including war, sexuality, economics, and the relation of religion and politics. In addition, the revival of interest in Aristotle's philosophy as providing a possible solution to the dilemmas of modern philosophy has increased interest in Aquinas as one of his most influential interpreters. Also, the increased diversity of opinion among Roman Catholic interpreters (Kerr 2002) since the Second Vatican Council (1962–65) has produced a large and controversial interpretive literature in recent decades. And the increased recognition of the influence of religious assumptions on political behavior has heightened interest in the relation between theology and political philosophy.

1. Biographical Background

Thomas Aquinas was born in 1225 in a castle (the ruins of which remain today) outside the town of Rocca Secca midway between Rome and Naples. The de Aquino family belonged to the lesser nobility of the Kingdom of Sicily, ruled by Frederick II, who had been crowned Holy Roman Emperor in 1208. Like many of his predecessors Frederick was in continual conflict with the pope, the borders of whose temporal possessions, the Papal States, extended to an area north of the Aquino castle. The family sent Thomas to be educated at the nearby Benedictine abbey, Monte Cassino, but when it was occupied by Frederick's soldiers, Thomas was sent to the new university that Frederick had founded in Naples. There he was influenced by the Dominicans, a new reforming religious order founded by St. Dominic to convert the Muslims. Over the family's opposition, which included detention by his brother for more than a year, he joined the Dominican order and went to study with them at the University of Paris. In Paris, the works of Aristotle, newly translated from the Arabic, were the subject of controversy between those who saw his teachings as contrary to the faith, and those who, like Aquinas's Dominican mentor, Albert (later to be known as St. Albert the Great) sought to develop a synthesis of reason and faith, based on the assumption that "grace does not contradict nature but perfects it." After studies in Paris and Cologne Aquinas received his degree from the University of Paris in 1252, and he taught and wrote in Paris and in

Italy for the rest of his life. He died in 1274 several months after he discontinued work on the unfinished Part III of his most important work, the *Summa Theologiae* (Summary of Theology).

Aquinas's *Summa contra Gentiles* (*Summary against the Gentiles*) (1264) and the *Summa Theologiae* (1265–73) were theological rather than political works, but the latter work, especially Part II, contains discussions of law, politics, and ethics, as well as the relations of religion and politics that were part of Aquinas's impressive synthesis of both Christian (the Bible, the Church Fathers, and St. Augustine) and classical (Aristotle, Neo-Platonism, Roman law) thought. His only specifically political writings were the first section of a work entitled *On the Governance of Princes* (*De Regimine Principum*) or *On Kingship* (*De Regno*) and his commentaries on Aristotle's *Politics* and *Nicomachean Ethics*.

2. The Influence of Neo-Platonism

Aristotle was not the only classical element in the Thomistic synthesis. His cosmological world view is Christian, but the particular vision of the universe that he lays out in his writings is derived from Neo-Platonism, and especially from (Pseudo) Dionysius the Areopagite. Dionysius, or Denys as his name is sometimes translated, was a fifth-century monk who was the principal agent of the transmission to the Middle Ages of late Platonism. The Neo-Platonists gave a theological interpretation to Plato's logical writings, identifying Plato's references to the One and to the Idea of the Good with a divinity that has created the universe as an overflowing of divine Goodness from which the world and everything in it descends in a hierarchical fashion and to which it strives to return. The works of Dionysius are frequently cited in the *Summa Theologiae*. They had acquired an almost scriptural status since he was mistakenly taken to be the "Dionysius" mentioned in the *Acts of Apostles* (*Acts* 13) who was converted to Christianity by St. Paul on the hill of the Areopagus in Athens. Man fits into the Neo-Platonic hierarchical world view between the angels and lower animals and is designed to find his fulfillment (happiness) in returning to God. He will do so through contemplation of God in this world and union with Him in the beatific vision after death.

Plato's dualism, the sharp division between the physical world of appearances and the unchanging ideal world of the spirit also influenced the Church Fathers including Augustine who read Plato in Latin translation. Their vision of an ordered hierarchical and unchanging world of the spirit, and the changeable and unstable world of the flesh, was easily combined with the Christian message in a world view that contrasted heaven and earth, and argued that the power of God's grace could overcome the weakness of the flesh.

3. The Metaphysics and Epistemology of Aristotle

The sudden irruption into Latin Europe of the entire corpus of the works of Aristotle in the first half of the thirteenth century presented a different view of the world. Aristotle outlined four causes, which he called formal, material, efficient, and final—i.e. the basic matter, the organizing principle or form, the instrument that acts to produce change, and the purpose or goal that is demonstrated by its structure and functioning. Instead of a static disjunction between the physical and spiritual worlds, matter is given a dynamic character in developing towards its inbuilt purposes and the changing world

has a patterned meaning available to human inquiry. Those purposive changes indicate that there must be a substance that caused them, whom Aristotle calls the Prime Mover, that is eternal, infinite, intelligent, and good. Because man is composed of the body as matter, and the soul as form, he can think and reason to the purposes that are evident from the intelligent design of the world.

What Aristotle's writings presented to Muslims, Jews, and Christians was an organized philosophy that seemed to have many similarities to the biblical message, as well as important differences from it. Two responses were possible. One could emphasize those differences and, suspicious of reason, adhere to the sacred texts as the sole reliable source of both religious and philosophical knowledge. This was the initial reaction of Church authorities in Paris when the translations of Aristotle first appeared; they tried, usually in vain, to prevent the reading and teaching of the new texts. In the case of Islam, after a period in which the study and commentary on Aristotle flourished, this was also the reaction as the religious and civil authorities made the fateful decision that the Koran would be the sole source of religious and philosophical truth.

Aquinas was aware that the Muslim interpreters had attributed to Aristotle some doctrines that were contrary to Christianity, e.g., the eternity of the world, and the existence of a world-soul, thus denying the individuality of the soul. However, he saw the possibility of a synthesis of Aristotelianism and Christianity that could give a philosophic foundation to the doctrines of divine revelation. Although he was aware that the human intellect and will had been damaged by the effects of sin, he believed that God reveals Himself to man through both revelation and reason, and that rightly understood the two will not conflict.

4. Christianizing the Politics and Ethics of Aristotle

An important change that Aquinas made was to make the Prime Mover a personal God and place in the mind of God the purposes of nature that had been expressed in the doctrine of final causation. Where Aristotle had said "nature does nothing in vain," Aquinas attributed rational purposes to the mind of God as expressed in the teleological (from *telos*—end, goal, purpose) structure of the created universe. Ideally those purposes should not contradict one another since an intelligent Creator has established a harmonious universe. Human beings are composed of body and soul, matter and form, and the soul shares with God reason that enables us to observe from human behavior and goals the intentions of a rational Creator.

What this meant for politics is that a philosophical foundation was established for government. From the human tendency to work together in the family, the community, and larger entities, Aquinas concludes that "Man is by nature a political and social animal" thus correctly translating Aristotle's *zoon politikon* which has a broader meaning than the political alone. Politics is not a negative enterprise concerned mainly with curbing the evil tendencies resulting from sin, as in Augustine, but a way of achieving common goals for the betterment of society. The family is a natural institution that has an important educational role and monogamous marriage is the best way to carry out the purposes of God and society in developing human beings who find their fulfillment in working with others for human betterment. Reason can also demonstrate the existence of a beneficent, eternal, spiritual God who guides human development by divine providence. Since he intends the good of his creation he wishes men to be happy and ultimately to find the fullness of happiness in union with him. This is done by living

a life of practice of the virtues described in the *Nicomachean Ethics*, and, more importantly, of the supernatural virtues of faith, hope, and charity. Aquinas's ethical naturalism should not be exaggerated. There is always revelation as a guide and corrective of reason. But the intellectual basis for rational argument about politics and ethics has been established and reason legitimated as a source of norms. In contrast with what was occurring in the Islamic world at the same time, rational analysis of politics and ethics is legitimated in Christian Europe.

5. Natural Law: Context and Content

The most important and influential application of Aristotle's teleological approach is Aquinas's theory of natural law. However, Aquinas also draws on many other sources besides Aristotle, and his work is set in a theological and cosmological framework that is important for the continuing debate about the contemporary relevance of Thomistic natural law.

The discussion of natural law is part of a broader set of arguments in Questions 90–97 of the First Part of Part II of the *Summa Theologiae* that are sometimes referred to as Aquinas's *Treatise on Law*. Like the rest of the *Summa* it is organized in the format that scholastic philosophy (philosophy of "the schools" or "schoolmen") employed, which reflected the structure of the oral disputations at the university: a topic presented as a Question, followed by a series of Objections, then a Response to the Question, followed by answers to the Objections.

Aquinas begins with a definition of law as "an ordination of reason (i.e. a rational command) for the common good made by the one who is in charge of the community." He then applies the definition to God's rule of the world, describing divine providence as the operation of the reason of God in providing for the good of his creation. This, he calls the Eternal Law which he describes as the "plan of divine wisdom moving all things to their appropriate ends"—thus linking Aristotle's teleology to the mind of God. He then defines Natural Law as "the participation in the Eternal Law by rational creatures" who have been created with a natural ability—which he calls a "natural inclination" to understand God's purposes as expressed in the moral order of the world. The basic principles of the natural law are known directly and he describes them as "indemonstrable." When they are applied to specific circumstances they become that foundation for Human Law, which he later divides into civil law and the law of nations. Aquinas then argues that Divine Law, by which he means the law contained in the Bible, provides additional certainty about basic moral principles, and it also covers areas that Human Law, which is concerned with external conduct, cannot regulate because they involve internal motivations or direct man to his supernatural goal, union with God.

Aquinas then discusses each of these types of law in a separate question. His description of natural law in Qu. 94 has been a subject of many debates both over its interpretation, and its relationship to the other forms of law outlined above. Beginning with the self-evident proposition that good is to be done and evil avoided, Aquinas then argues that since good is an end, we can know about it from examining the "natural inclinations" to moral conduct which God and nature have instilled in us as subjects of the natural law. Aquinas singles out the instinct to self-preservation that is shared with all beings, union of man and woman and the upbringing of children that is shared with all animals, and the natural inclination shared with all rational creatures to live in society, and to worship God. The basic principles related to these inclinations are known by

all, but more specific conclusions and applications are less certain. In addition, specific provisions that are convenient for society (Aquinas cites private property and slavery) may be added to the natural law. If we are looking for examples of the way in which the natural inclinations are translated into moral or legal principles we must wait until the next part of the *Summa*, which deals with individual issues. At this point we know he will link them to insights into more or less permanent needs, aims, and characteristics of human nature that have been designed by God to produce human fulfillment and happiness.

What happens if the various laws appear to conflict? Aquinas believes that God, who has created a harmonious and rational universe, would not allow a genuine conflict to emerge. If a human law conflicts with natural law, it is unjust—even "an act of violence" and need not be obeyed. Aquinas does not indicate who decides in cases in conflict. He suggests numerous candidates, depending on the area in question, including the pope, the civil authority, and the individual.

6. The Best Forms of Government

The law is made by the one who is in charge of community. What does this mean in practice? Aquinas has often been portrayed as an advocate of monarchy and he clearly endorses it in the unfinished work On Kingship. Monarchy is the best form of government because it is most effective in providing unified decisions about the common good, and it reflects the unity of God and the patterns of nature (including, for example, borrowing from Aristotle, the "king-bee"). Yet he makes favorable comments about the Roman Republic, and observes that while monarchy is the best form of government, tyranny is the worst. His conclusion is that monarchy should be limited and he promises to suggest ways to limit it. He does not deliver on the promise, but in his discussion of the legal provisions of the Old Testament ("the Old Law") in the *Treatise on Law* of the *Summa* he declares that "the best form of government for a city or a kingdom is one in which one person rules in accordance with virtue and under him are others who govern in accordance with virtue, and all have some part in government because all are eligible to govern and those who govern are chosen by all" (*Summa Theologiae*: II-II, Qu. 105). He describes the result as a combination of monarchy, aristocracy, and democracy. Similarly in his discussion of human law, after describing the characteristics of monarchy, oligarchy, and democracy, he states that "here is another form which is a mixture of these, and this is the best form of government" (ST: I-II, Qu. 95, a.4). Aquinas is usually described as an advocate of monarchy but his discussion of the role of the people was cited, centuries later, by the proponents of European and Latin American Christian Democracy.

Some late-medieval writers, especially in the Conciliar Movement of the fourteenth and fifteenth centuries that attempted to limit the absolute power of the pope, made an argument for the consent of the people or the bishops to Church decisions (Sigmund 1995). For Aquinas as a member of a religious order under the pope, writing in a period of legal and doctrinal centralization, the pope was the final authority in matters of Church doctrine. And in chapter 14 of *On Kingship* his authority seems to go much further when Aquinas argues that because the Church guides men to a higher end kings should be subject to priests, especially the pope. However, there are passages in the *Summa* that seem to limit clerical jurisdiction to spiritual matters, recognizing what appears to be an independent area of authority of the civil ruler (ST: II-II, Qu. 60, a.6). On the other

hand, Aquinas's treatment of heresy recommending that the heretic be turned over to the secular arm "to be removed from the world by death" (ST: II-II, Qu. 10, a.6) seems to proceed from a belief in subordination of the state to the Church in this area.

On the question of resistance to unjust laws, Aquinas's answer is ambiguous. On the one hand, deference to one's betters would seem to be the solution, and in the case of extreme tyranny, Aquinas's solution in *On Kingship* seems to be prayer to God. On the other hand, Aquinas says that individuals are not obliged to obey superiors if they issue commands that violate the individual's rights in "matters relating to human nature … since by nature all men are equal" (ST: II-II, Qu. 104, a.5).

One of the best-known sections of the *Summa Theologiae* is the second part of Part II where Aquinas deals with specific moral issues. Earlier in the thirteenth century, the Council of Lyons had made annual confession of sins to a priest obligatory for all the faithful. There was, therefore, a need for guidance to confessors who were responsible for assigning penances and would be interested in advice as to the gravity of various sins. "Penitentials" or guides for confessors already existed but Aquinas carried out a more systematic discussion based partly on the teleological approach which he had taken from Aristotle and utilized as the foundation for his theory of natural law. Many of his arguments continue to be discussed today, because of their influence both on Christian moral theology, and on writers on issues of national and international law such as just war, property, and social welfare policy.

On the question of the use of coercion in the area of faith, Aquinas takes the position that no one should be forced to accept Christianity because, by definition, the nature of faith is such that it is only genuine if it is freely accepted. Parental responsibility for the education of their children means that Jewish parents have a right to bring up their children in their faith, until they are mature enough to make religious decisions of their own free will. The religious practices of Jews and pagans are to be tolerated, particularly if suppressing them would cause harm to the public good. Pagan rulers have the right to rule unless they interfere with the faith because the Divine Law does not abolish human law "which is based on reason."

This represents a relatively liberal attitude in a society that assumed religious uniformity but Aquinas made significant concessions to existing practice. The Jewish religion is tolerated because it prefigures Christianity and so that we can learn from "our enemies." Heretics have committed a sin of promoting false doctrine that is more serious than counterfeiting money. If they persist after two admonitions, "the salvation of others" should be protected by handing them to a secular judge who is to impose the death penalty. Apostates who have accepted the true faith and later rejected it are to be deprived of their right to rule and their subjects absolved from their oaths of fealty. Pagan rulers may also be deposed if there are religious reasons for the church to do so.

7. War, Property, Lying, and Sex

These arguments, liberal for the age in which Aquinas wrote, are no longer relevant in a society that accepts religious freedom and rejects the use of coercion in matters of faith. However, Aquinas's analysis of the conditions of just war continues to be discussed in the literature on the subject. For a war to be just it must be carried out by a legitimate public authority (thus rejecting the common resort to private violence in his day), it must have a just cause which he defines broadly, and it must continue to be carried out for good purposes. Aquinas is also the *locus classicus* for the doctrine of double effect,

that is, that destructive weapons can be used in a just war even if it is known that their use will have an unintended bad effect, such as the loss of life by innocent civilians (called today "collateral damage"). He also appeals to the doctrine in his discussion of the legitimacy of self-defense in which the action of killing an aggressor as the only way to save one's life is foreseen but not intended.

Aquinas then turns to the question of private property. Here he has a problem of reconciling Aristotle's defense of private property as natural to rational human beings, provided it is used in a socially responsible fashion, with the statements by the Church Fathers and some legal sources that community of property is the ideal. Aquinas resolves the contradiction by adopting Aristotle's arguments about the advantages of a private property regime, but describing it as an addition to the natural law created by human reason and agreement. Aquinas also adds that God intends that material goods be used to supply the necessities of all mankind, so that property holders are obliged to provide for the needs of the poor, and a person in imminent danger of starvation can take what he needs from another and not be guilty of theft.

The classical and biblical traditions did not conflict in their opposition the payment of interest for money lending—the sin which Aquinas calls usury. Aquinas repeats Aristotle's teleological argument that since money is not productive there is no loss by lending it as there would be in the case of the loan of a cow or a house. He then interprets the Old Testament prohibition of the taking of interest from one's neighbors as now applicable to all, although he recognizes that usury is tolerated by human law "because of the imperfection of men."

Aquinas's discussion of lying is an example of the use of the teleological approach to justify an absolute prohibition. Lying is intrinsically evil because it violates the nature of speech which is to communicate truth to others. Since evil may not be done so that good may come of it, lying is not permitted even in extreme cases such as the likelihood that a lie could save many lives. To mitigate the harshness of this advice, Thomas quotes Augustine who allows the use of "prudent dissimulation," short of an outright lie.

The teleological approach is also used as basis for sexual morality. Aquinas begins with a surprising question, "Is virginity lawful?" He asks this question because it might appear that the natural inclination to form families and bring up children seems to be frustrated in such a case, calling into question the religious vows of celibacy and chastity. Aquinas responds that the duty to continue the human race is a general one and one may give up sex and family life to devote oneself to "the contemplation of the divine for the improvement and welfare of mankind."

For Aquinas, the nature and purpose of sexual relations are such that they may only be enjoyed in marriage. Aquinas then ranks sexual sins by degrees of sinfulness. Fornication is contrary to the nature of man because the children that may result cannot be brought up by both a mother to nurse them and a father to defend and educate them. For similar reasons adultery is wrong. Other forms of sexual activity outside of marriage are described as unnatural and "opposed to the natural order of the sex act" including in order of increasing seriousness non-coital sex, masturbation, homosexuality, and incest.

Our focus on what is allowed and what is forbidden by natural law may give a misleading impression of Aquinas's approach to morality. Natural law is only analyzed in a single question out of the hundreds in the *Summa*. Far more attention is given to divine and human law—and even more to the development of virtue, especially the theological virtues of faith, hope, and charity. Moreover, Aquinas constantly emphasizes the

necessity of cooperating with God's grace in order to live a good life—and the importance of the Sacraments and the message of Christ in leading human beings to their ultimate goal, eternal life and the face-to-face vision of God.

8. Later Influence of Thomistic Political Thought

Some of Aquinas's statements were included in a long list of condemned propositions issued after his death by the diocese of Paris which were believed—erroneously—to reflect the heretical doctrines espoused by the Muslim philosopher, Averroes. The condemnation was retracted and Aquinas was canonized as a saint in 1323. His writings, in particular the *Summa Theologiae*, were promoted by the Dominican order but rival theologies, notably that of William of Ockham, were also used in seminaries and universities. However, in the sixteenth century, in response to the Reformation, the *Summa* became the standard theological text and was placed on the altar along with the Bible at the Council of Trent (1545–63). In Spain and Italy there was a renewed interest in his work on the part of the Dominicans and the members of the newly founded Jesuit order. Spanish writers used Aristotelian-Thomistic arguments to argue—unsuccessfully—against the enslavement of the Indians in America. In England, the Anglican theologian, Richard Hooker, wrote The *Laws of Ecclesiastical Polity* (1593) which drew on Aquinas's natural law to develop arguments for community consent to government that were quoted in the seventeenth century by John Locke in his *Second Treatise of Civil Government* (1689). *On the Laws and God, the Lawgiver* (1608) by the Spanish Jesuit Francisco Suarez also drew on Aquinas's theories to argue that political authority is derived from God through the consent of the people.

In the nineteenth century, the papacy under Pope Pius IX (1846–78) opposed the liberal revolutions in Europe, especially in Italy where the Italian unification movement led to the seizure of the Papal States. Pius's successor, Leo XIII, sought to develop a theological and philosophical alternative to liberalism based on an updated Thomism. In 1879 Leo issued the encyclical *Aeterni Patris* making the study of Aquinas the basis of courses in theology and philosophy in Catholic seminaries and universities. His response to the twin challenges of socialism and capitalism was to publish *Rerum Novarum* in 1891 which used appeals to nature to defend private property and the family which he saw as threatened by socialism, while also criticizing capitalism as based on usury and greed. Quoting St. Thomas, the pope advocated the promotion of associations to protect "natural rights," including the right of workers to form unions in order to secure a living wage for themselves and their families. Forty years later in *Qudragesimo Anno*, Pope Pius XI developed these ideas further, arguing for what he called the principle of "subsidiarity" that protects and promotes intermediate groups between the individual and state as well as councils in which labor and management can work together for the common good. What this meant in political terms was that the church encouraged the organization of Catholic trade unions, and supported social legislation. What came to be known of Social Catholicism claimed to offer a middle way between capitalism and socialism, which could trace its ancestry back to St. Thomas Aquinas and the medieval guilds.

The message of the social encyclicals was taken up by Catholic parties which had emerged in a number of European countries to defend Catholic institutions against what were seen as efforts by socialists on the left and nationalists on the right (e.g. Otto von Bismarck) to bring them under state control. Those parties were committed to democracy

although official church teaching accepted democracy and religious pluralism as tolerable in a situation of religious diversity although the ideal was still a Catholic state.

In the first half of the twentieth century a Neo-Thomist movement emerged in France and Germany that not only made major scholarly contributions to the study of Aquinas, but also attempted to reinterpret his political and moral teachings to support democracy, religious freedom, and human rights. The most influential writer on Aquinas's political and social thought was the French philosopher, Jacques Maritain (1881–1973). Beginning in the 1930s, Maritain argued that Aquinas' writings could be reinterpreted in support of democracy, religious freedom, and human rights. During World War II he published *The Rights of Man and the Natural Law* (Maritain 1943) which included a full list of human rights that could be developed out of Aquinas's writings. In a later work, *Man and the State* (Maritain 1951), he made his theory more broadly acceptable by arguing that Aquinas's references to natural inclinations as the foundation of natural law meant that despite deep religious and philosophic differences, human beings possess an intuitive knowledge of the basic provisions of the natural law that had made it possible for the United Nations to adopt the *Universal Declaration of Human Rights* in 1948.

Maritain's writings defending democracy, human rights, and religious pluralism also provided the ideological foundation of the new Christian Democratic parties that emerged in Europe after World War II and in Latin America in the 1950s and 1960s. They contributed to the reorientation of Catholic political thought associated with the reforms of the Second Vatican Council (1962–65) including the decrees *Gaudium et Spes* endorsing democracy, and *Nostri Aetatis* supporting religious freedom, both of which employed Thomistic natural law arguments.

Cases involving direct application of Neo-Thomist social thought include worker "co-determination" in Germany in the 1950s and the establishment of worker-owned and -managed firms in Peru in the 1970s. In Germany a law supported by both the Christian Democrats and the Social Democrats requires boards of directors of business firms over a certain size to include representatives elected by the workers. In Peru the military government in power at the time financed the establishment of worker-owned and -managed factories and required all firms to distribute 10 percent of their annual profits in shares to the workers. (The Peruvian experiments were discontinued after the return to democracy, and in the German case ways have been found to circumvent the law.) Also in Latin America the new Christian Democratic parties that were active in the 1960s in many Latin American countries cited St. Thomas on the "social function" of property to justify their support of agrarian reform.

The tight link between Roman Catholicism and Thomism was loosened as a result of the opening to other religions and philosophies initiated by the Second Vatican Council (1962–65). Catholic universities no longer required courses in Thomism, and their philosophy departments became more diversified. Papal documents tended to cite the Bible rather than St. Thomas, although the popes continued to argue that reason guided by faith could attain moral truth. The attempt to promote a third position between capitalism and socialism was abandoned, and in 1991 *Centesimus Annus*, an encyclical published, as its title indicates, 100 years after *Rerum Novarum*, Pope John Paul II recognized the advantages of the market system, while insisting that it should be subject to national and international regulation.

Aquinas's writings on natural law received renewed attention because of the writings of John Finnis, Professor of Philosophy at Oxford and Notre Dame. His book, *Natural*

Law and Natural Rights (Finnis 1980), increased interest in the Thomist approach at a time of growing interest in human rights theories. A later book by Finnis, *Aquinas: Moral, Political, and Social Theory* (Finnis 1998), highlighted the liberal elements in his thought by selective quotations from his writings that downplayed his views on the execution of heretics, his derogatory comments on women, and his (tepid) justification of slavery.

Finnis was an early advocate of what became known as the New Natural Law. Finding the teleological approach that derived moral norms from natural functioning and structures unpersuasive, representatives of the New Natural Law developed its moral theories from a list of Basic Goods that were or could be recognized by all as underlying moral experience. More traditional Thomist writers denied that this could be found in Aquinas's writing, but the New Natural Law writers argued that the natural inclinations of which he wrote, and his description of basic values as indemonstrable indicated that their theory could be derived from his writings.

9. Criticisms

The most frequent criticism of Aquinas's political and ethical philosophy is that it is not philosophy at all, but a set of arguments designed to support pre-established theological assumptions. The editor of the Cambridge edition of his Political Writings accuses him of being "fundamentally dishonest" because he is "committed in advance to a closed system of moral and religious beliefs" (Dyson 2002: xxxiv). Such a view can lead to the arbitrary rejection of philosophical arguments by attacking the motives of the authors, or worse, the refusal to take seriously any philosophical arguments by religiously committed philosophers. The hope of natural law reasoning is that there can be dialogue on the basis of rational approaches that are available to all. To give one example, the Thomist-influenced writings of Alasdair MacIntyre (e.g., MacIntyre 1990) have been widely discussed and debated in recent years. Natural law was at its outset a classical rather than Christian approach, its modern natural rights formulation is not dependent on religious assumptions, and its arguments about God are based on reason, rather than revelation.

A second criticism of the argument from natural law is that identifying morality with nature means that such beliefs are fixed and not subject to criticism or new understandings. It also requires an identification of "natural" with "moral" when the two terms have different meanings. In reply one can point to the development of natural law and natural rights theory which has been anything but static. Jacques Maritain's theory of natural law that distinguishes between a natural law based on inclination and the evolving content of written versions of it suggests the possibility of coming to new understandings of the implications of moral values as a result of human experience.

A third criticism would be David Hume's argument that natural law theory commits a logical error in failing to distinguish between factual and normative statements. Nature just "is." Is does not contain an "ought" and to assign an "ought" to a fact is fundamentally arbitrary. The defender of Thomist natural law would reply that it is possible for an observer to arrive at an understanding of common human needs and potentialities the fulfillment of which are required for what Aristotelians would call human flourishing (George 1999) and, as in human rights theory, a division can be made between those that are fundamental and those that constitute an ideal.

In the case of Thomas Aquinas, however, these criticisms are *not* central to his basic

moral theories. Aristotle, Cicero, and medieval theories of politics and ethics are only a part of his approach to moral conduct. It is more important to develop habits of virtue and to cooperate with God's grace. Politics and ethics are only means to higher goals. But, as described above, Aquinas's theories in those areas have made important contributions to political and social theory.

Related Topics

Natural Law and Rights Theory, War, Authority and Legitimacy, Toleration, Religion in Public Life

References

Dyson, R.W. (2002) *St. Thomas Aquinas Political Writings*, New York: Cambridge University Press.
Finnis, J. (1980) *Natural Law and Natural Rights*, New York: Oxford University Press.
—— (1998) *Aquinas: Moral, Political, and Social Theory*, New York: Oxford University Press.
George, R.P. (1999) *In Defense of Natural Law*, New York: Oxford University Press.
Kerr, F. (2008) *After Aquinas: Versions of Thomism*, Oxford: Blackwell Publishers.
MacIntyre, A. (1990) *Three Rival Versions of Moral Inquiry*, Notre Dame, IN: University of Notre Dame Press.
Maritain, J. (1943) *The Rights of Man and the Natural Law*, New York: Scribner's.
—— (1951) *Man and the State*, Chicago, IL: University of Chicago Press.
Sigmund, P. (ed. and trans.) (1995) *Nicholas of Cusa, The Catholic Concordance*, Cambridge: Cambridge University Press.

Further Reading

There are two useful edited editions of Aquinas's political thought. Paul E. Sigmund (ed. and trans.), *St. Thomas Aquinas on Politics and Ethics* (New York: W.W. Norton, 1988) is intended for undergraduates and also includes selections from the major critical and interpretative secondary literature. R.W. Dyson (ed. and trans.), *St. Thomas Aquinas: Political Writings* (Cambridge: Cambridge University Press, 2002) provides longer selections, most still in scholastic form, and is intended for graduate students and scholars. Timothy McDermott, *Summa Theologiae: A Concise Translation* (Allen, TX: Christian Classics, 1989) is an accessible and useful one-volume translation that omits the citation of authorities. James A. Weisheipl, *Friar Thomas D'Aquino* (New York: Doubleday, 1974) is a good biography. John Finnis, *Aquinas: Moral, Political, and Social Theory* (New York: Oxford University Press, 1998) draws on a wide range of his writings to support a liberal interpretation of his politics.

4
MEDIEVAL POLITICAL THOUGHT

Cary J. Nederman

Before one can undertake a survey of medieval political thought, it is necessary to confront squarely the historiographical question of what counts as "medieval." The present tendency is to associate the origins of the Middle Ages in Europe with a deep historical break—what Robert I. Moore (2000) has termed "the first European revolution"—occurring around the year 1000, marked by a series of fairly rapid social, political, economic, religious and cultural changes. This period saw the formation of most of the main regional and institutional units that would persist well into the modern world and of the languages (post-classical Latin and the vernaculars) and cultural artifacts that came to define European civilization. Of particular relevance to political thought was the emergence of codified legal systems—based on, but not slavishly beholden to, Roman law—as well as the circulation and study of classical non-Christian texts, first from Roman antiquity, and then gradually from ancient Greece (especially Aristotle, newly translated into Latin during the course of the thirteenth century). Not only were these sources investigated in their own right, but they were applied to and became fodder for the emergent political conflicts of the first half of the second millennium. Such controversies arose from a number of different disputes about the ordering of political authority: between the church (whether the pope, bishops, or the community of priests and believers generally) and temporal authorities (such as territorial princes, urban municipalities, and imperial power), between the Roman pontiff and the Holy Roman Emperor (both claimants to a sort of universal jurisdiction), and between and within divergent regional units of secular power (from the universal emperor, to kings and lesser princes, to incorporated cities).

The medieval world, viewed thus, represented a genuinely new phase in the history of political thought. But how long did it last? The generally accepted assessment among scholars today is that significant strains of medieval thought persisted well beyond, and were not thoroughly repudiated by, the Renaissance and Reformation. Rather, the dividing line between the Latin Middle Ages and modernity, while not entirely irrelevant, is blurry at best. On the one hand, the fifteenth and sixteenth centuries did witness the appearance of important new principles—such as the sovereign state, institutionalized systems of constitutional limitation on power, and robust conceptions of the political efficacy of the individual—that were largely foreign to medieval European thought. On the other hand, key political themes and languages pioneered before 1400 remained alive and resonant with earlier modern thinkers (Burns and Goldie 1988: 1–4). In sum, it is fruitless to look for some specific moment or event (such as the emergence of

Florentine civic humanism or Martin Luther's nailing of his famous theses on the door of the church in Wittenberg) as the definitive end of the Middle Ages. Transformation in patterns of political thought was incremental, gradual, and in many ways imperceptible during the early modern period (Nederman 2009).

Given the historical complexities involved in defining the medieval worldview, any brief attempt to characterize the political thought of the Latin Middle Ages must necessarily be illustrative rather than comprehensive. Political authors of the period between the eleventh and the fifteenth centuries generated a large body of writings on various matters of public affairs, addressing the church as well as the secular sphere and the interrelations between them, and employing multiple genres, sources and discourses. The following sketch will attempt to do justice to these complicated matters by concentrating on the single salient theme of political order, which was addressed, in particular, by reference to the ideal, widespread during the Middle Ages, that communal organization and governance formed a sort of organic unity, often expressed by the image of the body politic. A diverse range of medieval thinkers held that all communities of whatever nature or size shared a public goal or set of interests that brought them together and held them in place—a fixed common good that transcended the private or personal wishes or desires of particular men or groups. The nature of such a common good could be, and was, much in dispute, but its existence was not in doubt. In focusing on this framework, there is no intention to imply an exhaustive treatment of its application in medieval political thought, which would be impossible in brief compass. Rather, a survey of a few key thinkers who proposed an organically constituted vision of the political system must suffice. But the concept of political order conceived in organic terms does seem typical of the way in which arguments about public issues proceeded during the Middle Ages. Moreover, while medieval discussions of the body politic certainly make use of preceding source materials of both pagan and Christian provenance, they reflect unique extensions and applications of earlier ideas.

1. Organic Political Order: Foundations

The theme of communal order during the Latin Middle Ages was generally framed in terms of a conception of social and political institutions as organic unities: head and members are interdependent and reciprocally interrelated segments of a living body, characterized by hierarchy but also by inclusiveness and mutual respect. The *constitutio* of the medieval political organism was, metaphorically, a matter of physiological health. Of course, organic ideas of political community enjoyed a history prior to Christian times in the writings of authors such as Plato and Aristotle and were also attractive to political thinkers working within non-Christian religions such as Islam. Moreover, the medieval European reliance upon organic discourse and imagery resonates with an earlier neo-Platonic Christian cosmology, which regarded the universe to be an ordered and living whole composed of microcosm and macrocosm—not so much a "chain of being" as an animated but differentiated totality. According to a line of Christian thought that runs back to the Church Fathers, God created the universe as a unity, composed of many different parts, each of which is itself an organic unit that both partakes of the greater whole and has a distinct purpose within the larger scheme (Constable 1995: 251–360). Within each organic unit, as within the universe as a whole, there is higher and lower, superior and inferior. Yet this hierarchical ordering does not exclude the lesser members of the body from a proper and worthwhile role within the whole.

Medieval authors supplemented the hierarchical implications of ancient pagan and Christian organic thought with an emphasis on reciprocity, that is, mutual support and respect among the various parts of the whole necessary for the good of the whole. This implied, in turn, the need for the active inclusion of all members for the sake of the welfare of the whole. Each segment makes its own unique contribution; the organism cannot survive without the coordination and cooperation of every function. When transferred by analogy to the political realm, this vision of order coalesces into the position that the ruler is restricted in what he can do by the reciprocity built into the body politic; the participation of the other members of the organism in the business of the common welfare both sets limits on the actions of the head—he cannot properly interfere with their unique tasks—and establishes a sort of "check" on the head in cases where it oversteps its appropriate scope of activity.

Medieval organic political thought was not, however, a single, static idea. Rather, it is perhaps more accurate to characterize the image of the body politic as a site of intellectual and political struggle among many contending interpretations. On the one hand, some of its features were taken as given by all parties to the discourse. For instance, it was taken for granted that some ruling part (generally, but not always, the head) was necessary for the body's health; a headless body was doomed to misery and destruction. Likewise, it was incontestable that law was the medium through which the members of the organism cooperated and to which all parts must submit; law formed the muscles and sinews, or perhaps the backbone, of the animate polity. And no one seriously challenged the existence of some good or welfare common to all parts of the body and greater than all of them taken individually. Yet even while creating a framework for discussion, the organic model raised a large number of questions that stimulated wide-ranging debates among medieval political theorists. Could the head be corrected or removed if it engaged in conduct detrimental to the health of the body? By whom? In what way(s)? What duties and powers did members retain under ordinary circumstances? How did they coordinate their specialized roles and tasks with one another? These questions were by no means closed or settled in the medieval mind, as the discussion below reveals. The body politic provided a matrix within which central issues about the nature of rulership, governance and community could be addressed in a manner consonant with deeply held cosmological and theological beliefs.

2. John of Salisbury

No medieval version of the organic metaphor was more innovative and influential than that presented by John of Salisbury (1115/20–1180) in the fifth and sixth books of his major treatise, the *Policraticus* (completed in 1159). Of course, until recent years, it was generally assumed that the *Policraticus*'s conception of the body politic was an adaptation of a now lost treatise of instruction addressed to the Emperor Trajan, mistakenly ascribed by John to Plutarch. We now know that John devised the *Institutio Traiani* as a cloak under which to cover his own originality. John's forgery of an authoritative basis for his conception of the body politic, and the alleged provenance of the source he invents, are telling. Rather than selecting a Christian author as the originator of the organic metaphor, John attributes it to a relatively obscure (in the twelfth century) pagan writer and makes its addressee a pagan Roman Emperor famed for his superlative virtue. John thereby intimates that the principles pertaining to a properly arranged social body are by no means inherently Christian in bearing. Rather, since the precepts

of the well-ordered republic were identified by a pagan philosopher and transmitted to an exceedingly moral but unbelieving emperor, it is presumably the case that any community containing wise rulers and just subjects is capable of apprehending and applying them, regardless of religious convictions. Reason, not revelation, plays the leading role in designing the body politic.

This becomes evident in John's depiction of the constitution of the political organism. He begins with the simple naturalistic observation that the commonwealth may be likened to a "body which is animated" (John of Salisbury 1990: 66). The differentiation of the offices of political society may thus be represented in a manner analogous to the distinct parts of human anatomy. The body politic is ruled by the prince, who "occupies the place of the head" (John of Salisbury 1990: 69). At the heart of the commonwealth lies the senate, composed of the counselors whose wisdom the ruler consults. The senses correspond to the royal judges and local agents such as sheriffs and bailiffs who exercise jurisdiction in the king's name. The financial officers constitute the stomach and intestines of the body, while the two hands are formed by the tax-collector and the soldier. Finally, John compares the feet to the artisans and peasants "who erect, sustain and move forward the mass of the whole body" (John of Salisbury 1990: 67). Each part of the organism, according to the *Policraticus*, has its own definite function which is fixed by its location within the overall scheme of the body. None of the members can be excluded or removed without serious damage to the whole. The security of the body politic, therefore, can only be maintained by means of a joint commitment to a public good which benefits every part without distinction, so that "each individual may be likened to a part of the others reciprocally and each believes what is to his own advantage to be determined by that which he recognizes to be most useful for others" (John of Salisbury 1990: 126). John's political body is one in which, beyond all social differentiation, there is "mutual charity reigning everywhere" (John of Salisbury 1990: 142), because all segments are attuned to the same enduring common purpose which encompasses the valid interests of the whole.

How does religious faith and ritual fit into John's model of the well-ordered body politic? He acknowledges that, like all bodies, the commonwealth is guided by a soul; specifically, the place of the soul belongs to "they who institute and mold the practice of religion, and who transmit the worship of God," namely, the priesthood (John of Salisbury 1990: 66). There must, in John's view, be a religious component to the political organism. Even in "superstitious" societies such as pagan Rome, where "Augustus Caesar himself was constantly subject to the sacred pontiffs" (John of Salisbury 1990: 67), religion plays a guiding role in public life. John clearly thinks that the priests of the Christian church, as representatives of the one true religion, ought to form the soul in the body politic of his own day. Yet his point of departure is ultimately neutral with regard to which religion directs the communal order.

Consequently, John's criteria for judging the welfare of the community remain primarily temporal ones. The presence of a cooperative spirit within diverse sections of the community stems from John's identification of the "health" of the body politic with the practice of justice by the organs and members. Indeed, his definition of the common good in terms of justice contains a recognition of a correlative obligation on the part of all members of the commonwealth: "Inasmuch as the duties (*officia*) of each individual are practiced so that provision is made for the corporate community, as long as justice is practiced, the ends of all are imbued with the sweetness of honey" (John of Salisbury 1990: 131). Every organ of the body must conduct itself according to the dictates of jus-

tice if the polity is to thrive as a corporate whole. John's conception of justice in public affairs applies to the physical good or harm of the body's members. Following Cicero's formula in *De officiis*, the *Policraticus* asserts

> Justice consists chiefly in this: do not do harm and prevent the doing of harm out of duty to humanity. When you do harm, you assent to injury. When you do not impede the doing of harm, you are a servant of injustice
>
> (John of Salisbury 1990: 62)

The essence of justice pertains to a responsibility toward the well-being of others; this responsibility entails not simply a negative obligation to refrain from the commission of injury, but a positive duty to protect others from harm as well. Since justice is the salient characteristic of the common good, it determines the manner in which each of the bodily members performs its functions. Hence, not only kings and magistrates, but even the lowliest of parts, are viewed by John as agents of public justice. This arises from the fact that the function of each segment of the organism is absolutely necessary for the welfare of all the others.

3. Marsiglio of Padua

During the middle of the thirteenth century, Aristotle's Nicomachean Ethics and Politics were translated into Latin and soon circulated widely throughout Europe. Within a few years, Aristotle's works had become the official textbooks for the study of political questions, albeit often supplemented by the also translated commentaries of Arab philosophers (such as Avicenna and Averroes) and the writings of Jewish thinkers (especially Moses Maimonides). This transition was facilitated by the rise of university education and a formal scholastic curriculum. Thereafter, the political theorists of the Latin Middle Ages acknowledged Aristotle as the supreme authority concerning matters of political life and governance, even when they ultimately disagreed with his conclusions. Aristotle, who had, in his youth, received training as a physician and who evinced a life-long concern with the role of nature, had famously advocated the doctrine that man was a political animal and that the state arose in order to facilitate a key aspect of human nature. This doctrine led to a nearly universal endorsement of some form of "political naturalism" among key scholastic thinkers of the late thirteenth and early fourteenth centuries (such as Thomas Aquinas, Giles of Rome, John of Paris, and Ptolemy of Lucca) and a concomitant reinforcement of the preceding tradition of drawing an analogy between the living organism and the political order (see Struve 1978 and Briguglia 2006).

Perhaps the most extensive use during the classic period of scholasticism of the organic comparison of the organization of political society to the physiology of an animate body—although by no means a wholly conventional one—can be found in the *Defensor pacis* (completed in 1324) by Marsiglio (also know as Marsilius, the Latinized version of his name) of Padua (1275/80–1342/43). Marsiglio's key theme was the threat to the peace of temporal communities posed by the claims of the pope and his supporters to forms of earthly jurisdiction and dominion. In order to make the point that the papacy had no business interfering in the "public" business of the secular realm, Marsiglio constructed in the first discourse of the *Defensor pacis* a detailed and extensive account of the foundations of communal order that owes little or nothing to spiritual

or even moral purposes. His training as a physician and his position at the University of Paris as, primarily, a teacher of natural philosophy, quite reflexively led him to employ organic doctrines in order to underscore his political lessons.

The analogy between community and organism is initially invoked in the *Defensor pacis* in order to define and distinguish between tranquility and intranquility. Marsiglio asserts that the basis for his comparison of the community to an animate body was their structural and functional similarity. Referring to Aristotle's references in the *Politics* concerning analogues between animate nature and political order, Marsiglio proposes a direct correspondence between organic model and political society:

> Just as an animal well disposed according to nature will be composed of given proportionate parts ordered reciprocally, each communicating its works to the others and to the whole, so the city (*civitas*) is constituted in such fashion when it is well disposed and instituted according to reason.
>
> (Marsilius of Padua 1956: I.2.3)

An animal is healthy when it behaves in accordance with its own proper and natural disposition, that is, when all its parts perform their appropriate tasks as dictated by the needs of the whole organism. Conversely, a diseased or unhealthy animal is one that lacks an overall coordination of the operations of its component parts, so that the natural balance inherent in the body is disturbed. By the same reasoning, tranquility in the political organism would also be the correct mutual ordering of the parts of the community, since the "health" of the civil body is its peace. In turn, intranquility may be viewed as "the diseased disposition of the city or kingdom, comparable to the infirmity of an animal, which impedes all or some parts from doing their tasks consistently, either in general or to completion" (Marsilius of Padua 1956: I.2.3). The organic metaphor is thereby introduced by Marsiglio as a diagnostic model, a way of delineating the good health of the body politic and of remedying the causes of political illness when it infects the organism.

Marsiglio conceives of the health of the political body in strikingly naturalistic and material terms, enshrining temporal advantage as a fundamental and entirely legitimate goal of human conduct. Indeed, he establishes "as the starting-point of everything to be demonstrated ... that all human beings not deformed or otherwise impeded naturally desire a sufficient life, and avoid or flee what is harmful thereto" (Marsilius of Padua 1956: I.4.2). Marsiglio cites with approval Cicero's proclamation in *De officiis* that the basic purpose of all living creatures is self-preservation. The advantage of human beings is achieved by gaining those conditions of existence that confer upon them a physically adequate life. Marsiglio makes fleeting mention of the Aristotelian conception of living well, constituted by the exercise of the practical and theoretical virtues, but the material sufficiency of human life receives the overwhelming measure of his attention.

Marsiglio formulates a philosophical reconstruction of the origins of human association and of government that serves as an explanation of both the purpose of civil life and its relation to human nature. The preservation of individuals is achieved most fully and naturally under conditions of human cooperation in the context of an ordered and organized community. Marsiglio holds that "human beings came together in the civil community in order to pursue advantage and a sufficient life and to avoid the opposite" (Marsilius of Padua 1956: I.12.7). The pursuit of mutual advantage grounds the account of the origins of communal life in the *Defensor pacis*. Specifically, successful self-preser-

vation demands perpetual interaction between human beings, each of whom makes a particular contribution through the exercise of his specialized task. Marsiglio expressly specifies these functions in terms of the occupations necessary to maintain the physical well-being of the community: farmers, merchants, craftsmen, and warriors. Since all of these jobs contribute essentially to achieving a sufficient temporal life, none are to be denigrated. Marsiglio insists that the intercommunication of functions required for the public welfare necessitates a politics of inclusion, and he thus construes citizenship in a remarkably extensive fashion.

Marsiglio intends citizens within the community to take an active role in their own governance, locating popular consent at the center of his theoretical framework. Such consent arises directly from the functional character of the community. First, all whose interests are served or affected by a community must agree to the conditions of association (that is, law and rulership). Second, having so consented, all such citizens are bound to obey the law and the determinations made by rulers in accordance with it. As a consequence, the *Defensor pacis* holds that the legitimacy of both laws and rulers depends wholly upon their "voluntary" character, that is, the extent to which those subject to their jurisdiction have publicly and overtly consented to their authority. What Marsiglio in fact regards to be distinctive of citizenship is the ability to express one's will in the political venue, by judging for oneself the validity of prospective rulers and laws. This implies for Marsiglio not merely formal, corporative assent, but an extensive privilege on the part of each individual citizen to examine prospective laws and rulers. Each and every member of the community reserves to himself final judgment about all matters of public regulation, realized by an act of explicit consent.

To explain the necessity of a unified citizenry maintaining control over the community, Marsiglio invokes a different version of the organic analogy to address how the body remains healthy in functional differentiation yet guards against disturbances of the communal peace. This organic account derives from Aristotle's treatise *De motibus animalium*, which purports to demonstrate that all animals are ordered in relation to a unique soul that fixes the goals for the sum of the organism's motions. Marsiglio modifies Aristotle's account of animal motion to explain how it is possible for a multiplicity of functions and activities to coexist and cooperate within the communal whole. In *Defensor pacis*, he associates the soul with the unified community of citizens. The corporate soul generates the sum of the offices within the civil body (including that of the government), it appoints particular occupants to the central office of the ruler, and it ultimately enforces the ruling part's conformity to and observance of the laws of the community (Marsilius of Padua 1956: I.15.6–8). The responsibilities of the soul within the political body are supervisory and regulatory, ensuring that the governors fulfill the terms of their office. This soul has but a single will and purpose—the sufficiency of life to which all humans aspire—reflected in public consent to law and government. Since the proper disposition of the animal requires that one and only one "central authority" can ultimately supervise the operation of the bodily parts, so, too, the organic metaphor enshrines a political arrangement that takes a single communal body (regardless of constitutional composition) as its focal point.

The force of the bodily analogy's application in the *Defensor pacis* is aimed toward supporting Marsiglio's primary polemical goal: resistance to secular incursions by papal power. If the goal of earthly society and government is the provision of the material necessities of life such as are provided by the different arts and functions, then the church and its officials have no business interfering in or directing secular affairs. Indeed,

the *Defensor pacis* counts the office of the priest as one of the "parts" of the city, and thus under the guidance and control of the "soul" of the body, the civil legislator itself. What the priesthood properly teaches concerns only matters supernatural and eternal, pertaining to the "inward" orientation of the immortal soul in its orientation toward God and salvation. Thus, priests who command actions that conflict with the laws or organization of the organic community overstep their proper and appropriate function. Otherwise, the harmonious equilibrium of all the members of the political body would be disturbed.

4. Nicole Oresme

Although he has long been acknowledged as one of the towering figures of fourteenth-century philosophy, it is only during the last generation or so that Nicole Oresme (1320/25–1382) has made any discernable impact among historians of political thought. Perhaps the reason for this is that his major contributions to medieval political ideas were nontraditional in language and genre: one, a sumptuously illuminated French-language translation of and commentary on Aristotle's *Politics* (*Le Livre de Politiques*); the other, a treatise of advice on monetary policy that seemed more a contribution to technical scholastic economics (*De moneta*). While Oresme certainly stands within the general tradition of late-medieval Aristotelianism, he was eclectic in his use of sources and distinctive in his interpretation and extension of Aristotle. Perhaps nowhere is the departure from strict adherence to Aristotle more evident than in Oresme's use of organic imagery in both of his main political writings.

Oresme seems to have touched initially upon this implication of organic thought in *De moneta*, probably written in 1356 or 1357. A brief tract that examines critically the practice of monetary alteration by governments, *De moneta* develops the case that money belongs to the community as a whole and therefore its value cannot be changed at the will of rulers. Indeed, Oresme holds that coin clipping and similar practices of devaluation constitute tyrannical acts that endanger the economic welfare of the realm. In order to demonstrate the difference in the order of the community when governed by a tyrant and by the true king, Oresme invokes in the penultimate chapter of *De moneta* the organic metaphor. He commences this exploration with a conventional reference to the "Plutarch" *Institutio Traiani*, known to him through John of Salisbury's *Policraticus*, which circulated widely during the fourteenth century. According to Oresme, Plutarch's letter to the Emperor Trajan shows how "the republic or kingdom is thus like a sort of human body" (Oresme 1957: 43), the condition of which depends upon the ends toward which the ruler directs it—justice in the whole or the sectarian good of the head or some other part.

However, Oresme immediately marries this traditional, head-oriented conception of the body politic to a passage from the fifth book of Aristotle's *Politics* and transforms it into a metaphor for the economic health of the community appropriate to his discussion of currency debasement. On Oresme's presentation, the body is construed as the standard for the distribution of wealth in the community. Each segment of society must have revenues sufficient for it to perform its function, and no member may grow rich at the expense of others.

> Just as the body, therefore, is badly disposed when the humors flow too freely into one member of it, so that that member is often inflamed and overgrown

> while others are withered and shrunken and due proportion is destroyed, and such a body cannot live long, so likewise is a community or kingdom when riches are attracted beyond measure (*ultra modum*) by one part of it.
>
> (Oresme 1957: 43)

In particular, Oresme is concerned that debasement by the monarch will lead to a sort of hydrocephalic condition within society.

> For a community or kingdom whose rulers increase their riches, power and station enormously in comparison with subjects is like a monster, like a man whose head is so large, so heavy, that it cannot be sustained by the rest of the body. And just as such a man cannot support himself or live long, so neither can a kingdom survive when its ruler draws to himself excessive riches, as is done by the mutation of money.
>
> (Oresme 1957: 43–4)

The common good marked by the balanced and reciprocal relation between the parts is viewed in terms of the income received by each. Oresme apparently takes for granted that every functional part of the body is deserving of a "natural" (or perhaps normal) measure of revenue. When one segment profits at the expense of others, the organic process breaks down. Likewise, presumably, the good health and longevity of the body is promoted and enhanced when every member receives its due. Oresme's lesson is that the superior authority of rulers is abused when they interfere with the "natural" operations of the economic constitution of the social body. The organism ruled by the true king thus enjoys economic equilibrium; that governed by a tyrant enjoys no such internal balance.

Oresme renews the application of the organic vision of the political origins of communal equilibrium in his French edition of and commentary on Book Five of the *Politics*, which he probably composed about fifteen years after *De moneta*. Oresme's discussion of the problem of revolutionary political change departs markedly from Aristotle's text, mainly in directing attention toward the sources of public dissent in royal regimes, whereas the original concentrated on oligarchic and democratic constitutions. Broadly speaking, Oresme shares Aristotle's view that the main reason for political upheaval is inequality, in particular the belief that social rewards are unjustly distributed. But where Aristotle had regarded such grievances as reflective of a profound misunderstanding of justice itself on the part of both democrats and oligarchs, rather than as a statement of legitimate protest, Oresme apparently holds that rulers ought to heed the complaints of their subjects. In particular, he focuses on economic injustices to which members of the community rightly object, such as when a prince "imposes taxes or exactions upon his people or lays hold of the fiscal revenues and the public treasury in order that they be applied to his own private profit and not to the common business" (Oresme 1970: 208). Moreover, rulers must take care to regulate the economic and social power of those under their control to ensure the maintenance of balance, following the principle of the organic body, lest disproportion engender discontent and sedition.

In this connection, Oresme turns directly to the organic principle of communal order that he had introduced in *De moneta* and expands it at considerable length. He begins once again with reference to the *Policraticus*, presenting a sketch of the full body, ranging from the royal head, through the eyes and ears of the judges, the senatorial heart, the hands of the soldiers, down to the feet composed of the laborers. Oresme's account, while slightly less detailed than John's, captures its hierarchical overtones. But having

established the design of the body politic, Oresme reasserts the principle of organic health as harmonious distribution and circulation of corporeal goods:

> A body is badly disposed when one of the members attracts to itself too much of the nourishment and the humors ... Similarly, the polity is badly ordered and cannot endure long when, whether through taxes or bad agreements or laws that are badly made or badly enforced, one of the members of the body takes too much nourishment into itself, because such a member is too large beyond just measure. And such a polity is like a monster and like a sick body.
>
> (Oresme 1970: 209)

All parts of the organism must share in sufficient nourishment in order to survive. This does not, of course, mean that all must have the same amount, but it does imply that no body can survive which starves some of its members while gorging others. The natural principle of equalization likewise applies to the animate community.

The deployment of the organic image in *Le livre de politiques* thus echoes the central themes that Oresme had emphasized in *De moneta*: the dangers of inequality to the rule of a prince and the order of a community. The former work, however, broadens out the application of the basic principle: it is not simply the disproportionate possession of money, but any inequality in the distribution of public goods, that threatens to harm the society. *De moneta* had concentrated primarily on the ruler's manipulation of his office in order to profit himself personally at the expense of his subjects. Oresme's commentary acknowledges that this remains a problem, but it also evinces concern about the destruction of organic reciprocity that occasions any unequal distribution of primary social goods. Indeed, a crucial role of any ruler who wishes to maintain his own government without (legitimate) dissention must be the equitable regulation of the assignment of wealth and power in society. One part of the community controlling too great a stake in the whole cannot help but endanger the harmony of functions that characterizes political order.

5. Conclusion: Order and Function

John of Salisbury, Marsiglio of Padua, and Nicole Oresme afford clear examples of the myriad ways in which medieval thinkers conceived of public order and the common good while still employing a shared vocabulary and conceptual framework. Many other authors could also be cited to illustrate the simultaneous diversity and coherence of medieval political thought. A more extended version of this analysis would include later medieval writers such as Christine de Pizan (who constructed an entire treatise around the body politic metaphor), Nicholas of Cusa, and John Fortescue, as well as twelfth-century masters such as Bernard Sylvester and Robert Pullan, not to mention a large host of Parisian schoolmen. Even with the proliferation of uses to which the body metaphor was put during the Middle Ages, however, the invocation of the organic metaphor depended upon the priority of reciprocity, the principle that all parts of the body share a single public good that can only be achieved when each member orients itself toward the welfare of the whole. Such "communal functionalism," as it has been labeled (Nederman 1992), pervades the entire range of organic theories and provides the framework for their commonality. Communal functionalists of all stripes emphasize that the duties and tasks necessary for the health of the body must be accorded some specified place in the coordination of the common good.

Yet the question of precisely how the segments of the body interrelate remained contentious in communal functionalist thought, as we have seen. Generally speaking, some communal functionalists tended to emphasize the direction of the head as the central source of reciprocal cooperation: the head directs the limbs and organs of the creature toward the attainment of a collective good that none can achieve on its own. By contrast, the equalizing interpretation of the communal functionalist principle inclined to view the head as dependent upon the assent of the bodily parts and looked to homeostasis to coordinate between the members. The head therefore occupies a secondary role in determining the substance of the common good and how it is to be attained: the ruler is more a traffic cop than an engineer. Communal functionalism was sufficiently capacious to permit the development of quite divergent conceptions of how public life ought to be organized. Within the broad framework of medieval organic political theory, with its guiding principle of reciprocity, there remained plenty of room for deep dispute about the fundamentals of political order.

Related Topics

Aristotle's Social and Political Philosophy, Aquinas, Natural Law and Rights Theory, Religion in Public Life

References

Briguglia, G. (2006) *Il corpo vivente dello Stato: Une metafora politica*, Milan: Bruno Mondadori.

Burns, J.H. with M. Goldie (1988) *The Cambridge History of Political Thought 1450–1700*, Cambridge: Cambridge University Press.

Constable, G. (1995) *Three Studies in Medieval Religious and Social Thought*, Cambridge: Cambridge University Press.

John of Salisbury (1990) *Policraticus: Of the Frivolities of Courtiers and the Footprints of Philosophers*, ed. and trans. C.J. Nederman, Cambridge: Cambridge University Press.

Marsilius of Padua (1956) *The Defender of Peace*, trans. A. Gewirth, New York: Columbia University Press.

Moore, R.I. (2000) *The First European Revolution: 970–1215*, Oxford: Blackwell.

Nederman, C.J. (1992) "Freedom, Community and Function: Communitarian Lessons of Medieval Political Theory," *American Political Science Review* 86: 977–86.

—— (2009) *Lineages of European Political Thought: Explorations Along the Medieval/Modern Divide from John of Salisbury to Hegel*, Washington, DC: Catholic University of America Press.

Oresme, N. (1957) *De moneta*, trans. C. Johnson, London: Thomas Nelson and Sons.

—— (1970) *Li Livre de Politiques de Aristote*, ed. A.D. Minut, Philadelphia: American Philosophical Society.

Struve, T. (1978) *Die Entwicklung der organologischen Staatsauffassung im Mittelalter*, Stuttgart: Hiersemann.

Further Reading

There are three recent surveys of medieval political ideas that are all commendable. A. Black, *Political Thought in Europe 1250–1450* (Cambridge: Cambridge University Press, 1992) emphasizes the multiple languages employed by medieval authors. J. Canning, *A History of Medieval Political Thought 300–1450* (London: Routledge, 1996) is especially useful on questions of church–state conflicts during the Middle Ages. J. Coleman, *A History of Political Thought: From the Middle Ages to the Renaissance* (Oxford: Blackwell, 2000), while weak on the period before the thirteenth century, is the most comprehensive for the later Middle Ages. J.H. Burns, ed., *The Cambridge History of Medieval Political Thought c.350–c.1450* (Cambridge: Cambridge University Press, 1988) covers the Middle Ages in greater detail, although some of the chapters are now somewhat out of date.

5
MACHIAVELLI

Vickie B. Sullivan

Born in 1469, Niccolò Machiavelli was a child of the Italian Renaissance as it flourished in Florence, but the influence that his writings exerted far transcended his time and place. He was a boy when the Pazzi conspiracy attempted to rid Florence of Medicean rule but succeeded in assassinating only Giuliano de'Medici, leaving his brother Lorenzo alive. Patron of great artists and poets, Lorenzo the Magnificent dominated the city's political life until his death in 1492. His son, Piero inherited power in Florence, but when King Charles VIII of France invaded Italy in 1494 to lay claim to Naples, resistance to Piero's rule grew to such an extent that he eventually fled the city. As a result of this French invasion, Florence became a republic. At first, the new Florentine republic was under the sway of Friar Girolamo Savonarola, but after his execution as a heretic in 1498, the republic lost its theological fervor. At this point, Machiavelli's life in service to the republic began. He served as Florence's secretary of the Second Chancery and of the Ten of War and represented the republic on diplomatic missions to France and to the Emperor, for example. When in 1512, the Medici returned to rule his native city, Machiavelli's life changed dramatically. No longer was his native city a republic and no more was he to be its civil servant. To compound his troubles, because he was suspected of conspiring against the restored Medici, he was arrested and tortured. After his release, he returned to his farm in Sant'Andrea in Percussina, just outside of Florence.

It was there, during Machiavelli's expulsion from political life, that he wrote the works that would make his name endure. After great effort, he won enough confidence from the Medici, who not only controlled Florence but who had also accumulated power in Rome, for Cardinal Giulio de'Medici to commission him to write the *Florentine Histories* for presentation to the Medicean pope, Leo X. Machiavelli presented the final product to his patron in 1525 when Giulio had become Pope Clement VII.

Machiavelli published only three works during his lifetime: a poem, the *First Decennial*; a comic play, *Mandragola*; and a dialogue, the *Art of War*. His two most famous and influential works are *The Prince*, a short treatise, which he wrote in 1513, and the *Discourses on Livy*, a much longer work focusing primarily but not exclusively on the history of the Roman republic, which was written in the latter part of that decade. Neither was published until about five years after his death in 1527. His entire corpus was proscribed by the first papal Index of Prohibited Books of 1557 and confirmed by the second in 1559 (Anglo 2005: 172, 6).

1. Readers and Interpreters of His Thought

This proscription by the Papacy did not prevent Machiavelli's works from having a wide readership among subsequent political philosophers and thinkers, including Francis Bacon, James Harrington, Algernon Sidney, Benedict Spinoza, Montesquieu, Jean-Jacques Rousseau, David Hume, and Friedrich Nietzsche. In addition, his readers and commentators have included such political leaders as Emperor Charles V, Frederick II of Prussia, John Adams, and Benito Mussolini. The interest of political leaders in his work seems natural enough as Machiavelli declares his intention to give leaders practical advice on the ends and means of rule. The character of that very advice, though, has attracted philosophers and scholars because they judge it as representing a pivotal point in the history of Western thought, although they differ fundamentally in their assessment of its meaning and significance. *The Prince* offers "the means whereby religion, piety and every natural inclination of virtue can be easily destroyed," and its author is "an enemy of the human race" and "a son of Satan" (Pole 1997: 275); "we are much beholden to Machiavel" for declaring "what men do, and not what they ought to do" (Bacon 1920: §2.21.9, 201); by "pretending to give lessons to Kings," Machiavelli conveyed "great ones to the people" and therefore, his notorious treatise *The Prince* is, in fact, "the book of republicans" (Rousseau 1994: §3.6, 177); he is the purveyor in the modern world of the notion of raison d'état, which declares that the state must be protected even at the cost of the subversion of law and of conventional morality (Meinecke 1957); he broke with all previous political philosophy to found modern political philosophy (Strauss 1989: 83–4); he played a critical role in the history of thought by transmitting elements of classical and Renaissance philosophy in the form of classical republicanism to subsequent thinkers (Pocock 1975; Skinner 1998); he rejected the pronouncements of philosophy—as well as of Christianity—to embrace the rapacious imperialism that the ancient Roman republic represents with its concomitant elevation of action over contemplation (Coby 1999; see also Hulliung 1983).

While some write to condemn and others to condone Machiavelli, all commentators must confront his reputation as an immoralist—whether that means acknowledging it or rejecting it. Indeed, his very name came into conventional discourse to signify the type of individual who uses devious and immoral means to achieve his or her own immediate interests. For instance, this implication is clear when one of Shakespeare's characters asks, "Am I politic? am I subtle? am I a Machiavel?" in an attempt to exonerate himself from so disreputable an association (*Merry Wives of Windsor* 3.1.91–2). Machiavelli excused—indeed, praised—great founders of states, who used nefarious means for the most exalted of purposes with the maxim: "when the deed accuses him, the effect excuses him" (Machiavelli 1996: §1.9, 29). It is not clear that Machiavelli would similarly sanction the use of such methods by private individuals in an established society to attain their particular private interests; nevertheless, his comic play *Mandragola* does feature a character who benefits not only himself but others by enlisting a corrupt friar in subverting the marriage vows of a hitherto virtuous matron.

2. Misfortune as Opportunity, His Teaching on Human Activity, and Machiavelli's Knowledge

In prominent places in both *The Prince* and the *Discourses*, Machiavelli expresses abiding dissatisfaction—with his own personal fate, with an Italian peninsula divided and

conquered, and also with the historical situation broadly understood as he ascertains in contemporary times a marked decline in political and military virtue from the height of Roman antiquity. Despite his often ostentatious lamentation over these various misfortunes, his works quite prominently instruct that the circumstances that others would term unfortunate or bad luck must be viewed positively because they present an opportunity for ambitious political leaders to magnify their greatness in overcoming hardships. In analyzing how human beings might overcome such circumstances, Machiavelli broaches issues of free will and human autonomy in the face of extraneous forces—whether divine or natural. Moreover, in contemplating the passivity that results from belief in such omnipotent forces, he endeavors to inspirit his readers to take control of their destinies. Given this teaching regarding the ultimate usefulness of adversity, his many prominent laments suggest that his times are propitious for innovation and its attendant glory (cf. Ridolfi 1963: 219; Viroli 2000: 215). He announces that he will bring his knowledge, deriving from things both ancient and modern, to bear on his instruction of leaders and potential leaders in undertaking to acquire that glory.

The epistle dedicatory of *The Prince* highlights his agony over his own circumstances as he presents himself as an abject office seeker and ends with an appeal to the current Medicean ruler in Florence: "if your Magnificence will at some time turn your eyes from the summit of your height to these low places, you will learn how undeservedly I endure a great and continuous malignity of fortune." Here, Machiavelli presents himself as a victim of fortune, but he also expresses his hope that his authorship of this treatise is a means by which to overcome that situation (Machiavelli 1985: 4; see also his letter to Francesco Vettori, December 10, 1513 in Machiavelli 1985: 107–11).

Whereas the preface to *The Prince* laments his own personal hardships, the preface to the first book of the *Discourses* laments a much broader misfortune—that is, what he regards as modern political weakness. In his times, Machiavelli writes, political and military actors do not even attempt to make use of those ancient methods, whose memories remain. In explaining the cause for this lapse, he detects a general malaise. For instance, he refers in this context to "the weakness into which the present religion has led the world," but finally attributes the gulf between ancient and modern practice to the belief that imitation of the ancients is impossible. In response to this belief, he will comment on all those books of Livy "that have not been intercepted by the malignity of the times" (Machiavelli 1996, preface: 5–6). His mention of the malignity of the times refers to the ascendancy of Christianity, a point that he makes quite clear later in the *Discourses* when he notes that the incomplete record of pagan Rome available in his own times is due to the efforts of the early Christians, namely "Saint Gregory," who attempted but failed to obliterate completely the memory of pagan times by "burning the works of the poets and historians," including those of Livy (Machiavelli 1996: §2.5, 139). Therefore, he links the misfortune of his times to the reign of Christianity.

His times are, indeed, lamentable, but Machiavelli teaches that misfortune can and should be regarded as opportunity. His four great exemplars of *The Prince*, Theseus, Moses, Cyrus, and Romulus, founded in Machiavelli's words "new modes and orders." In speaking of their accomplishments, he emphasizes the hardships they faced at the outset of their undertakings: Moses's people were enslaved in Egypt; those of Theseus, founder of Athens, were "dispersed"; those of Cyrus were under the rule of the Medes; and Romulus was exposed at birth and not received in Alba. These circumstances worked to their advantage, however: "And as one examines their actions and lives, one does not see that they had anything else from fortune than the opportunity, which gave them the

matter enabling them to introduce any form they pleased." Of course, it is not merely unfortunate circumstances that allowed these great founders to succeed; they had to recognize the opportunity and to capitalize on it: "Without that opportunity their virtue of spirit would have been eliminated and without that virtue the opportunity would have come in vain" (Machiavelli 1985: §6, 23). Their circumstances constituted an opportunity to be seized.

In this way, Machiavelli encourages human beings to take their earthly lives seriously. In the famous chapter 25 of *The Prince*, entitled "How Much Fortune Can Do in Human Affairs, and in What Mode It May Be Opposed," he endeavors to vindicate the free will of human beings as a means of curing the general malaise of his times. Here he describes that malaise quite precisely:

> It is not unknown to me that many have held and hold the opinion that worldly things are so governed by fortune and by God, that men cannot correct them with their prudence, indeed that they have no remedy at all; and on account of this they might judge that one need not sweat much over things but let oneself be governed by chance.

To overcome the resignation that beliefs in omnipotent and transcendent forces foster he rouses human beings to action:

> in order that our free will not be eliminated, I judge that it might be true that fortune is arbiter of half of our actions, but also that she leaves the other half, or close to it, for us to govern.

In a most emphatic fashion, Machiavelli encourages human beings to take their fate in their own hands when he makes his notorious comment that because fortune is a woman, "she is the friend of the young, because they are less cautious, more ferocious, and command her with more audacity" (Machiavelli 1985: §25, 98 and 101). Machiavelli's overt message is to conquer fortune.

In considering Machiavelli's injunction to conquer fortune, one should recall that he depicts himself very much as a victim of the malignity of fortune. His own misfortune, then, puts him paradoxically in favorable circumstances, akin to his four great founders, as these unfavorable circumstances, over which he agonizes, would furnish him with an opportunity. His resolve to seize this opportunity by helping to ameliorate the situation of the moderns is evident in the preface to the first book of the *Discourses* when, after describing their deplorable situation, he says that he will comment on Livy's history in order to foster a "greater understanding" of political affairs. It is his knowledge that can foster this change, and he specifies that it is not only his knowledge of "ancient" but also of "modern things" (Machiavelli 1996, preface: 6). Thus, one finds that the work discusses not only ancient events but also contemporary ones. Similarly, he makes this combination of ancient and modern knowledge prominent in *The Prince* when he declares that he will share with a prince his knowledge, which derives from his "long experience with modern things and a continuous reading of ancient ones" (Machiavelli 1985: 3). His emphasis on this combination of sources of knowledge strongly suggests that although he seeks a return of sorts to antiquity, it is not an unmitigated return but, rather, one informed fundamentally by his analysis of the corruption of modern times, its origin, and the manner by which it can be overcome (Sullivan 1996). Although

expelled from Florentine politics, his life as a writer, the life to which bad fortune delivered him, is also his opportunity, as his writings convey to others his discoveries regarding the proper response to the broader malignity of the times and make him the teacher of captains, rulers, and political thinkers of the future.

3. The Christian Church and the Misfortune of His Times

Machiavelli also laments the condition of Italy, and his call to conquer fortune in *The Prince* is entwined with his analysis of "Why the Princes of Italy Have Lost Their States" as well as his famous "Exhortation to Seize Italy and to Free Her From the Barbarians" (Machiavelli 1985: §24, 96 and §26, 101). His writings blame the Christian Church for Italy's abject condition.

In chapter 1.12 of the *Discourses* he lodges two particular complaints against the Church: first, that the presence of the papal government in Rome keeps Italy divided; and second, that the corruption of the Churchmen harms their flock. In treating the first charge, Machiavelli emphasizes that the Church maintains "a temporal empire there." His analysis maintains that the Church has been incapable of unifying Italy under its own aegis because it lacks its own arms; nevertheless, it is strong enough to oppose any other temporal power that endeavors to control and unite the province, a result it fears because such a unifying power would cause it to lose its "dominion over its temporal things." The Church's presence in Rome, then, prevents the unification of Italy: "The church"

> has been the cause that [Italy] has not been able to come under one head but has been under many princes and lords, from whom so much disunion and so much weakness have arisen that it has been led to be the prey not only of barbarian powers but of whoever assaults it.
>
> (Machiavelli 1996: §1.12, 38; see also Machiavelli 1988: §1.9, 20 and §1.23, 34)

He regrets this situation: "truly no province has ever been united or happy unless it has all come under obedience to one republic or to one prince" (Machiavelli 1996: §1.12, 38). The Church is responsible for Italy's unhappiness.

In addition to condemning the Church for fostering Italy's division, he also denounces in this chapter and elsewhere what he sees as the harmful character of the corruption of the clergy. In *Discourses* 1.12 he specifies that "because of the wicked examples of that court, this province has lost all devotion and all religion—which brings with it infinite inconveniences and infinite disorders" (Machiavelli 1996: §1.12, 38). Here Machiavelli claims that the very institution that is charged with inculcating devotion to religious values actually undermines them. Elsewhere in the *Discourses* Machiavelli elaborates on that very corruption to declare that the bad behavior of the Churchmen results from their lack of belief in divine retribution: "they do the worst they can because they do not fear the punishment that they do not see and do not believe" (Machiavelli 1996: §3.1, 212). They may preach the horrors of eternal damnation, but they do not believe their own sermons. Instead, the fear they engender in others is the instrument by which they maintain their noxious rule.

In more than one passage in the *Discourses*, Machiavelli actually contemplates the possibility of the demise of Christianity's reign and with it that of the Churchmen.

At one point he says that "our religion" "would be altogether eliminated" if it had not been for the intervention of Saint Francis and Saint Dominick, who "with poverty and with the example of the life of Christ ... brought back into the minds of men what had already been eliminated there." Their actions allowed Christianity to find new life: "Their new orders were so powerful that they are the cause that the dishonesty of the prelates and of the heads of the religion do not ruin it." Moreover, Machiavelli declares that that very perseverance harms the population of Christian states by allowing the corruption of the clergy to continue unchecked:

> Living still in poverty and having so much credit with peoples in confessions and sermons, they give them to understand that it is evil to say evil of evil, and that it is good to live under obedience to them and, if they make an error, to leave them for God to punish.
>
> (Machiavelli 1996: §3.1, 211–12; see also §2.5, 138–9)

In this clever sentence, the evil, of which it is evil to speak, is the evil that the clergy commits. Although responsible for evil, in Machiavelli's view, the members of the clergy have convinced others that they must not even speak of it, and as he informs us, the Churchmen do not believe in God's punishment, but the members of their flock do. Machiavelli's criticism of the Church is shocking for its sarcastic boldness.

He also sees fit to criticize Christianity for rendering "men less strong" than they were in pagan times, because modern people "esteem less the honor of the world" as they seek after the "paradise" of Christian teaching. He then seems to soften the sharpness of his criticism of the harmful effects of Christian doctrine by asserting that an alternative interpretation of the religion could permit "the exaltation and defense of the fatherland" (Machiavelli 1996: §2.2, 131–2). A political outcome is the touchstone by which he evaluates moral and religious teachings.

4. Machiavelli's Teaching on Morality

The malignity of the times—the source of his dissatisfaction but also of his challenge, and hence, of others' potential greatness and also his own—arises from the modern education, which hobbles the ability of human beings to respond effectively to political necessities. In order to usher in a more effective politics, he needs to counter that enfeebling education. Machiavelli teaches that politics demands a resolution to act in addition to an ability to deceive, a hard heart, and a sharpened sword. His examination of the disastrous consequences of traditional understandings of the demands of morality combined with his recommendations for much more resolute actions form his new teaching on morality, which he promises will bring much more desirable political results.

In *The Prince* he quite boldly describes the source of the problem he has identified:

> [M]any have imagined republics and principalities that have never been seen or known to exist in truth; for it is so far from how one lives to how one should live that he who lets go of what is done for what should be done learns his ruin rather than his preservation.

In speaking here of those prior thinkers who have offered imagined republics and principalities, he engages the traditions of classical philosophy, such as Plato's *Republic*,

for example, and Christian morality, such as Augustine's *City of God*, to declare that political ruin is the result of their imaginings (Mansfield 1985: x–xi). This is a most damning judgment because politics, and the existence that it secures in this world, is the only criterion Machiavelli accepts for determining and judging actions. Indeed, in the same passage, he explicitly makes the outcome the touchstone for judgment, teaching that "if one considers everything well, one will find something appears to be virtue, which if pursued would be one's ruin, and something else appears to be vice, which if pursued results in one's security and well-being." A ruler must "learn to be able not to be good," he declares, because the goodness taught by these authoritative traditions enfeebles political actors (Machiavelli 1985: §15, 61–2). Unlike scriptural and philosophical teachings, which not only acknowledge but teach that often the wicked prosper and the good do not (Berlin 1980: 26), Machiavelli declares instead, in the words of one commentator, that "[t]here is no gap between the demands of politics and the demands of virtue" because "[w]hatever politics demands is virtuous" (Orwin 1978: 1218).

Indeed, Machiavelli insists that "works and … intentions" have "to be judged by the end" and that, as a result, a leader must focus on attaining that end and ignore entirely the terms of condemnation others might use to decry the leader's means (Machiavelli 1996: §3.3, 215). What Machiavelli regards as the demands of politics thoroughly shapes his moral understanding. This result is evident in another chapter of *The Prince* when he informs a ruler, or a potential one, that there are such things as "cruelties" "well used," which he defines as cruelties that "are done at a stroke, out of the necessity to secure oneself, and then are not persisted in but are turned to as much utility for the subjects as one can" (Machiavelli 1985: §8, 37–8). Machiavelli finds that his contemporaries recoil from this truth and political disaster is the result.

This state of affairs is quite evident when he considers in the *Discourses* the way in which a republic should handle a subject city rent by factions. Three alternatives exist, he explains: the ruling city can "kill the heads of the tumults"; remove the leaders from the city; "or … make them make peace together under obligations not to offend one another." He reports that his native city always attempted the third alternative. Although Florence wished to foster peace, precisely the opposite resulted from its pursuit of that end, as he declares that "always greater tumults and greater scandals arose from it." He insists that the first alternative is the best, pointing to the success the ancient Romans had with this policy. By contrast, contemporary republics fail to follow the Roman example. Reflecting on this failure, he explains that "[b]ecause such executions have in them something of the great and generous, however, a weak republic does not know how to do them." This harshest approach is great and generous, in his estimation, and he laments that his republican contemporaries cannot bring themselves to employ cruelties well used. He also reflects in this chapter on the reason for the weakness that prevents his contemporaries from committing such acts that he views as both great and generous. He concludes that "the weakness of men at present, caused by their weak education and their slight knowledge of things, makes them judge ancient judgments in part inhuman, in part impossible" (Machiavelli 1996: §3.27, 274–5). The difference between the education of his contemporaries and that of the ancient Romans resides, of course, in the fact that the tenets of Christianity inform the modern education. Christianity teaches human beings to seek peace, which is precisely what the Florentines endeavored in their misplaced efforts to reconcile warring factions. He teaches that the harshest political measures can sometimes bring the most favorable results, and in the process he offers an antidote to the modern moral education.

5. The Harshest Examples as Models

Machiavelli's political thought is replete with historical examples of leaders taking extreme actions. These examples further instruct leaders on the demands of a vigorous and effective politics.

No harsher political measure exists than that taken by the first Roman Brutus. Brutus's sons conspired to overthrow the fledgling republic in order to bring back the privileges they had enjoyed under the kings. Machiavelli acknowledges that Brutus's deed "is an example rare in all memories of things to see the father sit on the tribunals and not only condemn his sons to death but be present at their death." Although Machiavelli notes how rare such a deed is, he insists on its necessity: "This will always be known by those who read of ancient things: that after a change of state, either from republic to tyranny or from tyranny to republic, a memorable execution against the enemies of present conditions is necessary" (Machiavelli 1996: §3.3, 214). In Machiavelli's hands, Brutus's deed becomes a prescription by which to treat a perennial political problem: "there is no remedy more powerful, nor more valid, more secure, and more necessary, than to kill the sons of Brutus" (Machiavelli 1996: §1.16, 45).

He also uses the example of Brutus's deed when addressing the necessity of overcoming complacency in a republic. He maintains that a republic must be drawn "back toward its beginning" by means of examples of harsh punishment that are so "excessive and notable" that they induce fear in citizens. Such terrifying punishments ensure that citizens do not become complacent in their obedience of the republic's laws. Among such efficacious punishments he lists Brutus's execution of his sons (Machiavelli 1996: §3.1, 210).

In addition, his *Discourses* highlight the crimes of one of the great founders of *The Prince*, Romulus. Legend tells of Romulus's slaying of his brother Remus when the brothers were competing to decide who should found and name the new city. Machiavelli, though, adds to Romulus's guilt in making Rome's founder directly responsible also for the later death of his colleague in rule, the Sabine king Titus Tatius. Although in his rendition Romulus's crimes are even greater than those which his historical source describes (Sullivan 1996: 127), he will excuse him. It is in this context that he offers his famous statement: "When the deed accuses him, the effect excuses him." He continues, "when the effect is good, as was that of Romulus, it will always excuse the deed; for he who is violent to spoil, not he who is violent to mend, should be reproved." Indeed, Machiavelli specifies that Romulus's action of organizing the Senate shows that "what he did was for the common good and not for his own ambition" (Machiavelli 1996: §1.9, 29). Of course, Romulus's actions for the common good also inured to his own benefit as they made him the founder of a great city, whom Machiavelli glorifies.

6. The Dilemma of the Good Man and the Bad

Judgment dictates that these harshest methods be used only for good ends—that these cruelties be well used. Machiavelli must educate his readers' judgment regarding these means and these ends because he finds an obstacle to their correct use:

> Because the reordering of a city for a political way of life presupposes a good man, and becoming prince of a republic by violence presupposes a bad man, one will find that it very rarely happens that someone good wishes to become

> prince by bad ways, even though his end be good, and that someone wicked, having become prince, wishes to work well, and that it will ever occur to his mind to use well the authority that he has acquired badly.
>
> (Machiavelli 1996: §1.18, 51)

Neither the good man nor the bad man can overcome the divide to use the correct means for the desired ends. Both need Machiavelli's instruction.

Wicked men fail to see how crimes can at once further the common good and win their perpetual fame. Of course, Machiavelli's times were replete with disreputable characters who seemed to have no need of instruction in the necessity of eschewing morality. The notorious Borgias come to mind, and, indeed, Pope Alexander VI and his son Cesare Borgia are prominent characters in *The Prince*. Even these wicked men are in need of Machiavelli's instruction, though (Scott and Sullivan 1994).

Machiavelli recounts the tale of another bad man of his times, Giovampagolo Baglioni, "tyrant of Perugia," who, akin to all Machiavelli's great founders, was offered a most enviable opportunity, but unlike Machiavelli's heroes, he failed to capitalize on it. Machiavelli describes Baglioni as "a villainous man, who was taking his sister for himself, who had killed his cousins and nephews so as to reign." This bad man's opportunity came when Pope Julius II, intending to overthrow Baglioni, impetuously entered Perugia before his armies had arrived to guard him. Machiavelli comments that the

> rashness of the pope and the cowardice of Giovampagolo were noted by the prudent men who were with the pope, and they were unable to guess whence it came that he did not, to his perpetual fame, crush his enemy at a stroke and enrich himself with booty, since with the pope were all the cardinals with all their delights.

Machiavelli clearly believes that Baglioni should have mustered all of his wickedness in order to slay the pope:

> So Giovampagolo, who did not mind being incestuous and a public parricide, did not know how—or, to say better, did not dare, when he had just the opportunity for it—to engage in an enterprise in which everyone would have admired his spirit and that would have left an eternal memory of himself as being the first who had demonstrated to the prelates how little is to be esteemed whoever lives and reigns as they do; and he would have done a thing whose greatness would have surpassed all infamy, every danger, that could have proceeded from it.

Machiavelli openly and shockingly declares, then, that such an assassination would have been a great deed that would have resounded to Baglioni's "perpetual fame." Although Baglioni possessed the attributes of a bad man, he lacked the acumen to recognize how his capacity for villainy could be used to benefit both himself and others: "But it was concluded that it arose from men's not knowing how to be honorably wicked or perfectly good; and when malice has greatness or is generous in some part, they do not know how to enter into it" (Machiavelli 1996: §1.27, 62–3; see Machiavelli 1988: §6.29 for a similar opportunity that although recognized, was poorly executed).

Conversely, good men may have the intention to seek the political good, but they lack the ability to take the harsh measures to secure their cities. Of such a good man

who fails and ruins his city, he provides the example of his own boss in republican Florence, Piero Soderini, who served as the gonfalonier of the republic from 1502 until its collapse, and who, Machiavelli explains, shied away from necessary and harsh action. Machiavelli certainly regards him as a good man, but one who "did not know that" "goodness is not enough" (Machiavelli 1996: §3.31, 280). Machiavelli comments that Soderini, in fact, recognized "the appetite that was in the sons of Brutus for returning to another government." He "deceived himself," however, because he thought he could "overcome" their opposition "with his patience and goodness." This was a mistake that led to the republic's downfall, and Machiavelli comments that "he should never allow an evil to run loose out of respect for a good, when that could easily be crushed by that evil." Soderini might have been another Romulus or Brutus, who took bad means for good ends and won perpetual fame as a result, but, instead, Soderini ruined his republic (Machiavelli 1996: §3.3, 214–15). Machiavelli's writings educate regarding the desired ends—security for the people and glory for the leader and founder—and the required means, the means of a bad man turned to these good ends.

7. Martial Republicanism

Machiavelli's readers have long reflected on whether his allegiance lies with republics or principalities, as the *Discourses* seems devoted to republics and *The Prince* to sole rule, but a salient theme that unites the two works is Machiavelli's enthusiastic endorsement of territorial expansion. In both works, he advocates for the desirability and necessity of armed increase and points to the superior acquisitive abilities of republics.

Chapter 3 of *The Prince*, for instance, is devoted to a consideration of the invasions of Italy by two kings of France. He analyzes the mistakes that each king made in his enterprise and the manner by which the peninsula might be invaded and held in a more effective manner. In the context of his discussion of these attempts at armed conquest, he makes a general comment regarding human nature: "it is a very natural and ordinary thing to desire to acquire." The ancient Romans, he notes, were particularly effective at acting on this human desire as they conquered and maintained their acquisitions: "For the Romans did in these cases what all wise princes should do" (Machiavelli 1985: §3, 14 and 12). In holding up the Romans here for princes to imitate, he is, in fact, using the Roman republic as a model for wise princes.

Not only are republics worthy models for princes, in Machiavelli's view, but they actually far surpass principalities in their capacity for successful acquisition of territory. Indeed, when considering the territorial acquisitions made by Philip and his son Alexander the Great, Machiavelli turns to consider those that can be made by a republic:

> A republic should do so much more, as through the mode of electing it has not only two in succession but infinite most virtuous princes who are successors to one another. This virtuous succession will always exist in every well-ordered republic.
>
> (Machiavelli 1996: §1.20, 54)

Therefore, although a principality might be able to mimic a republic to some extent, that mode of government cannot surpass a well-ordered republic in long-term acquisitive prowess.

Machiavelli's interest is intensely focused on a martial republic capable of winning and maintaining territory. He establishes that interest early in the *Discourses* when he examines the various types of republics, those that are content to stay within their initial territory and those that are constituted to expand. After weighing the alternatives, Machiavelli decrees that both foreign relations and domestic considerations demand that a republic seek expansion (Machiavelli 1996: §1.6, 23; Sullivan 2004: 34–8).

The choice for empire, in turn, dictates, in Machiavelli's view, that such a republic will be characterized by tumults, as was the Roman republic. Such tumults are a result of the facts that a conquering army must be large and hence the republic requires a large population from which to draw its army (Machiavelli 1996: §2.3, 133–5), and that this large population will make demands for political representation to which the republic must accede. In the chapter of the *Discourses* that he entitles "That the Disunion of the Plebs and the Roman Senate Made That Republic Free and Powerful," Machiavelli declares, "I say that every city ought to have its modes with which the people can vent its ambition, and especially those cities that wish to avail themselves of the people in important things"—namely, the acquisition of an empire (Machiavelli 1996: §1.4, 16–17). Machiavelli advocates quite vociferously for a martial, and hence a democratic, republic.

Machiavelli's argument in favor of the conquering prowess of republics would hold particular appeal to later republican thinkers such as James Harrington during the English civil wars and their aftermath (Hulliung 1983: 4–6; cf. Pocock 1975: 386). Although the republic that Machiavelli explicitly endorses is harsh, often bloody, and unabashedly imperialistic, his republicanism would be a means by which commentators could excuse his reputation and rehabilitate the thought of the man whom some had decried as evil.

Related Topics

Plato's Political Philosophy, Aristotle's Social and Political Philosophy, Aquinas, Republicanism, Virtue Ethics and Political Philosophy, Natural Law and Rights Theory

References

Anglo, S. (2005) *Machiavelli—The First Century: Studies in Enthusiasm, Hostility, and Irrelevance*, Oxford: Oxford University Press.

Bacon, F. (1920) *The Advancement of Learning*, ed. W. A. Wright, Oxford, Clarendon Press.

Berlin, I. (1980) "The Originality of Machiavelli," in *Against the Current: Essays in the History of Ideas*, ed. H. Hardy, New York: Viking Press.

Coby, J. P. (1999) *Machiavelli's Romans: Liberty and Greatness in the "Discourses on Livy*," Lanham, MD: Lexington Books.

Hulliung, M. (1983) *Citizen Machiavelli*, Princeton, NJ: Princeton University Press.

Machiavelli, N. (1985) *The Prince*, trans. H. Mansfield, Chicago, IL: University of Chicago Press.

—— (1988) *Florentine Histories*, trans. L. Banfield and H. Mansfield, Princeton, NJ: Princeton University Press.

—— (1996) *Discourses on Livy*, trans. H. Mansfield and N. Tarcov, Chicago, IL: University of Chicago Press.

Mansfield, H. (1985) Introduction to *The Prince* by N. Machiavelli, trans. H. Mansfield, Chicago, IL: University of Chicago Press.

Meinecke, F. (1957) *Machiavellism: The Doctrine of Raison d'État and its Place in Modern History*, trans. D. Scott, London: Routledge and Kegan Paul.

Orwin, C. (1978) "Machiavelli's Unchristian Charity," *American Political Science Review* 72: 1217–28.

Pocock, J. G. A. (1975) *The Machiavellian Moment: Florentine Political Thought and the Atlantic Republican Tradition*, Princeton, NJ: Princeton University Press.
Pole, R. (1997) *Apology*, trans. N. Webb in *Cambridge Translations of Renaissance Texts*, vol. 2, ed. J. Kaye, Cambridge: Cambridge University Press.
Ridolfi, R. (1963) *Life of Niccolò Machiavelli*, trans. C. Grayson, Chicago, IL: University of Chicago Press.
Rousseau, J. J. (1994) *On the Social Contract* in *"Social Contract," "Discourse on the Virtue Most Necessary for a Hero," "Political Fragments," and "Geneva Manuscript,"* trans. J. R. Bush, R. D. Masters, and C. Kelly, vol. 4, *Collected Writings of Rousseau*.
Scott, J. T. and V. B. Sullivan (1994) "Patricide and the Plot of *The Prince*: Cesare Borgia and Machiavelli's Italy," *American Political Science Review* 88: 887–99.
Shakespeare, W. *Merry Wives of Windsor*.
Skinner, Q. (1998) *Liberty before Liberalism*, Cambridge: Cambridge University Press.
Strauss, L. (1989) "The Three Waves of Modernity," in *An Introduction to Political Philosophy: Ten Essays by Leo Strauss*, ed. H. Gilden, Detroit, MI: Wayne State University Press.
Sullivan, V. (1996) *Machiavelli's Three Romes: Religion, Human Liberty, and Politics Reformed*, DeKalb, IL: Northern Illinois University Press.
—— (2004) *Machiavelli, Hobbes, and the Formation of a Liberal Republicanism in England*, Cambridge: Cambridge University Press.
Viroli, M. (2000) *Niccolò's Smile: A Biography of Machiavelli*, trans. A. Shugaar. New York: Farrar, Straus and Giroux.

Further Reading

Contemporary scholarship on Machiavelli is riven by two very different approaches to reading Machiavelli's work, one textual and the other contextual, which have produced diametrically opposed conclusions as to the character of Machiavelli's thought and of its significance. For the textual approach, see L. Strauss's interpretation (1958), which makes central Machiavelli's claims to have discovered new intellectual territory and argues that Machiavelli turned away from both classical philosophy and Christian thought. See also H. Mansfield, *Machiavelli's Virtue* (Chicago: University of Chicago Press, 1996) and H. Mansfield, *Machiavelli's New Modes and Orders: A Study of the "Discourses on Livy"* (Chicago: University of Chicago Press, 2001) (originally published by Cornell University Press in 1979). By contrast, the contextual approach, broadly referred to as classical republicanism, highlights his participation in an ideology derived from classical antiquity that views republics as a unique locus for the development of the human personality and the flourishing of liberty. J. G. A. Pocock's influential interpretation (1975) identified Machiavelli as a type of republican thinker indebted to an Aristotelian emphasis on civic participation in forming the human personality. Q. Skinner's contextual approach focuses on a Roman intellectual tradition, derived primarily from Cicero and Sallust, which he terms "neo-Roman," that links negative liberty to virtuous civic participation. See Q. Skinner, "The Idea of Negative Liberty: Machiavellian and Modern Perspectives" and "Machiavelli on *Virtù* and the Maintenance of Liberty" both in Skinner (2002) *Renaissance Virtues*, vol. 2 of *Visions of Politics* (Cambridge: Cambridge University Press).

6
HOBBES

S.A. Lloyd

The seventeenth-century English philosopher Thomas Hobbes (1588–1679) famously boasted that the science of politics was no older than his own work *De Cive*. While many readers, from his time to ours, have objected to his absolutist conclusions (Lawson 1657: I.4–5; Locke 1690; Gauthier 1969: 169–70), his attempt to lay a sound scientific foundation for those conclusions is widely recognized. His appreciation for the method of Euclidean geometry and his admiration of his contemporary fellow natural philosopher Galileo speaks to Hobbes's commitment to a scientific understanding, not just of the non-human natural world, but of the human social world. Hobbes argued that persons characterized by a realistic human psychology in the social circumstances we typically face would find their rational interests best served by jointly submitting to the authority of an undivided, unlimited, absolute sovereign authority; further, such submission is dictated to them as a matter of moral and religious duty. Some interpreters have taken Hobbes to have offered a powerful early version of Social Contract theory, later developed by Locke, Rousseau, and Rawls. Others have viewed Hobbes as having modernized his inherited natural law tradition.

1. Hobbes's Political-Philosophical Project

Hobbes sought to discover rational principles for the construction of a civil polity that would not be subject to destruction from within. Having lived through the period of political disintegration culminating in the English Civil War, he came to the view that the burdens of government by an absolutist monarch are "scarce sensible, in respect of the miseries, and horrible calamities, that accompany a Civill Warre" (*Leviathan* XVIII.20). Because virtually any government would be better than a civil war, and, according to Hobbes's analysis, all but absolute governments are systematically vulnerable to dissolution into civil war, people ought to submit themselves to an absolute political authority. Continued stability will require that they also refrain from the sorts of actions that might undermine such a regime. For example, subjects should not dispute the sovereign power and they should not use their private interpretations of religious duty as a pretext for sedition. In general, Hobbes aimed to demonstrate the reciprocal relationship between political obedience and a peaceful, stable, enriching social environment. When people mutually covenant each to the others to obey a common authority, they have established what Hobbes calls "sovereignty by institution" (*Leviathan* XX.2). When, threatened by a conqueror, they covenant for protection by promising obedience, they have submitted to "sovereignty by acquisition" (*Leviathan* XX.1). These are equally legitimate ways of establishing sovereignty, according to Hobbes, but for most of us, who are born into

ongoing political societies, the more pressing question is not of how sovereignty might originally be established but, rather, what our attitude toward it must be if it is to remain stable and functional. We should regard ourselves *as if* we had owned and authorized all the actions of our existing sovereign authority. Political legitimacy depends not on how a government came to power, but only on whether it can effectively protect those living under it. Political obligation ends when protection ceases.

2. Unlimited and Undivided Government

Although Hobbes offered some practical grounds for preferring monarchy to other forms of government, his main concern was to argue that effective government—whatever its form—must have absolute authority. Its powers must be neither divided nor limited. The powers of legislation, adjudication, enforcement, taxation, war-making (and the less familiar right of control of normative doctrine) are connected in such a way that a loss of one may thwart effective exercise of the rest; for example, legislation without interpretation and enforcement will not serve to regulate conduct. Only a government that possesses all of what Hobbes terms the "essential rights of sovereignty" (*Leviathan* XVIII.16–17) can be reliably effective, since where partial sets of these rights are held by different bodies that disagree in their judgments as to what is to be done, paralysis of effective government, or degeneration into a civil war to settle their dispute, may occur.

Similarly, to impose limitations on the authority of the government is to invite irresoluble disputes over whether it has overstepped those limits. If each person is to decide for herself whether the government should be obeyed, factional disagreement—and war to settle the issue, or at least paralysis of effective government—are quite possible (*Leviathan* XVIII.4). To refer resolution of the question to some further authority, itself also limited and so open to challenge for overstepping its bounds, would be to initiate an infinite regress of non-authoritative "authorities" (*Leviathan* XXIX.9). To refer it to a further authority itself unlimited, would be just to relocate the seat of absolute sovereignty, a position entirely consistent with Hobbes's insistence on absolutism. To avoid the horrible prospect of governmental collapse or bloody civil war, people should treat their sovereign as having absolute authority to settle all disputes, including those over its exercise of sovereign powers.

3. Hobbes's Method

Hobbes judged that earlier writers on politics had failed to discover the necessary properties of stable states, the rights of sovereigns and the duties of subjects, because they lacked proper scientific method. Extrapolation from past experience might, if we have had a lot of experience and are very good at extrapolating from it, yield prudence; but this lacks the certainty of that sapience gained only by correct reasoning from proper principles. Experiences can be incomplete or wrongly interpreted; even sense perceptions can lead to false judgments, as when we perceive a straight stick as bent when partially submerged in clear water.

The proper method for discovering truth is *definitional*. We reason from premises that are analytically true, that is, true by virtue of the meanings of their component terms. The meanings of those terms must themselves enjoy a fixed, settled signification, and we are to proceed syllogistically using settled rules of inference. The conclusions we reach thus enjoy the status of reliable, non-contingent truths. Hobbes defines a commonwealth as

one "person" (this could be an "artificial person" such as an assembly), authorized by all members to use their goods and their abilities as *it judges necessary* for their peace and common defense (*Leviathan* XVII.13). Both the essential rights of sovereigns and the duties of subjects are to be unfolded from this concept of a commonwealth (Lloyd 1992: 58–68).

However, not just any old arbitrary definitions will track the reality of human concerns and experience, nor offer us a practically useful science of politics. Further, many of the premises that must go into any realistic science of politics will have to include claims about what interests humans have, how they are motivated, how they interact in groups, and so on, that will not be true by definition but, rather, warranted by experience. It looks as if despite his stated definitional method, Hobbes will be forced to depend on some empirical observations if his political science is to have any interest for, or claim on, humans as we are.

One can address this apparent problem by reflecting on Hobbes's *reason* for aspiring to employ a definitional method. He wants his readers to have confidence in his conclusions and the system constructed on the basis of those conclusions. A definitional method, of the sort he perceived in Euclid's proof of the Pythagorean theorem—which Hobbes declared a totally convincing proof of an astonishing conclusion (Aubrey 2000: 150)—was what he wanted for his own political argument (*Leviathan* IV.12–13). He wanted the component conclusions of his political philosophy to be not doubted, and not contestable, at least by anyone prepared to consider them carefully and in good faith. He wrote in the Dedicatory Epistle to his *Elements of Law* that

> [t]o reduce this doctrine to the rules and infallibility of reason, there is no way, but, first put such principles down for a foundation, as passion, not mistrusting, may not seek to displace; and afterwards to build thereon the truth of cases … till the whole have been inexpugnable.

Because Hobbes was engaged in the practical project of showing his countrymen that they have good and sufficient reasons to submit to the authority of their existing government, and treat it as if it were absolute, it makes sense to distinguish between those improperly synthetic premises that are harmful to his project, and those that are not.

Presumably with this in mind, Hobbes permits himself some empirical premises, but only those that would not be doubted by anyone. One may think of such premises as *indubitable introspectables*, premises that "passion not mistrusting, will not seek to displace."

Such premises are those that each person can confirm by introspection to be true in her own case, and so not doubt. They will not include universal generalizations about what everyone desires, or what everyone thinks, because these cannot be known by individual introspection. Hobbes insists, firstly, that our science is too primitive to give us knowledge of human nature:

> For it is supposed that in this natural kingdome of God, there is no other way to know anything, but by natural reason; that is, from the principles of natural science; which are so farre from teaching us anything of God's nature, *as they cannot teach us our own nature*, nor the nature of the smallest creature living.
>
> (*Leviathan* XXXI.33)

Hobbes's second reason for thinking that no proper proof of political principles will include empirical generalizations about what humans desire or think is principled.

Science might tell us something of the *mechanisms* by which we operate—explaining what it is to want something or to believe something—but it cannot tell us what everyone wants or believes because "for these the constitution individual, and particular education do so vary, and they are so easie to be kept from our knowledge [by lying] that the characters of mans heart … are legible only to him that searcheth hearts," that is, to God (*Leviathan*, Introduction).

When Hobbes first conceived his political project, he thought to proceed from an investigation of nature to man to political society. However, as political tensions rose in his own society, and addressing political questions became more urgent, he discovered that his political doctrine did not stand in need of any prior inquiry into larger nature or specifically human nature (*De Cive*, Preface). Hobbes may have been among the first, and arguably the most significant, political philosophers to assert the independence of political theory. Rather than treating it as an application of the ethics for natural man, Hobbes treated it as the science of an artificial entity, the commonwealth.

4. Hobbesian Psychology

Any adequate science of politics requires a veridical conception of human reasoning, desires, interests, and motivation. Traditionally, interpreters understood Hobbes to espouse a narrowly self-interested, or egoistic psychology concerned only with the agent's health, wealth, reputation, or personal profit. Recent scholarship has rejected the attribution to Hobbes of an egoistic psychology (Gert 2010) and has shown Hobbes to have designed a political theory to cope with people's *transcendent interests*, interests in pursuit of which they are willing to risk or even embrace the death of their present bodies (Lloyd 1992: ch. 8; 2009: 68–70). Hobbes also stressed the motivating force of pride and the sense of honor.

5. The State of Nature

To establish his conclusion that our interests are best served by submitting to an absolutist government, Hobbes invites us to consider what life would be like in a state of nature, that is, a condition without government, and where each decides for herself how to act, and is judge, jury and executioner in her own case whenever disputes arise. Hobbes terms this situation "the condition of mere nature," a state of perfectly private judgment, in which there is no agent with recognized authority to arbitrate disputes and effective power to enforce its decisions.

Hobbes's near intellectual descendant, John Locke, insisted that the state of nature was indeed to be preferred to subjection to the arbitrary power of an absolute sovereign. But Hobbes famously argued that such a "dissolute condition of masterlesse men, without subjection to Lawes, and a coercive Power to tye their hands from rapine, and revenge" would make impossible all of the basic security upon which comfortable, sociable, civilized life depends. If so, people have strong reasons to avoid a state of nature, which can be done only by submitting to some mutually recognized public authority, for "so long a man is in the condition of mere nature, (which is a condition of war,) as private appetite is the measure of good and evill" (*Leviathan* XV.40).

Although many readers have criticized Hobbes's state of nature as unduly pessimistic, he constructs it from a number of individually plausible empirical and normative assumptions. He assumes that people are sufficiently similar in their mental and physical

attributes that no one is invulnerable nor can expect to be able to dominate the others (*Leviathan* XIII.1). Hobbes assumes that people generally shun death, and that the desire to preserve their own lives is very strong in most people. While people have local affections, their benevolence is limited, and they have a tendency to partiality. Concerned that others should agree with their own high opinions of themselves, people are sensitive to slights (*Leviathan* XIII.5). They make evaluative judgments, but often use seemingly impersonal terms such as "good" and "bad" to stand for their own personal preferences (*De Cive* 14.17; *Leviathan* XLVI.11). They are curious about the causes of events, and anxious about their futures; according to Hobbes, these characteristics incline people to adopt religious beliefs, although the content of those beliefs will differ depending upon the sort of religious education one has happened to receive (*Leviathan* XII.2–6, 11, 23–4).

Hobbes ascribes to each person in the state of nature a liberty right which he terms "the right of nature" (*Leviathan* XIV.1). This is the right to do whatever one sincerely judges needful for one's preservation; yet because it is at least possible that virtually anything might be judged necessary for one's preservation, this theoretically limited right of nature becomes, in practice, an unlimited right to potentially anything, or, as Hobbes puts it, a right "to every thing" (*Leviathan* XIV.4). Hobbes further assumes as a principle of practical rationality, that people should adopt what they see to be the necessary means to their most important ends.

Taken together, these plausible descriptive and normative assumptions yield a state of nature potentially fraught with divisive struggle. The right of each to all things invites serious conflict, especially if there is competition for resources, as there will surely be over at least scarce goods such as the most desirable lands, spouses, etc. People will quite naturally fear that others might (citing the right of nature) invade them, and may rationally plan to strike first as an anticipatory defense. Moreover, that minority of prideful or "vain-glorious" persons who take pleasure in exercising power over others will naturally elicit preemptive defensive responses from others (*Leviathan* XIII.3–7). Conflict will be further fueled by disagreement in religious views (*Leviathan* XII.11–12, 23–4), in moral judgments, and over matters as mundane as what goods one actually needs, and what respect one properly merits. Hobbes imagines a state of nature in which each person is free to decide for herself what she needs, what she's owed, what's respectful, right, pious, prudent, and also free to decide all of these questions for the behavior of everyone else as well (Molesworth 1841: Vol. II, 9), and to act on her judgments as she thinks best, enforcing her views where she can. In this situation where there is no common authority to resolve these many and serious disputes, we can easily imagine with Hobbes that the state of nature would become a state of war, even worse, a war of "all against all." Hobbes defines a state of war as a tract of time in which there is a known disposition to violence and no guarantee against violence. (*Leviathan* XIII.8) But universal self-government by individual private judgment cannot guarantee against the violent resolution of disputes.

In response to the natural question whether humanity ever was generally in any such state of nature, Hobbes gives three examples of putative states of nature. First, he notes that all sovereigns are in this state with respect to one another. This claim has made Hobbes the representative example of a "realist" in international relations. Second, he opined that many now civilized peoples were formerly in that state, and some few peoples—"the savage people in many places of America" (*Leviathan* XIII.11), for instance—were still to his day in the state of nature. Third and most significantly, Hobbes asserts that the state of nature will be easily recognized by those whose formerly peaceful states have collapsed into civil war. While the state of nature's condition of

perfectly private judgment is an abstraction, something resembling it too closely for comfort remains a perpetually present possibility, to be feared, and avoided.

Some scholars stress the threat such a condition poses to individual self-preservation, the desire for which Hobbes is said to regard as necessary and overriding in any non-pathological human (Hampton 1986; Murphy 2000). Other scholars see the state of nature as problematic no matter what ends we attribute to humans, because their uncoordinated actions create sufficient interference that none can be sure of gaining and maintaining access to the material resources and unimpeded movement needed to pursue his ends (Lloyd 2009: 18–25). Humans are, by definition in Hobbes's view, rational agents, and *qua* rational and agents, they must desire to make their agency effective. This is a necessary desire—perhaps the only real candidate for a necessary desire, in contrast to desires for such things as wealth and temporal bodily self-preservation that might mean less to many people than the pursuit of honor, eternal salvation, or the welfare of their family or nation. Yet the universal right to all things undermines effective agency. It follows that we must abridge our right to all things.

However, the fact that it can be disastrous for everyone to exercise their right to all things by no means implies that everyone must give up that right. It is an open question who should abridge their right, and how much of it they should abridge. Hobbes's answer offers a simple reciprocity theorem: "Whatsoever you require that others should do to you, that do ye to them" (Leviathan XIV.5). Hobbes calls this theorem the core or sum of the laws of nature, and sometimes says simply that it *is* the law of nature.

Hobbes offers a number of differing formulations of this reciprocity requirement, as prohibiting doing what one would not have done to oneself; doing what one thinks unreasonable to be done by another to oneself; doing what one would not approve in another; reserving to oneself any right one is not content should be reserved to all the rest; and allowing to oneself that which one denies to another. Hobbes also offers positive formulations of the reciprocity theorem as commanding that you do to others whatever you require that they should do to you, and that we love others as ourselves. He writes, "*Thou shalt love thy neighbor as thyself* … is the natural law, having its beginning with rational nature itself" (*Philosophical Rudiments* XVII.8). Hobbes characterizes the required sort of love behaviorally rather than as an experience of affectionate feeling: "To love our neighbor as ourselves, is nothing else but to grant him all we desire to have granted to ourselves" (*Philosophical Rudiments* IV.12). It is "so to be understood as that a man … should esteem his neighbor worthy of all rights and privileges that himself enjoyeth … [He] should be humble, meek, and content with equality" (*Elements of Law* I.V.6).

The core of the laws of nature requiring reciprocity establishes who should lay down the right to all things, and which specific rights falling under it are to be laid down: "At the entrance into the conditions of peace, no man require to reserve to himself any right which he is not content should be reserved to all the rest," according to *Leviathan*'s tenth law of nature. A person reserves a right to himself when he insists that in some sphere of activity his private judgment should be determinative. If he would be unwilling to let others insist on the authority of their own private judgment in, say, property disputes with him, then the tenth law of nature requires that he agree with others jointly to defer to the judgment of a public arbitrator in such matters.

A sovereign authority then, can be understood as a public political judgment to which a citizen or subject is to defer his own private judgments in certain matters, namely those matters on which he would not be willing to permit everyone to act on their own private judgments. The laws of nature Hobbes specifically enumerates describe rules for

getting along with others. They concern only the behavior of "men in multitudes," that is, people living together in groups. There may be other rules of reason prohibiting certain self-regarding vices such as gluttony or drunkenness, but Hobbes does not include them in his discussions of the laws of nature. Those he does include facilitate peaceful, sociable, and comfortable living within communities.

The laws of nature articulate a set of natural duties rather than obligations, meaning that their claim on us does not depend on our having agreed or promised to follow them. One is to be faulted for ingratitude to good parents and other benefactors even though one has never promised not to be ungrateful. Those who have never agreed to treat others fairly are nevertheless blameworthy for behaving inequitably. Hobbes notes that even during war, gratuitous cruelty is forbidden by the law of nature; it would certainly be no moral defense to explain that one had never undertaken any obligation to refrain from such cruelty. The third law of nature, requiring the keeping of valid covenants, provides a natural duty to underwrite the obligations we undertake through our voluntary actions.

However, although the laws of nature articulate a set of "immutable and eternal" duties, Hobbes allows that when no one else is performing them, we are excused from unilateral performance (*Leviathan* XV.36). This has led some interpreters to argue that each of the laws has a two-clause structure (the latter clause suppressed) of "Do X, provided others are doing so as well" (Kavka 1986: 344). These normative rules enjoy a triple status of precepts of prudence, moral requirements, and those of God's laws discoverable without benefit of divine revelation. Hobbes holds that anyone who has attained the use of reason can be expected to acknowledge, and be held accountable to, the laws of nature. Only children and madmen are excused (*Leviathan* XXVII.23). Scholars disagree as to whether these laws are literally laws, understood as the commands of an authority, that is, of someone whom we are formerly obliged to obey, namely God (Martinich 2005: 120–1).

6. Deriving the Laws of Nature

How does the exercise of unaided natural reason result in everyone's arriving at the specific eternal and immutable rules Hobbes terms the laws of nature? *Desire based derivations* of Hobbes's laws of nature present those laws as necessary means for obtaining peace, which is in turn desired instrumentally as a necessary means to the individual's self-preservation (Curley 1994; Gauthier 1969; Hampton 1986). *A rule-egoist variant of desire based derivations* developed by Kavka (1986: 357–68) holds that Hobbes sought to justify the laws of nature by showing that the best *strategy* for securing self-preservation is always to conform to, or at least to accept and try to conform to, the laws of nature, even in those particular cases in which violating the laws can be seen to be in one's interest. *Duty based derivations* of Hobbes's laws of nature include *divine command interpretations* (Hood 1964; Martinich 1992; Taylor 1965; Warrender 1957) that understand the "oughtness" or normativity of Hobbes's laws of nature to be a function of their having been commanded for our observance by God. Some of these interpretations argued that Hobbes's moral theory is independent of his psychological theory and is a strict deontology. The *rationally required end interpretation* (Gert 2001) derives the laws of nature from a duty of self-preservation imposed on us by reason itself, which specifies required ends as well as finding means to our idiosyncratic ends. *Definitional derivations* of the laws of nature seek to use only definitions and premises analytic to the theory, in accordance with Hobbes's stated geometrical method. Some definitional derivations attempt to derive the laws of nature from Hobbes's definition of a law of nature (Deigh

1996; McNeilly 1968), while others offer derivations of Hobbes's specific laws of nature based on Hobbes's definition of man as a rational agent (Lloyd 2009).

7. Application of the Laws of Nature

Hobbes uses the laws of nature to settle a wide variety of particular moral judgments by applying reciprocity to factual situations, or as he calls them, cases in the law of nature. He condemns the efforts of missionaries to alter religion in another country on the ground that the missionary does that which he would not approve in another, namely, coming from abroad to try to convert those of one's own religion (*Leviathan* XXVII.4). Hobbes considers that valid covenants or contracts can make actions that were formerly morally indifferent now either morally required or impermissible; and because the third law of nature prohibits violating valid covenants, some people may be faulted for actions that others may perform blamelessly. For instance, the man who voluntarily enlists in the military undertakes an obligation not to plead fear as an excuse for running from battle (and is morally fault-worthy if he does so) (*Leviathan* XXI.16). The most important class of what Hobbes terms cases in the law of nature applies that law to the sovereign's behavior, yielding the norms of Hobbes's art of government.

Sovereigns are bound by *all* of the laws of nature because, these being divine, no person or commonwealth can alter or abrogate them. So, for instance, all are bound by equity "to which, as being a precept of the law of nature, a sovereign is as much subject, as any of the meanest of his people" (*Leviathan* XXX.15). It is true that the sovereign has final authority to interpret what equity requires of subjects and of himself; but that does not mean that equity just is whatever the sovereign thinks it is. Hobbes insists that "there is no judge subordinate, *nor sovereign*, but may err in a judgment of equity" (*Leviathan* XXVI.24). Sovereigns might be mistaken, but their interpretations and applications of the laws of nature are nonetheless to be treated as authoritative. Only God may hold them accountable for their errors.

Others of the sovereign's moral duties are office-specific. Hobbes writes:

> The office of the sovereign … consisteth in … the procuration of *the safety of the people*; to which he is obliged by the law of nature, and to render an account thereof to God, the author of that law, and to none but him. But by safety here is not meant a bare preservation, but also all other contentments of life, which every man by lawful industry, without danger, or hurt to the commonwealth, shall acquire to himself.
>
> (*Leviathan* XXX.1)

Hobbes insists that it is the sovereign's duty to leave people as much liberty as is safely possible, by which he means the maximum extent of liberty compatible with the maintenance of peace. All "harmless liberty" is to be allowed, and it is best that the sovereign's reasons for prohibiting subjects' activities be fairly intuitive so that people do not accidentally violate the laws through ignorance. Harmless liberties include such things as "the liberty to buy, and sell, and otherwise contract with one another; to choose their own abode, their own diet, their own trade of life, and institute their children as they themselves think fit; and the like" (*Leviathan* XXI.6).

What Hobbes terms the "true liberties of subjects" form a special case. As discussed in Chapter XXI of *Leviathan*, these are the rights which cannot be transferred by valid cov-

enant. Because they cannot be transferred, subjects are morally blameless when acting in defense of these rights; although sovereigns may legitimately command actions that threaten the interests these rights protect, and may punish disobedience, subjects may legitimately disobey those commands. Commands to kill, wound, or maim oneself, to abstain from things necessary to preserve one's life, to incriminate oneself or one's loved ones, or not to resist punishment may permissibly be disobeyed. It might seem that these sorts of liberties are far from harmless, but Hobbes suggests that acknowledging them really does not threaten public safety, for two reasons. First, it is not *necessary* to public safety that subjects obey commands to, say, execute themselves, because there are plenty of others who are bound to carry out the sovereign's execution orders. Second, the improbability that individuals will *reliably obey* such commands would render any presumed obligation on them to do so of no practical importance. Subjects need not, even if they could validly, divest themselves of the true liberties, because the sovereign's ability to secure the public good does not depend on that.

It is also a part of the sovereign's duty under the law of nature of procuring the common good that it regulate doctrine and religion so that subjects do not develop false beliefs that pit moral or religious beliefs, or the claims of private conscience against their duty of civil obedience. Taking educational measures against the "poison of seditious doctrines" (*Leviathan*) is especially important for, as Hobbes insightfully stressed, sovereign power ultimately depends on people's opinions of the sovereign's rights, especially on the opinions of those forces who are supposed to coerce compliance with the sovereign's commands. This is why Hobbes writes that

> the actions of men proceed from their opinions; and in the well-governing of opinions, consisteth the well-governing of men's actions, in order to their peace, and concord … It belongeth therefore to him that hath the sovereign power, to be judge, or constitute all judges of opinions and doctrines, as a thing necessary to peace, thereby to prevent discord and civil war.
>
> (*Leviathan* XVIII.9)

Hobbes took regulation of doctrine to require the education of ordinary people in their civic and moral duties by preachers trained in correct doctrine in the universities (*Leviathan* XXX.14), and suggested that his book *Leviathan* would make a suitable university text (*Review and Conclusion*).

Finally, the law of nature requires sovereigns to provide for the future good of their subjects both by establishing a mechanism for an organized succession, and, surprisingly, by doing whatever they sincerely think necessary for securing subjects' eternal good. Hobbes writes:

> Forasmuch as eternal is better than temporal good, it is evident, that they who are in sovereign authority, are by the law of nature obliged to further the establishing of all … they believe the true way thereunto. For unless they do so, it cannot be said truly, that they have done the uttermost of their endeavor [for the good of the people].
>
> (*Elements of Law* II.9.2)

It does not necessarily follow from this that sovereigns must dictate the profession and practice of religion to their subjects, for a conscientious belief in the value of free faith

would condemn such imposition. Still, sovereigns are required to take whatever measures they think needed to help their subjects toward a happy afterlife.

8. Good and Evil

Hobbes's laws of nature settle standards of right and wrong but, it might be wondered, how does Hobbes understand good and evil? He consistently points out that individuals *call* things good or evil according to whether they like or dislike them. We use these terms to express our attitudes, even though we take ourselves to be talking about independent properties of actions, objects and states of affairs. This has led many scholars to attribute moral subjectivism to Hobbes. Naturally, as our likes and dislikes differ from those of others, so do our judgments of good and evil. For this reason, Hobbes thinks that sound moral conclusions can never begin from claims about what is good, because the term "good" has no settled universal signification. Further, but against subjectivist interpretations, although it is a fact that people *use* the terms to track their likes and dislikes, such use is improper, on Hobbes's view. Hobbes condemns the schools of the Grecians as unprofitable, in part, for the reason that

> [t]heir moral philosophy is but a description of their own passions. For the rule of manners, without civil government, is the law of nature; and in it, the law civil; that determineth what is *honest*, and *dishonest*, what is *just*, and *unjust*; and generally what is *good* and *evil*. Whereas they make the rules of *good* and *bad* by their own *liking* and *disliking*: by which means, in so great diversity of taste, there is nothing generally agreed on; but everyone doth, as far as he dares, whatsoever seemeth good in his own eyes, to the subversion of commonwealth.
>
> (*Leviathan* XLVI.11)

To take one's own affections as the standard for good and evil is only human, but it is an error, and not just an error, but a social catastrophe. The dismal state of nature is precisely the condition in which private appetite is the measure of good and evil, and is dismal for that very reason (*Leviathan* XV.40).

Hobbes identifies the insistence on taking one's own likes and dislikes as the standard of good and evil as a form of *hubris*, or excessive pride. Hobbes names his treatise on how to create a stable state *Leviathan* after "the king over all the children of pride" discussed in the biblical *Book of Job*. The sorts of private judgments that disrupt social order are judgments about what God wants, what is just, righteous, sinful, etc.; and a person willing to take up arms to impose those judgments, in defiance of the civil sovereign, must feel very certain that he knows better than anyone whose judgments differ from his own. This prideful self-assurance in the superiority of one's own opinions is condemned by God himself not only in the *Book of Job*, but in the *Genesis* account of God's anger with Adam and Eve for taking on themselves to judge good and evil by eating from the tree of knowledge, and subsequently judging their nakedness, in which God saw fit to create them, as shameful.

9. The Relation of Natural Law to Civil Law and to Divine Positive Law

Hobbes describes the relationship between the moral law and positive civil law as one of mutual containment. What it means to say that these types of law are of "equal extent"

and contain each other is subject to interpretation. If the laws of nature command submission to all civil laws, and if every civil legal system must include the laws of nature as elements of itself, then we might see them as of equal extent and containing each other. If each sovereign has ultimate authority to interpret the laws of nature, and the laws of nature themselves require subjects to treat the sovereign's interpretation of their content as authoritative, Hobbes's system will be a "self-effacing" natural law theory (Lloyd 2009: 264–6) rather than a form of legal positivism.

Divine positive laws, such as are the Ten Commandments, also require interpretation. Hobbes argues that the Judeo-Christian Scripture authorized the civil sovereign to be the final earthly interpreter of these laws (*Leviathan* XXX.5). Although the sovereign might err in interpreting them, and will be held accountable by God for any such errors, subjects' foremost duty to God is to treat all their sovereign's judgments as authoritative, even those concerning the interpretation of natural law and of divine positive law (*Leviathan* XLII.106). In this way, Hobbes seeks to reconcile the requirements of morality, revealed religion, and civil law.

Hobbes progressively expanded his discussion of Christian religion in each version of his political philosophy, until it comes in *Leviathan* to comprise roughly half the book. Interpreters disagree about how Hobbes understands the significance of religion within his political theory. Some have doubted the sincerity of his professed Christianity, arguing that by the use of irony or other rhetorical devices, Hobbes sought to undermine his readers' religious beliefs (Curley 1994; Johnston 1986).

Others argue that Hobbes is trying to demonstrate the compatibility of his political theory with core Christian commitments, since it might seem that Christians' religious duties forbid their affording the absolute obedience to their governors that Hobbes's theory requires of them. This can be a potent source of disorder, especially when their religious interests are transcendent.

Related Topics

Anarchism, Locke, Contractarianism, Natural Law and Rights Theory, Authority and Legitimacy, Rights

References

Hobbes wrote several versions of his political philosophy, including *The Elements of Law, Natural and Politic* (also under the titles *Human Nature* and *De Corpore Politico*) published in 1650, *De Cive* (1642) published in English as *Philosophical Rudiments Concerning Government and Society* (herein cited by chapter and paragraph) in 1651, the English *Leviathan* (herein cited by chapter and paragraph) published in 1651, and its Latin revision in 1668.

Aubrey, John (2000) *Brief Lives*, ed. John Buchanan-Brown, Harmondsworth: Penguin Books.

Curley, E. (1994) "Introduction to Hobbes's *Leviathan*," in E. Curly (ed.), *Leviathan*, Indianapolis: Hackett Publishing Company, Inc.

Deigh, John (1996) "Reason and Ethics in Hobbes's *Leviathan*," *Journal of the History of Philosophy* 34: 33–60.

Gauthier, D. (1969) *The Logic of "Leviathan": The Moral and Political Theory of Thomas Hobbes*, Oxford: Clarendon Press.

Gert, Bernard (2001) "Hobbes on Reason," *Pacific Philosophical Quarterly* 82: 243–57.

—— (2010) *Hobbes*, Cambridge: Polity Press.

Hampton, Jean (1986) *Hobbes and the Social Contract Tradition*, Cambridge: Cambridge University Press.

Hood, F.C. (1964) *The Divine Politics of Thomas Hobbes*, Oxford: Clarendon Press.

Johnston, David (1986) *The Rhetoric of Leviathan*, Princeton, NJ: Princeton University Press.
Kavka, Gregory S. (1986) *Hobbesian Moral and Political Theory*, Princeton, NJ: Princeton University Press.
Lawson, George (1657) *An Examination of the Political Part of Mr. Hobbs his Leviathan* (printed by R. White for Francis Tyton).
Lloyd, S.A. (1992) *Ideals as Interests in Hobbes's "Leviathan": The Power of Mind over Matter*, Cambridge: Cambridge University Press.
—— (2001) "Coercion, Ideology, and Education in Hobbes's *Leviathan*," in Andrew Reath, Barbara Herman, and Christine Korsgaard (eds), *Reclaiming the History of Ethics: Essays for John Rawls*, Cambridge: Cambridge University Press, pp. 36–65.
—— (2009) *Morality in the Philosophy of Thomas Hobbes: Cases in the Law of Nature*, Cambridge: Cambridge University Press.
Locke, John (1690) *Second Treatise of Government*.
Martinich, A.P. (1992) *The Two Gods of Leviathan: Thomas Hobbes on Religion and Politics*, Cambridge: Cambridge University Press.
—— (2005) *Hobbes*, New York: Routledge.
McNeilly, F.S. (1968) *The Anatomy of Leviathan*, London: Macmillan.
Molesworth, Sir William, ed. (1841) *The Collected English Works of Thomas Hobbes*, (in 11 volumes), London: John Bohn.
Murphy, M. (2000) "Hobbes on the Evil of Death," *Archiv für Geschichte der Philosophie* 82: 36–61.
Taylor, A.E. (1965) "The Ethical Doctrine of Hobbes," in K.C. Brown (ed.), *Hobbes Studies*, Oxford: Blackwell.
Warrender, H. (1957) *The Political Philosophy of Hobbes: His Theory of Obligation*, Oxford: Clarendon Press.

7
LOCKE

Eric Mack

No explication of Locke's political philosophy will accommodate everything Locke said in the course of his forty years of moral and political theorizing. Nevertheless, one account does capture much more of what is philosophically central, distinctive, and interesting in Locke's political thought than alternative accounts. On this account, Locke defends an essentially secular, rights-based, liberal individualism. In this essay I present a streamlined and philosophically sympathetic version of this Locke; I do not attempt to explicate or rebut alternative understandings of Locke. I draw primarily from Locke's two most important and best-known works in political philosophy—the *Second Treatise* (*ST*) from his *Two Treatises of Government* and *A Letter Concerning Toleration* (*LCT*). I also draw upon Locke's early lectures, *Essays on the Law of Nature* (*ELN*), and some passages from Locke's *First Treatise* (*FT*) and his *Essay Concerning Human Understanding* (*ECHU*). My presentation starts with the Lockean state of nature and persons' natural and property rights and moves on to the inconveniences of the state of nature, the contractual formation of government, and the conditions under which political power may justly be resisted. I then turn to Locke's basic case for religious toleration and the place of religious toleration within Locke's overall political doctrine.

1. Freedom

At the core of Locke's liberalism is the endorsement of *freedom* and *law* and the condemnation of *slavery* and *tyranny*. An individual is free to the extent that he is not precluded from disposing as he sees fit of his own person and (legitimately acquired) possessions. A regime is lawful to the extent to which each individual's right to freedom is respected. The pre-political state of nature is lawful insofar as within it individuals respect the moral rights that individuals possess within it (and abide by other strictures of the law of nature). The political state is lawful insofar as the enactments and practices of those with political power respect the finer delineations of individual moral rights which it is the business of political authority to draw and enforce. In this way, freedom and law are complementary; an individual enjoys freedom to the extent that *others* in their transactions with him abide by the law of nature that governs the state of nature or abide by the positive law that provides a more determinate specification of the law of nature in the political state (*ST*: §135).

This view of the connection of freedom and law must not be confused with the Rousseauist-Kantian view that an individual's freedom consists in *his* compliance with the law. Furthermore, it differs from the Hobbesian-Benthamite view that freedom is a matter of acting as one desires. Locke rejects this Hobbesian-Benthamite view because on

it each person's freedom requires infringements on other persons' freedom and each person's exercise of his right to freedom requires infringements upon others' rights to freedom. "Freedom is not," as we are told, "*A Liberty for every Man to do what he lists*: For who could be free, when every other Man's Humour might domineer him?" (*ST*: §57). "To be subject to law that requires one to allow others to dispose of their own persons and possessions as they see fit is not an infringement upon one's liberty" (*ST*: §57).

To be denied freedom is to be treated as a slave. To be subjected to unlawful political power is to be subjected to tyranny. Slavery and tyranny are, therefore, the antipodes of freedom and law. Furthermore, Locke links the law of nature and the finer delineations of that law in the political state to *reason*. For the law of nature is grounded in our existence as rational beings and is identified by reason. Compliance with the rights and obligations that (in part) constitute the law of nature or with that law's finer positive articulations is at least a necessary condition of practical rationality. Thus, unlawful action in violation of those rights and obligations is also contrary to reason. Such action reflects arbitrary, i.e., unlawful and unreasonable, will. Freedom, then, is also a matter of not being subject to the arbitrary will of others; and the fuller statement of the basic background contrast that frames Locke's political thought is: freedom, law, and reason versus slavery, tyranny, and arbitrary will.

2. Property

The core claim that freedom is non-interference with one's disposal of one's "whole property" *seems* to imply that persons have rights to freedom in virtue of having rights over particular objects—most saliently, rights over themselves. Putting matters this way highlights the crucial role of *property* in Locke's explication of freedom and of the right to freedom. Conceptually, the realization of freedom and lawfulness is linked to the recognition and respect for property. In practice, recognition and respect for mine and thine is the key to social order, peace, and prosperity. Nevertheless, saying that persons have rights to freedom *in virtue of* having proprietorship over particular objects might make persons' moral claims to freedom too conceptually dependent upon an *independent* identification of their rights. Such an identification of those rights could not appeal to—since it would have to be the ground of—persons' moral claims to freedom. A more promising understanding of Locke's position is that each person's basic moral claim to freedom is to some considerable extent *best codified* in terms of ownership claims. Ownership is the characteristic mode through which the basic right to freedom is expressed.

It is often held that for Locke the right of self-ownership is the sole original and non-acquired right. All other non-acquired rights are aspects of this one sturdy foundational right and all acquired rights arise through the exercise of this self-ownership right. This is mistaken. For, in his key arguments for natural rights in chapter II of the *Second Treatise*, Locke argues for a wider moral right to freedom, a right against being subordinated to the will of others. He casts the conclusion of his arguments broadly and disjunctively. "[N]o one ought to harm another in his Life, Health, Liberty, or Possessions." Except to do justice on an offender, no one may "take away, or impair the life, or what tends to the Preservation of the Life, the Liberty, Health, Limb, or Goods of another" (*ST*: §6). A moral right over one's own person may be part of the codification of this broad right against being subordinated to others, but it does not seem to capture all of that right. For instance, self-ownership does not encompass each person's state of nature right that others abide by their agreements with him (*ST*: §14). Indeed, self-ownership as such is

not mentioned until Locke offers his explanation of acquired rights to particular possessions in the *Second Treatise*'s chapter "Of Property."

3. The Law of Nature: Two Programs

On what basis does Locke assert each individual's basic original claim to freedom? Answering this question is complicated by the presence in Locke of two distinct, but not clearly distinguished, programs for determining the content of the law of nature. We can label one "the divine command program" and the other "the inborn constitution program." Throughout his life, Locke subscribed to one thesis that is characteristic of authoritarian political doctrines, namely, law must be the product of the will of the law-giver. A norm is a matter of obligatory law if and only if it is commanded by a relevant legislator. "[I]n order that anyone may understand that he is bound by a law, he must know beforehand that there is a lawmaker, i.e., some superior power to which he is rightly subject" (*ELN*: 102). Hence, the most basic moral law, namely, the law of nature, consists of the norms that God commands us to follow. Locke always attempts to steer clear of crudely voluntarist versions of this divine command doctrine. It is not God's mere will nor God's propensity to punish non-compliance with his commands that makes his commands obligatory law. God's will binds us because of the *authority* God possesses in virtue of his wisdom and goodness or in virtue of his being our creator (*ELN*: 116–17). Still, even moderate voluntarism seems to imply that we are to determine the content of the law of nature by detecting what God wills us to do (*ELN*: 102).

The problem for this program is that we have very little access to God's intentions and will. We know that God wills us to do "something" (*ELN*: 105); but we do not have any direct knowledge about what he wills. Indeed, over his lifetime Locke increasingly narrows the range of knowledge we can have about God and his intentions. See especially Locke's second and third letters on toleration (Locke 1823). We do, however, have access to certain fundamental and morally fertile facts about human nature; we have knowledge about our "inborn constitution" (*ELN*: 125). Locke *might* have gone on to say that our knowledge of our inborn constitution is the basis for certain inferences about what God commands us to do or to leave off; and these inferences from human nature enable us to carry out the divine command program.

Yet, when Locke shifts focus to our inborn constitution, his explicit contention is that persons' inborn constitution determines how they should conduct themselves. God has willed that beings with our natures exist. But then it is our inborn constitution—and not God's will—that determines how beings with such a nature should act. The law of nature

> is a fixed and permanent rule of morals, which reason itself pronounces, and which persists, being a fact so firmly rooted in the soil of human nature. Hence human nature must needs be changed before this law can be either altered or annulled. … [S]ince man has been made such as he is, equipped with reason and his other faculties and destined for this mode of life, there necessarily result from his inborn constitution some definite duties for him, which cannot be other than they are.
>
> (*ELN*: 125)

Locke adds that "these duties of his necessarily follow from his very nature" and that, therefore, "natural law stands and falls together with the nature of man as it is at present" (*ELN*: 126). Certain features of our human nature ground the norms that constitute the law of nature and that is why, as the inborn constitution program asserts, we are to identify the content of the law of nature through a study of our inborn constitution. Moreover, when Locke actually sets out to ground natural rights in the early chapters of the *Second Treatise*, he proceeds by way of this inborn constitution program. It is either our nature or our reason as it inspects our nature and identifies its normative implications that provides the content of the law of nature.

In the pursuit of the inborn constitution program Locke focuses on two morally fertile features of our human constitution. First, each person seeks his own happiness and is rational to do so. Although all good, i.e., all happiness, is "the proper object of desire in general," not every instance of happiness moves each particular man. Rather, each man is moved only by those instances of happiness "which make a necessary part of his happiness."

> All other good, however great in reality or appearance, excites not a man's desires who look not on it to make a part of that happiness wherewith he, in his present thoughts, can satisfy himself. Happiness, under this view, everyone constantly pursues, and desires what make any part of it: other things, acknowledged to be good, he can look upon without desire, pass by, and be content without.
>
> (*ECHU*: II, 341)

In a fragment ("Thus, I Think") composed shortly before the publication of the *Two Treatises*, Locke declares that

> Tis a man's proper business to seek happiness and avoid misery. … I will therefore make it my business to seek satisfaction and delight and avoid uneasiness and disquiet and to have as much of the one and as little of the other as may be.
>
> (Goldie 1997: 296)

Within Locke's political writings, the stand-in for the desire for happiness and the rationality of its pursuit is the desire for (comfortable) self-preservation (*FT*: §88) and the rationality of its pursuit. And the crucial necessary *interpersonal* condition for the attainment of self-preservation is freedom, i.e., the protected liberty to dispose of one's life, limb, and possessions as one sees fit (*ST*: §17). Thus, the crucial implication for political theory of the fact that each rationally pursues his own happiness is that each rationally pursues his own freedom. The second normatively portentous feature of our constitution is that persons, sharing a common nature, are morally equal beings; whatever original moral claims are ascribed to one person must also be ascribed to all other persons.

Intuitively, the first feature underwrites the idea that persons each have ends of their own to which they are rationally oriented; no person is made for another's purposes. The subordination of another person to one's own purposes falsely presumes that the other is made for one's purposes and, thus, such subordination is unjustified (*ST*: §6). The first feature also underwrites the rationality of each person's laying claim to freedom from the interference of others and, thus, by way of the second feature, each is ration-

ally required to acknowledge a like claim for others (*ST*: §§4–7). Other persons exist as equal and independent moral beings (*ST*: §6). This fact about other persons has some sort of directive import for one's conduct toward them. Locke often expresses this directive import by saying that, if others are not offenders, one ought to preserve them as long as that preservation does not come into competition with one's own. However, Locke immediately parses this call for one's preservation of mankind as the requirement not to "take away, or impair the life, or what tends to the Preservation of the Life, the Liberty, Health, Limb, or Goods of another" (*ST*: §6).

The normative significance of the existence of other persons who, like oneself, rationally seek their own self-preservation and freedom is not that one must also promote their preservation or freedom. Rather the import is that one must refrain from depriving others of the freedom that is the crucial interpersonal condition for their respective attainment of self-preservation and happiness. To demand more of others is to demand their subordination to one's own purposes and will. Locke does assert that each is required not to subordinate others to his will because we are each God's property. But this argument—based as it is on *God's* alleged ownership of us—is not an argument for *our* having rights against being subordinated (*ST*: §6).

4. Rightful Holdings

Locke also holds that private property rights in particular extra-personal objects—in acorns, bits of turf, or fields—can arise in the state of nature. Since, all these rights must be acquired, Locke needs to provide an account of how such rights can and do arise. Moreover, since Locke was convinced by Robert Filmer's arguments (Sommerville 1991: 19–21) that private property rights could not arise through any general social contract, he needs to offer an explanation of such rights that does not appeal to any such compact (*ST*: §25).

Locke's explanation begins in the *First Treatise* (*FT*: §§86–8). Persons' natural rights to engage in self-preserving action (which is close to isomorphic with the natural right to freedom) must include a right to use and even acquire discretionary control over parts of the extra-personal world. For such use and (at least in some cases) such acquisition is necessary for the attainment of comfortable preservation. So persons have rights not to be precluded from using and even exercising ongoing discretionary control over portions of the extra-personal world; hence, the earth cannot originally be the joint property of mankind. For were it the joint property of mankind, each individual would need the consent of all others before making use of or appropriating anything; and, hence, no individual would have an original right to bring his powers to bear upon the world. Indeed, each individual would starve while waiting for everyone else to consent to his self-preserving actions (*ST*: §28). So there has to be *some* process by which individuals can generate private property rights without obtaining the consent of mankind.

In the *Second Treatise* Locke specifies the process that generates initial property rights. That specification builds on the original, non-acquired, right of self-proprietorship. Each self-owner has an original right over his own labor. When an agent brings his skills, insights, or energy to bear on some morally available material in the service of some intended transformation of that material, that agent invests his labor in the resulting object. Since the resulting object cannot then be taken from the agent without taking the labor that has been "mixed" with the object, taking that object without the agent's permission violates the agent's retained right to his labor. The agent has a

property right to the transformed object because taking it from him without his consent violates this retained right to his labor. Moreover, all raw material is morally available for acquisition through labor mixing precisely because the earth is originally unowned; for that is (almost) all that can be meant by saying that God has given the earth to all mankind in common. Locke also clearly believes that property rights established through the investment of one's labor can be transferred both through simple barter and complex commercial interactions. Yet he offers no explanation of how the recipient of a freely transferred holding acquires a right to that holding not merely against the person who transferred it to him, but *against the world at large*. For Locke, as for other seventeenth-century theorists, the generation of *initial* just holdings was the only pressing question.

Locke maintains, however, that there are limits on how much one may "*ingross*" (*ST*: §31) through labor mixing and voluntary transfer. One cannot establish rights over what will spoil in one's possession and, more significantly, private acquisition must leave "enough, and as good" for others. Nevertheless, the general tenor of Locke's "Of Property" is that the progressive development of private property does *not* transgress either of these limitations. The case for the non-transgression of the spoilage proviso is simple. Labor mixing involves taking pains. Prior to the existence of money, no sensible person will take pains to acquire more goods that are susceptible to spoilage than he (or his household) can consume or can barter for consumables. After the appearance of hard money, any sensible person will exchange goods that will otherwise spoil for gold or silver coins (*ST*: §46).

The case for the non-transgression of the enough and as good proviso is more complex and more interesting. Locke strongly associates the initial acquisition of portions of the earth with the shift from a hunter-gatherer to an agricultural mode of self-preservation. An individual makes this shift by settling down and making a bit of the earth his own. In turning to agriculture, that individual will need and will take possession of much less land than he previously used in hunter-gatherer fashion. In settling down, that individual in effect donates to others the difference between his past pro rata share of the land that they all hunted and gathered in and the amount of land he makes his own. Thus, at this crucial early phase in the rise of private property, the enough and as good proviso is readily satisfied.

Things get more complicated after the appearance of money. For money vastly increases the prospects of economic advancement by facilitating and drawing people into trade. Money radically intensifies and extends the market. It radically increases the number and range of desirable goods that one may acquire through trade and the incentives that one has to develop and exercise one's own industry so as to be in position to enter into gainful trades. Money also, of course, provides a store for the increased wealth that becomes available to one from industry and trade. The effect of money is, therefore, to increase enormously the development and deployment of human capital. This explains, according to Locke, the vast difference in material well-being between societies that lack private property, money, extensive trade, and commercial development and societies that are blessed with these institutions. Indeed, the "godlike" prince is the ruler who establishes "laws of liberty" that "secure protection and incouragement to the honest industry of Mankind" (*ST*: §42). Differences of "degrees of industry" among men naturally give rise to differences in their wealth and these differences are enlarged with the magnification of industry and trade in commercial society. Yet everyone gains absolutely in this transition. The economic game is a positive sum game precisely because

wealth derives from the development and exercise of human capital. And, if anything, one person's increased industriousness provides inducements and opportunities for others' enhanced industriousness. In fact, Locke's very strong contention is that the worst off day laborer in commercial England is materially better off than the king of a large, fertile, but hunter-gatherer domain in America (*ST*: §41).

Nevertheless, the incentives engendered by money and the world of commerce that money calls forth lead individuals so extensively to appropriate portions of the earth that enough and as good of the earth is not left for others to appropriate or even to use as would have been available had the earth remained an open commons. So it appears that in the second, money-engendered phase of a private property regime the enough and as good proviso is violated. Yet Locke denies this. Locke's explicit argument here falls back on his belief that in some relevant sense money arises through general agreement (*ST*: §36). Locke adds two further premises. First, if one agrees to a certain institution, one agrees to whatever are the obvious consequences of that institution and, second, it is an obvious consequence of the institution of money that enough and as good of the earth will not be left for some individuals. It follows that, wherever there is money, there has been universal agreement to enough and as good not being left for some individuals. Through the agreement that institutes money, this proviso has been set aside.

This is a bad argument for Locke. For money does not arise from general agreement. Moreover, Locke himself has vowed not to appeal to general agreement within his account of property rights. In addition, what one really wants to know is *why*, given its obvious consequences, everyone would agree to the institution of money. Locke's answer is that each would agree because each correctly anticipates being better off due to the introduction of money. Each correctly anticipates being better off with respect to the concern that motivates the enough and as good proviso, namely, how extensively one is "straitened" (*ST*: §36) by others in one's bringing one's powers to bear on the external world in pursuit of one's ends.

Initially, when one focuses on the use or appropriation of *raw material*, the concern about being straitened takes the form of concern about others not leaving enough and as good *raw material*. Yet Locke consistently emphasizes the relative unimportance of raw material in economic development. As one's environment develops economically, concern about straitening must increasingly take the more general form of concern about whether the overall receptivity of that environment—which now is largely constituted by the products of people's industry and their decisions about the disposition of those products—to one's exercise of one's industrious powers is at least as great as it would have been had private property and money not come on the scene. Since, on this more general measure of being straitened, an individual is not straitened by the development of property and commerce, the underlying motivation for the enough and as good proviso is satisfied and, hence, the proviso more broadly construed is satisfied—even if enough and as good *raw material* is not left for that individual to appropriate or use. (This sympathetic recasting of Locke's enough and as good proviso follows the general lines laid down in Robert Nozick's *Anarchy, State and Utopia* (1974: 178–82).)

Does this merely show that Locke's enough and as good proviso sets too low a baseline? Should one not be able to demand more of economic development than that it not diminish the receptivity of one's environment to one's laboring in the service of one's ends? Locke's answer is that to demand more than this simply on the basis of being in the presence of development engendered by other people's industry is to demand *unjustly* "the benefit of another's Pains" (*ST*: §34).

For Locke, first-order rights to dispose as one sees fit of one's person and legitimately acquired possessions imply second-order rights to enforce these rights, namely, to defend one's rights, to extract compensation for the violation of one's rights, and to punish violators of rights. In pre-monetary property regimes, the private exercise of these rights of enforcement causes little trouble because the boundaries that define people's rights are simple and readily perceived and one's neighbor really does not have that much to plunder or that much for one to seek through exaggerated demands for compensation (*ST*: §§51, 107). It is when property and contract become more varied and complex and there is more of value to contend over that frequent and serious inconveniences arise from individuals seeking to enforce the law of nature on the basis of their private judgments. As is well known, the Lockean remedy for these inconveniences is the establishment of government that provides known (positive) law, impartial judges, and reliable power to enforce the law and judicial decisions.

5. Government

This establishment involves two stages. The first is the creation of "political society" to which contracting individuals transfer their second-order rights to act as executors of the law of nature. The second is political society's establishment of a governmental structure and selection of key governmental officers who are contractually bound to political society to preserve that structure and to serve the purposes for which political society is created. It would be unrewarding to trace the details of Locke's tangled discussion of who consents at what point, within which stage, and with what resulting sort of obligation. But certain general features of Locke's position are worth emphasizing. With one important qualification, the only state of nature rights that are surrendered are the rights of defense, restitution, and punishment. First-order rights to life, liberty, and estate are retained by all—except that Locke imagines that each contracting individual agrees to contribute to the public enforcement of rights (*ST*: §§88, 140) and this presumably includes being subject to taxation to cover the costs of that enforcement. Moreover, the purpose of the establishment of known (positive) law, impartial judges, and reliable force is the protection of property in its broad sense of life, liberty, and estate (*ST*: §§88, 123, 131, 138). "The great end of men's entering into society [is] the enjoyment of their properties in peace and safety" (*ST*: §134). Since these rights do not owe their existence to government, their retention constrains all governmental activity; the function of positive law is merely to more finely articulate and enforce the law of nature.

> The Obligations of the Law of Nature, cease not in Society, but only in many Cases are drawn closer, and have by Humane Laws known Penalties annexed to them, to inforce their observation. Thus the Law of Nature stands as an Eternal Rule to all Men, *Legislators* as well as others.
>
> (*ST*: §135)

To think that men would surrender more of their rights to establish more extensive political authority is "to think that Men are so foolish, that they take care to avoid what Mischiefs may be done them by *Pole-Cats* or *Foxes*; but are content, nay, think it Safety, to be devoured by *Lions*" (*ST*: §93).

Numerous passages in the *Second Treatise* emphasize the role of majority rule within political society and within any legislature created by political society (*ST*: §§96, 97).

However, those majorities only have the authority to select among competing measures aimed at the protection of individuals' moral rights. For the community and any majority within it only acquire from consenting individuals the right to advance "the ends for which they unite into Society" (*ST*: §99, §222). Since "no Body has an absolute Arbitrary Power over himself, or over any other, to destroy his own Life, or take away the Life or Property of another" (*ST*: §135), no valid transfer of rights can result in the community or any majority within it having such an authority.

The establishment of governmental authority does not fundamentally alter persons' original equal moral standing. Acts of taxation by the government are morally permissible only insofar as they are akin to the enforcement of contractual commitments to pay for services rendered. Government officials who engage in actions that would be criminal if performed by a private citizen are themselves criminals.

> The Injury and the Crime is equal, whether committed by the wearer of a Crown, or some petty Villain. The Title of the Offender, and the Number of his Followers, make no difference in the Offence unless it be to aggravate it. The only difference is, Great Robbers punish little ones, to keep them in their Obedience; but the great ones are rewarded with Laurels and Triumphs …
>
> (*ST*: §176)

Due to the power at his command, the chief executive officer—the monarch—poses the greatest danger of this criminality (*ST*: §218). Such criminality vindicates resistance. In some cases, the chief officer—usually through his minions—engages in isolated violations of the rights of some subject. In these cases, if no legal recourse is available, resistance by the individual is permissible even though it is likely to be imprudent (*ST*: §208). In other cases, the chief officer either subverts the constitutional order established by political society or instigates widespread attacks on people's property. In these cases, the aggrieved party is that mysterious thing called "political society." Individuals as members of political society may both resist and replace the ruler. Indeed, actions by the ruler that justify resistance by political society also "dethrone" the ruler (*ST*: §239), so the people's resistence is not rebellion against legitimate authority.

But, who shall judge whether rights have (sufficiently) been violated? Locke insists that ultimately "*every Man is Judge* for himself, as in all other Cases, so in this" (*ST*: §241). Yet Locke also holds that the transfer of rights that creates political society involves the *surrender* of the right of private judgment (*ST*: §87), and it does seem that individuals can only exit the state of nature through some sort of mutual disavowal of private judgment. Perhaps we can reconcile these contentions by considering precisely why people agree to accept public articulations of rights and public decisions about common measures for the protection of those rights.

People agree to forego insistence upon their divergent judgments about which specifications and measures *best* protect property and, hence, they agree to abide by positive "Humane laws" (*ST*: §135) that draw closer the Law of Nature. For it is insistence on one's own judgment about precisely how the Law of Nature should be construed that generates the state of nature inconveniences. Lockean contractors do not, however, agree to be bound by *whatever* public articulations or common measures are issued or enforced. They only agree to be bound by public delineations and decisions about common measures that *can reasonably be construed* as best securing persons' rights. Private judgment on behalf of one's own favorite among the reasonable articulations or measures

is excluded. Nevertheless, each person retains the right to judge whether the conduct of political authorities can reasonably be construed as securing rights. If that conduct is privately judged to be within the range of the reasonable, the individual must take himself to be bound to comply. However, if governmental conduct is privately judged to be outside this range, the individual need not take himself to be bound to comply.

This position is suggested by Locke in *A Letter Concerning Toleration*.

> [T]he private judgment of any Person concerning a Law enacted in Political Matters, for the publick Good, does not take away the Obligation of that Law, nor deserve a Dispensation. But if the Law indeed be concerning things that lie not within the Verge of the Magistrate's Authority … men are not in these cases obligated by that Law, against their Consciences.
>
> (*LCT*: 48)

Locke's stance in the *Letter* is that any law restricting peaceful religious activity is beyond "the Verge of the Magistrate's Authority." No such law can reasonably be construed as protective of rights because no peaceful religious activity infringes upon anyone else's rights of life, liberty, or estate. "If any man err from the right way, it is his own misfortune, no injury to thee" (*LCT*: 31).

6. Religious Toleration

No agent's mismanagement of his own worldly affairs opens him up to legitimate coercive interferences; and the same anti-paternalist point applies to his mismanagement of extra-worldly affairs (*LCT*: 34). Nor does the judgment that another's religious practices are sinful ground legitimate coercive interference. For "it does not follow, that because it is a sin it ought therefore to be punished by the Magistrate" (*LCT*: 43). Only those acts may be punished that are "prejudicial to mens Rights [or] break the publick Peace of Societies" (*LCT*: 44). Moreover, legal restrictions on peaceful religious activities—backed as they are with punishments—reveal neither a concern for others' rights nor for the well-being of those subject to those punishments. Instead, such measures reveal "intemperate Zeal" and the desire for "Temporal Domination." It is not credible that someone who would deliver "his Brother … to be burnt alive, does sincerely and heartily concern himself to save that Brother from the Flames of Hell in the World to come" (*LCT*: 35).

No man (or Church) can transcend his (or its) own private judgment about what one must do to attain salvation: "every Church [and person] is orthodox to it [and him] self." Hence, the proposal that only those magistrates who adhere to the *correct* path for salvation have the authority to impose that orthodox path is merely "great and specious Words" (*LCT*: 32). For every magistrate will take this proposal as a vindication of his imposition of his own favored, hence, orthodox path. The solution lies not with any surrender of private judgment per se but, rather, with a recognition of the "Verge" of each individual's private judgment. In pursuit of salvation, each individual may dispose *of himself and his possessions* in accordance with his private judgment. Religious liberty, as with liberty in general, is a matter of respect for mine and thine. Indeed, one's protected liberty to engage in some religiously motivated activity has nothing to do with its religious motivation and everything to do with one's right to dispose as one sees fit of the resources employed in that activity. I may sacrifice a calf in my religious ceremonies

if, but only if, it is *my* calf (*LCT*: 42). Crucially, others' permitting me to engage in this ceremony does not involve their endorsement of it but, rather and merely, their recognition of my ownership of the calf.

The solution to religious conflict is the de-politicization of religion. Such de-politicization engenders peaceful, albeit competitive, coexistence among distinct churches. The search for religious truth is best advanced by such a marketplace in religious ideas (*LCT*: 46) for very much the same reasons as the search for material well-being is advanced by the economic market. A framework of positive law that clearly delineates people's individual rights and protects those rights against violation is called for by our existence as morally equal and independent beings. At the same time, the establishment of this framework—the institution of "laws of liberty"—enables individuals to create the complex of peaceful and mutually beneficial interactions that constitute a flourishing civil society.

Related Topics

Hobbes, Hume and Smith on Justice, Kant, Anarchism, Liberalism, Natural Law and Rights Theory, Libertarianism, Freedom, Rights, Toleration

References

Goldie, M., ed. (1997) *Locke: Political Essays*, Cambridge: Cambridge University Press.

Locke, J. (1689/1967) *Two Treatises of Government*, ed. Peter Laslett, Cambridge: Cambridge University Press.

—— (1689/1983) *A Letter Concerning Toleration*, ed. James Tully, Indianapolis: Hackett Publishing.

—— (1689/1959) *Essay Concerning Human Understanding*, ed. A. C. Fraser, New York: Dover Publications.

—— (1823) *The Collected Works of John Locke*, vol. 6 (containing the four letters concerning toleration), London, printed for Thomas Tegg *et al.*

Nozick, R. (1974) *Anarchy, State and Utopia*, New York: Basic Books.

Sommerville, J. (1991) *Filmer: Patriarcha and other Writings*, Cambridge: Cambridge University Press.

Further Reading

Stephen Buckle, *Natural Law and the Theory of Property* (Oxford: Oxford University Press, 1991) situates Locke among other seventeenth- and eighteenth-century figures. Eric Mack, *John Locke* (London: Continuum Publishing, 2009) provides a more extended explication of Locke along the lines of this essay. A. John Simmons, *The Lockean Theory of Rights* (Princeton: Princeton University Press, 1992) and *On the Edge of Anarchy* (Princeton: Princeton University Press, 1993) provide a comprehensive study of Locke's political thought. Richard Tuck, *Natural Rights Theories* (Cambridge: Cambridge University Press, 1979) traces pre-Lockean rights theory.

8
ROUSSEAU

Christopher Bertram

Jean-Jacques Rousseau's principal contribution to political philosophy is through his work *Du Contrat Social (Of the Social Contract)* in which he introduces and defends the idea that a legitimate state is one whose members are governed by their own general will. In this, he provides a distinct alternative to the social contract theories advanced by Hobbes and Locke: to Hobbes because he rejects the idea that sovereignty can be represented rather than exercised by the people directly, to Locke because the rights citizens enjoy issue from their own choice rather than being prior and independent constraints on the legitimacy of government. Rousseau's political philosophy is embedded in broader concerns in social theory, history, moral psychology and philosophical anthropology and in his obscure doctrine that human beings are good by nature but corrupted by society. His political philosophy is one of his attempts to show a means by which modern human beings might yet institute a set of relationships conducive to the good of the type of creatures that they are. His work is crucial to modern understandings of democracy, legitimacy, citizenship and public reason and has inspired theorists of direct and participatory democracy, though his own enthusiasm for democratic governance in his native Geneva was somewhat limited.

1. Biography

Rousseau was born in 1712 in Geneva, an independent city-state. The Genevan state had been refounded by the theologian Calvin in 1543 with Calvinist Protestantism as its established religion. It was nominally subject to the sovereignty of its citizens. In practice, however, it was an oligarchy dominated by the few families from which membership of its effectively governing body, the *Petit Conseil*, were drawn. Even the citizens as a whole were but a small proportion of the population of the state, most of whom were immigrants from neighboring territories or the descendants of such immigrants. Since Rousseau's father Isaac was a citizen, Jean-Jacques would normally inherit that status on his majority so long as he remained in communion with the established church.

Rousseau documents his own childhood and life in his autobiographical *Confessions* and in a number of other works, which together make a major contribution to the art of modern autobiography and to our modern understandings of selfhood. Following the death of his mother shortly after his birth, Rousseau was brought up by his father until Isaac was forced to flee the city as the result of a dispute. Thereafter, Jean-Jacques was pensioned with a clergyman and his family outside the city before being apprenticed to an engraver. At sixteen, he left the city to escape his abusive master and came under the influence of a Roman Catholic convert, Mme de Warens, who sent him to Turin

for instruction in the Catholic faith. After a period working for de Warens and becoming her lover, Rousseau then embarked on a series of employments, including working as a tutor in the extended family of the Enlightenment philosopher Condillac. Travel to Paris, in the vain hope of winning approval for a novel numerical scheme of musical notation, brought Rousseau into the circle of figures such as Denis Diderot, co-editor, with the Marquis d'Alembert, of the *Encyclopédie*. He also held a brief appointment as secretary to the French Ambassador to Venice, and experience of that republic apparently gave him the idea to write on political institutions.

An obscure middle-aged autodidact on the fringes of Enlightenment salons, Rousseau achieved sudden fame and notoriety when he entered, and won, an essay competition set by the Academy of Dijon in 1749. The resulting work, the *Discourse on the Sciences and the Arts*, argued that progress in the arts and sciences had come at the expense of the moral integrity of the human species, a thesis that Rousseau then defended in a series of replies to objections. He gained further fame as a contributor to a variety of fields of intellectual and artistic endeavor, including the composition of a successful opera, *Le Devin du Village*, in 1753. The social claims made in the *Discourse on the Sciences and the Arts* were refined and expanded in the comprehensive and sophisticated theory of human development he advanced in the *Discourse on the Origins of Inequality Among Men* (1755). Despite growing tensions with Diderot and with the circle around Voltaire, Rousseau became an increasingly prominent and successful figure. His novel *Julie* was published in 1761 (and became one of the most read fictional works of the eighteenth century), and was followed in 1762 by his treatise on education, *Émile*, and *On the Social Contract*.

The publication of Rousseau's most intellectually sophisticated works should have been the high-point of his career, but *Émile* and the *Social Contract* provoked extreme hostility and Rousseau was forced to flee to escape imprisonment or worse. *Émile* was condemned in both Paris and Geneva, principally for its heterodox religious content. The *Social Contract* was also condemned in Geneva where the ruling *Petit Conseil* took it to be not just a work of abstract speculation but also an intervention in the city-state's persistent constitutional wranglings. After spending some time in Switzerland under the protection of Frederick the Great of Prussia, Rousseau travelled to England at the invitation of David Hume in 1766 and spent several months there. Notoriously, he started to suffer from episodes of acute paranoia and believed Hume to be part of a plot to have him killed.

Returning to France in 1767, Rousseau spent most of the rest of his life in obscurity, but continued to work both on his autobiographical writings and, at the request of nationalist activists, on constitutional plans for both Corsica and Poland, which carry enduring interest as attempts by him to make his political principles concrete for particular states. Rousseau died in 1778, but received posthumous recognition as an inspirer of the French Revolution in 1794 when his remains were transferred to the Panthéon in Paris.

2. Moral Psychology and Speculative History

Rousseau's political philosophy, at least as outlined in the *Social Contract*, tacitly relies on views about human psychology, the good, and the course of development of both the human species and the human individual that he sets out elsewhere in various places, but most systematically in the *Discourse on the Origins of Inequality* and *Émile*. The details

of these ideas vary somewhat among the various presentations but the general shape they take is reasonably consistent in his mature writings. Importantly, they have often been the subject of acute misunderstanding, leading to a view of Rousseau as hankering after the restoration of some primitive state of society.

Central to Rousseau's view is the idea that the human good consists overwhelmingly in the achievement of relations of equality and respect with others, an achievement that requires granting to others and receiving from others recognition as free and equal. Other human aims, such as the satisfaction of the desires that we might happen to have or the acquisition of material goods or resources are of much less importance, and their pursuit may divert us from what truly matters. Tragically, for Rousseau, the very historical and social processes that bring this relational aspect of our being to the fore—thereby making genuine morality, freedom and responsibility possible—also make it overwhelmingly likely that this craving for recognition will be frustrated and that modern humans will fall prey to reactive attitudes such as anger, resentment, vanity and jealousy, and that these attitudes will lead them into mutually destructive conflict (see Dent 1988 and Neuhouser 2008).

Rousseau argues for two principal strategies for escaping these corrosive features of modern life. The first, found mainly in *Émile*, looks at the way in which a child's psychological development might be structured so as to immunize it from these aspects of life and make possible a satisfactory love-relationship in adulthood. The second is a political strategy that is realized in the *Social Contract*, whereby attaining the status of citizen in a just republic would, by making all subject to a law that they themselves have chosen, free individuals from the kind of oppressive dependence on the particular desires and choices of others that fosters such unhappiness. To understand this, we need to look at Rousseau's account of how this characteristically modern craving for recognition first emerges.

In the *Discourse on Inequality*, Rousseau outlines a speculative history of the human race. It is unclear how far he intended this to approach any kind of empirical accuracy, though he did draw on contemporary scientific accounts to some degree. Against Hobbes, Rousseau rejects the idea that human beings are naturally drawn into conflict with one another; against thinkers influenced by Aristotle, he denies that human beings are naturally sociable. Instead, he postulates that the original humans were solitary creatures who wandered through the forests, occasionally meeting to copulate, but with no real social connection. Like other animals, human individuals had natural powers proportioned to their needs, which were essentially for food, warmth and shelter. The lives of such "people" were essentially governed by their natural instincts.

These natural instincts or drives were twofold: first, they were guided by what Rousseau calls *amour de soi* to care for their own good *qua* natural beings; second, they were subject to an impulse to relieve the suffering of others provided they could do so without prejudicing their own survival. This second impulse Rousseau calls *pitié* and in some of his accounts, most notably *Émile*, he presents it as a kind of other-directed modification of *amour de soi*. Additionally, humans differ from other animals in the fact that they are not tied instinctually to particular ways of satisfying their needs (such as a single food source) and in being able to learn and acquire new means to secure their ends.

Very gradually, these creatures are drawn into elementary forms of co-operation around such activities as hunting, but the most significant change happens when they begin to establish settled communities. The catalyst for this change is sex, and sexual competition. People start to compare potential mates for the first time and also to real-

ize that they are the subjects of evaluation and comparison. This gives rise to a sense of selfhood, and a new mode of being in the world, a consciousness of self as appearing in the eyes of another and of being loved, desired, or spurned and rejected. People compare themselves to others, seek preferment, and derive a sense of their own value from the valuation that others place on them. Rousseau calls this new feature of our psychology *amour propre*: it is an impulse to look after one's interests as a social being, to be loved, desired, respected.

It is easy to read the *Discourse on Inequality* and to form the impression that the emergence of *amour propre* is, for Rousseau, a wholly regrettable development. But this is a mistake. The first societies in which *amour propre* makes its appearance are, for him, the happiest ones in human history. People are jealous, angry and resentful for the first time, and apt to go as far as murder from such motives, but they also have a new self-awareness and are creatures capable of reason and morality; they also have needs that are still easy to satisfy from the bounty of nature alone. Unfortunately, these societies are not stable, and further technological and social developments produce changes which make people increasingly interdependent, change their needs, and exacerbate their longing for recognition whilst making it ever harder to satisfy.

The prime culprits here are agriculture, metallurgy and the division of labor. Human beings increase their productive power but in ways that change their relationships to one another. Being more specialized, they have to secure the help of others to satisfy needs that have increased to match their greater wealth, needs that can no longer be satisfied by solitary effort. To protect the farms that they have planted they need to establish systems of right and property that exclude those who have not contributed to production. And to co-ordinate their new systems of production, they have increased recourse to hierarchies of authority. People now need one another both materially—in order to be fed, sheltered and kept warm—and psychologically, and the interplay of the two proves toxic.

We can see how this works by taking some two-person cases. One person needs the co-operation of another to secure some end, such as getting enough to eat. He therefore has an incentive to act towards that other person so as to secure that person's co-operation. Perhaps the need is reciprocal because the two need to trade to achieve their goals. Now each of them has such an incentive, and each knows that the other has it, and so on. Each also craves respect, as a component of their *amour propre*, but cannot be sure that the other is sincere in the expression of it, because each knows that the other has good reason to feign respect even when he does not feel it. In hierarchies, with superiors and subordinates, the effect is magnified. Now the superior is in a position to command expressions of respect and affirmations of his superior worth and subordinates may be obliged to comply. But precisely because these expressions are obligatory for the subordinate, the superior cannot take genuine satisfaction in them as confirmation of his superiority. Again, his *amour propre* is left unsatisfied. Everyone in this system becomes preoccupied with how they appear in the eyes of others and how they compare to others; they deform their very selves in order to be pleasing, but cannot achieve the goal they most prize because they cannot be sure that the good opinion that they need is a genuine expression of recognition rather than a cynical piece of self-serving behavior. Rousseau sees the modern humans trapped in this system as suffering from an *amour propre* that has become "inflamed": they are prone to anger, jealousy, contempt, resentment and frustration because of the elusive nature of their central psychological need.

At the same time that *amour propre* is taking on this form, the original impulse that led to people to care for one another, *pitié*, is losing its effectiveness, though it is never

extinguished. *Pitié* is not, for Rousseau, a moral sentiment, because it is based in instinct rather than reason, but it can play something of the role of morality by reducing conflict and inducing people to care about the suffering of others. Where humans' interest in recognition is dominant, however, the reactive attitudes very often drown out the voice of compassion. As people have become more sophisticated, moreover, they have acquired the capacity for complex reasoning. In principle this makes genuine morality possible, but in practice it permits people to devise for themselves persuasive rationalizations of self-interested conduct and to give themselves reasons to ignore the promotings of *pitié* as merely sentimental. Only when self-interest is not at stake, such as when people go to the theater—or, nowadays, watch a TV soap-opera—do people allow their sympathetic imagination to express itself and feel for the victims of pain or injustice. But this too, according to Rousseau, is problematic because it allows them to take satisfaction in the idea that they are basically compassionate and decent people when this facility is practically inert.

By the end of the *Discourse on Inequality* Rousseau has reached a point not unlike Hobbes's state of nature, where human beings are brought into conflict by the forces of competition, diffidence and glory. But he is keen to claim that this is an artificial and historical state-of-affairs rather than something rooted in the nature of human beings. Indeed, he insists that human beings are "naturally good" by which he means not that they have the natural capacity for *moral* goodness but that the proportionate relationship between powers and needs that the original humans possessed, coupled with *pitié*, means that there is nothing in human nature that impels us to conflict and cruelty. The Hobbesian condition that has developed, however, calls for a solution, as people tire of lives that are brief and violent. Unfortunately, Rousseau argues in the *Discourse on Inequality*, the most likely outcome is a class state, dominated by the propertied, who propose to all a system of equality before the law which in reality serves to concretize their dominant social position. The poor and propertyless agree to this because they crave an end to insecurity but fail to think through the fact that the new regime compels them to lives of proletarian dependence on the whims of the wealthy. Thus we arrive at a social contract as the foundation of a political order, but a fraudulent one, because it is not based on a genuinely general will of the people but on the interests of the privileged and ensures the continuation of the toadying and condescension, resentment and contempt and practical unfreedom that Rousseau sees as characteristic of modern life.

3. The General Will

Rousseau's solution to this fraudulent social contract is to propose a genuine one in its place. The genie of self-consciousness and *amour propre* cannot be put back and human beings cannot revert to a stupid and peaceful anarchy. Creatures such as we have become need to be subject to authority if they are to co-operate and prosper. But they also need a set of social arrangements that complement and to some extent reshape their human nature by freeing them from being oppressively dependent on the arbitrary desires of others, by granting them the recognition that they crave, and by allowing them the freedom that is a prerequisite of moral responsibility. Rousseau's aim in the *Social Contract* is therefore to delineate a political order in which interdependent citizens can be subject to authority and yet free at the same time, because the authority they are subject to is the product of their own wills. The central concept that Rousseau introduces to reconcile freedom and authority is that of the *general will*.

Rousseau's idea of the general will is suggestive and problematic at the same time. The idea that it conjures up is that of a self-governing democratic community, and that is indeed an important part of Rousseau's vision. However, Rousseau's account of it is systematically ambiguous between two different conceptions, and this ambiguity has generated a great deal of the controversy that has dogged the interpretation of his work. The ambiguity in question is between an idea of the general will as the democratic decision of the citizenry, and an alternative understanding where it is some kind of transcendent principle corresponding to the objective public interest of the community, which holds quite independently of what citizens believe or decide. Which of these views of the general will is preferred is crucial for our understanding of Rousseau's success or failure in his goal of reconciling freedom and authority, although each has its problems. If citizens can be coerced by the law in the name of some objective principle that they do not recognize, then a defender of Rousseau's claim to reconcile freedom and authority will need to rely on some surprising and extreme views about freedom (views that have some license from Rousseau's rhetoric). If the general will is simply what people decide, his task is arguably easier, but he still confronts the problem that minorities in the state will be coerced to comply with laws they opposed, a familiar problem in democratic theory.

We can get a handle on what the general will is by asking what it wills. The answer is that it wills laws that apply generally and universally to citizens and thereby limit their freedom. So the object of the general will is not the particular policy decisions that governments make but, rather, the framework of rules that define the terms of association of all the members of the state. Rousseau makes the case that these rules will not be onerous ones because of the process by which they are decided. The general will "comes from all" and "applies to all" and is therefore the result of citizens contemplating a set of restrictions and regulations that will also apply to themselves. Although each citizen has a "particular will" favoring his own private interest and which is often opposed to the private interests of other citizens, each realizes that they have no hope of legislating their private will and no interest in allowing others to enforce theirs. All have an interest in a fair system of regulation that respects their equal status and allows each as much freedom as is compatible with similar freedom for others. Rousseau believes that the generality and universality of this decision process will also have important socializing effects on the citizens, because they will gain the habit of thinking from the perspective of the community rather than simply looking out for their own private benefit: they will come to reason as citizens and as members of a public.

Rousseau's hopes for the non-onerous nature of the laws chosen by the general will depend on a number of background assumptions, which may not hold. The first of these is that citizens are sufficiently similar to one another in wealth, culture and conditions of life. The reason for this is the obvious one that laws with identical content can have very different implications for people in different circumstances. A law that forbids everyone from sleeping on benches in the park will not impact on the wealthy citizen living in a mansion but will limit the options of a homeless person. Similarly, planning and building laws that enable people to coexist in a dense urban environment could be onerous in the country where people are spaced far apart on large plots. If there is too much variation in wealth or type of life then citizens will have difficulty in taking the standpoint of the general will and of extrapolating from their own case to the wider experience of the rest of society. Worse, perhaps, they will see the ways in which laws might advantage or disadvantage the social group of which they are a member

compared to other groups, and it will be that sub-societal shared identity that guides them in their legislative choices. The second assumption is that citizens will only be guided in their choices by one of two things: either by a consideration of self-interest constrained towards impartiality by the nature of the choice situation or, as the result of long practice, by the internalization of a citizen identity. If, however, citizens have "nosy preferences" about how others should live their lives—such as particular views about which religion is conducive to salvation—then this could also lead to the passing of oppressive and onerous laws. Rousseau's concern for toleration in the chapter on civil religion—the final full chapter of the *Social Contract*—is perhaps, in part, meant to address this possibility.

In Book 2, Chapter 3 of the *Social Contract*, Rousseau makes a famous distinction between the general will and the "will of all" and argues that the vote of all the citizens does not necessarily reflect their genuinely common interest. Although many commentators treat it as such, it is important to realize that the phrase "the will of all" does not denote a technical term for Rousseau, so it cannot be assumed that whenever it occurs in his writings he is referring to this eventuality. In this discussion, however, he is making the point that democratic decisions are fallible and that the laws that citizens actually will may fail to correspond to their shared public interest. One important reason for this failure is the growth of inequality or diversity in forms of life just discussed, but another is a lack of citizenly virtue on the part of individuals. An interesting point about this is that it suggests a lack of confidence on Rousseau's part in the mechanisms of constrained choice that otherwise lead citizens to converge on the general will. The thought that citizens are led to will the general will because they think about the way in which universal laws will limit their own freedom does not require them to be other than self-interested; here the idea is that they need to have transformed their motivations and acquired a new social and moral identity.

This thought presents Rousseau with a problem; the individuals who come together to form a new society have been shaped psychologically by the competitive and individualistic society found at the end of the *Discourse on Inequality*. They are also in the grip of inflamed *amour propre* and hence predisposed to pride, vanity, resentment, jealousy and the whole panoply of the reactive attitudes, all of which makes it hard for them to see the good that mutually beneficial laws can bring to them. They are therefore very poor human material to forge a new and more co-operative society from. Rousseau provides a number of solutions, of which the most prominent is the figure of the lawgiver or legislator. The lawgiver, based on classical figures such as Lycurgus or Numa, is an inspirational figure who supposedly gets the citizens to will what is in their common interest, despite the fact that, not having been formed by just laws, they cannot see this. Since the new citizens are not amenable to rational argument, he must "persuade without convincing" (*Social Contract* 2.7). There are, to put it mildly, myriad problems with this stratagem; two seem particularly pertinent. The first, and most obvious one, is regress: if citizenly virtue is the product of subjection to just laws, then how did the lawgiver himself acquire it? Second, if citizens vote for the laws not from free deliberation about relevant reasons but, rather, because they have been non-rationally "inspired" to do so, in what sense can they be said to be subject to laws by their own free choice? In short, the lawgiver seems both impossible to believe in and actually subversive of what Rousseau has to say about freedom and legitimate authority. Having made these negative points, however, Rousseau should perhaps be given some credit for thinking about the social preconditions for a just society and about how these depend on some shared sense

of identity. This is really what the lawgiver tries to accomplish, and in the *Considerations on the Government of Poland* Rousseau has much to say about how shared activities such as festivals and games can help to foster a sense of solidarity such that the citizens who think about the general will, have a sense of thinking about something concrete, where "we, the people" is not a mere theoretical abstraction.

4. Government and Sovereign

An unusual feature of Rousseau's political theory, and one that leads to much misunderstanding, is his distinction between government and sovereign. For Rousseau, the people are sovereign and only they have the authority to bind themselves with general and universal laws. However, laws also need interpreting and applying to particular cases and Rousseau delegates this task to a special agency, the government, composed of individuals whom he calls "magistrates." The people as a whole decide whether there should be a law regulating, for example, political speech, and what its content should be, but it is the magistrates who are charged with deciding whether a particular instance of speech is permitted or prohibited according to the principles the people have approved.

Rousseau's conception of *sovereignty* is radically democratic. The whole people have to be involved in the making of the laws and, at least in the *Social Contract* (2.13), laws cannot be made by representatives or delegates. This leads many people to suppose that Rousseau is an opponent of representative government, but he is not, at least according to his use of that term, and the extent to which a Rousseauvian republic resembles a modern state with a representative assembly may well depend on exactly how we conceive of the distinction between sovereign and government working in practice. On the one hand, Rousseau sometimes argues that the laws should be simple and few in number. This suggests a view where the citizens decide on the basic laws governing their mutual relations, the constitutional structure as it were, and then have very little to do. On the other hand, he sometimes embraces a more active conception of citizenship in which the individual seems to be almost constantly engaged in legislating, and this would involve a very different practice from any state with which we are familiar.

In Book 3 of the *Social Contract* Rousseau examines the case for democracy, aristocracy and monarchy as forms of *government* and comes out in favor of aristocracy (government by the few). He rejects democracy on the grounds that an intimate involvement with the detail of particular cases will corrupt the judgment of the citizenry and cause their will to be less general when it comes to legislation. Monarchy, though efficient, leads to a system whereby office is held by those who can ingratiate themselves most effectively with the monarch. Aristocracy, by contrast, allows the top offices in the state to be occupied by the most talented individuals.

The question naturally arises of how democratic Rousseau's ideal state would be, taking sovereignty and government together. It is probably impossible to give a definitive answer on this matter given the ambiguity of many of the relevant texts, but proponents of Rousseau as a radical democrat have to argue away some difficult passages. For example, on one matter Rousseau seems utterly consistent throughout his writings: he reserves legislative initiative to the government and hence rejects the idea that ordinary members of the assembly should have the right to propose laws. He defends this point in the dedication to the *Discourse on Inequality*, in the *Social Contract* (4.1) and, perhaps most significantly, in the *Letters from the Mountains*, a piece where he is in open conflict with the Genevan oligarchy and highly critical of what he sees as their usurpation of

popular power. Citizen sovereignty is therefore highly circumscribed by the ability of the magistrates to shape the political agenda as they see fit, a power limited only by the right the citizens have to reject office-holders at periodic assemblies which cannot be cancelled (see Fralin 1977).

5. Does Rousseau Solve the Problem He Sets for Himself?

Rousseau's aim in the *Social Contract* is to reconcile freedom and authority, so that each person both enjoys the protection of the state but obeys "only himself and remains as free as before" (1.6). He claims that the act of association involves an exchange of "natural freedom" for "civil freedom" and also gains "moral freedom" involving "obedience to a law one has prescribed for oneself" (1.8). Notoriously, in a deliberately paradoxical formulation, he tells his readers that the citizen compelled to obey the general will is thereby "forced to be free" (1.7).

Rousseau's unclarity on some of these claims within the text of the *Social Contract* has led some commentators (e.g. Berlin, in his 1958 lecture, published in Berlin (1969)) to present him as a proponent of "positive liberty." But in fact, Rousseau's commitment is to something like "republican freedom," according to which a person is free so long as they are not subject to the arbitrary will of particular private persons. Subjection to the law, applying impartially to all, frees people from this direct dependence on others, according to Rousseau. Since the natural freedom people enjoyed in the state of nature has been eroded over the course of human development to culminate in a nightmare of mutual personal dependence and subjection (both material and psychological) at the end of the *Discourse on Inequality*, it is hard not to agree that mutual constraint within a framework of law and positive rights is an improvement. Rousseau's parallel claim that citizens are subject to a law they have chosen themselves is more difficult to defend. At the beginning of Book 4 of *Social Contract* he argues that citizens who find themselves in a minority simply thereby discover that they were mistaken about the content of a general will which they will anyway. That seems altogether too simplistic a response. Moreover, since the legislative process may be guided both by the non-rational persuasion of the lawgiver and the initiative of the magistrates, it can seem as if the citizen's conscious selection and rational approval of laws counts for less than Rousseau openly claims.

6. Impact and Influence

Despite their ambiguity, Rousseau's writings have been, and remain, enormously influential across whole swathes of human endeavor. In political life, the idea of the general will formed part of the ideology of the French Revolution, although it can hardly be said that the Jacobin dictatorship or even the National Assembly met Rousseau's criteria for popular sovereignty (see Wokler 2000). In philosophy, Kant acknowledged Rousseau's influence on his moral theory and Hegel's theories of both psychological development and the state can be read as direct developments of Rousseau's ideas (Neuhouser 2000). Modern political theories, such as John Rawls's theory of justice, also bear the stamp of Rousseau's influence, with the choice apparatus of the original position being an idealized and reworked form of the constraints that led to the emergence of the general will in Rousseau's theory. Even though three hundred years have passed since Rousseau's birth, it seems likely that his ideas will continue to be a fertile source of thinking about the problems of justice, legitimacy and human association.

Related Topics

Liberalism, Republicanism, Contractualism and Political Liberalism, Authority and Legitimacy, Democracy, Rights

References

Berlin, I. (1969) *Four Essays on Liberty*, Oxford: Oxford University Press.
Dent, N. (1988) *Rousseau*, Oxford: Blackwell.
Fralin, R. (1977) *Rousseau and Representation*, New York: Columbia University Press.
Neuhouser, F. (2000) *Foundations of Hegel's Social Theory*, Cambridge, MA: Harvard University Press.
—— (2008) *Rousseau's Theodicy of Self-Love*, Oxford: Oxford University Press.
Wokler, R. (2000) "The Enlightenment: The Nation State and the Primal Patricide of Modernity," in N. Geras and R. Wokler eds., *The Enlightenment and Modernity*, London: Macmillan.

Further Reading

Bertram, C. (2004) *Rousseau and the Social Contract*, London, Routledge. (A commentary on the *Social Contract*.)
Rousseau, J.-J. (1997) *The Discourses and Other Early Political Writings*, ed. V. Gourevitch, Cambridge: Cambridge University Press. (A useful selection of Rousseau's writings.)
Rousseau, J.-J. (2012) *Of the Social Contract and Other Political Writings*, ed. Christopher Bertram, trans. Quintin Hoare, London: Penguin. (A selection of Rousseau's later writings.)

9
HUME AND SMITH ON JUSTICE

Stephen Buckle

1. The Relation between Hume and Smith

At first glance, the philosopher David Hume and the economist Adam Smith seem to have little in common: the former a noted skeptic who, in the words of Bertrand Russell, destroyed empiricism by rendering it consistent (Russell 1961: 634); the latter the godfather of modern economics, the apostle of free trade and of the legitimacy of economic self-interest. However, these orthodoxies are seriously misleading. There is in fact a close connection between the two: Hume's philosophy propounds an evolutionary account of justice and social progress on which Smith builds; and Hume's account of the relation between justice and virtue shows how to resolve the apparent disjunction in Smith's thought between moral sympathy and economic self-interest. Nevertheless there are differences: Hume's theory has a markedly Epicurean character, emphasizing the artificiality of justice and its roots in social utility; Smith, in contrast, develops a broadly Stoic conception grounded in the natural regard for self-preservation.

To begin with biographical details: Hume, twelve years older than Smith, published his first great work, *A Treatise of Human Nature* (1739–40), as Smith was graduating from the University of Glasgow, and taking a scholarship to Oxford. Smith was unimpressed by Oxford, and the feeling appears to have become mutual, because he was disciplined for possessing a copy of Hume's *Treatise* in his room. (The work had already acquired a reputation for religious unorthodoxy, and was regarded as unsuitable reading material for a student of an institution whose main purpose was to train the clergy.) Smith not only possessed the book; he read it carefully. We know this because while he was there he began work on an essay (not published in his lifetime) which shows clear signs of Humean influence. This essay, entitled *The Principles which Lead and Direct Philosophical Enquiries; Illustrated by the History of Astronomy*, considers the human capacity for scientific knowledge by mobilizing a psychological theory derived from Hume. Unsurprisingly, then, when he returned to Scotland he soon met Hume and a life-long friendship, both personal and philosophical, developed.

2. Hume (and His Precursors)

Hume's *Treatise* is commonly regarded as an examination of arcane questions in epistemology and morals, but it also offers an account of the development of the basic institutions of human society: justice, or the rules concerning what belongs to each

person; promising and contracts; and government and the court system. The account is markedly at odds with religious orthodoxies: it strips justice (and related social institutions) of any natural foundation, treating it as the result of a long and undirected process of social evolution. Hume built on these ideas in a series of short essays intended for a wide readership. The most important of them was the *Political Discourses* (published in 1752), which included pioneering essays on economics and politics. They exercised a profound influence on Smith, convincing him of the connections between economic growth and the creation of a well-ordered and secure society; without them, *The Wealth of Nations* is almost unimaginable. To see the connections properly, we need to begin with Hume's account of morality and justice.

In the *Treatise*, and more forcefully in *An Enquiry concerning the Principles of Morals* (1751), Hume argued that human morality is not the result of the rational discovery of eternal truths, but of natural human sympathetic responses to the actions and character of ourselves and others. The cardinal virtue of this moral responsiveness is benevolence: literally, the willing of the good, and thus the desire of the happiness of all. However, this natural responsiveness cannot explain the virtue of justice, since justice is concerned with the necessity of respecting the boundaries between people—in particular, of respecting their property. Justice is thus a "hands-off" virtue, in contrast to the "hands-on" virtue of benevolent concern for others' welfare. Hume brings out this difference by describing justice as "the cautious, jealous virtue," in contrast to the "generous" sentiments of benevolent fellow-feeling (Hume 1998: 3.3, Appendix 1).

How, then, is justice to be explained? To understand Hume's answer and its significance, it will help to place it in its historical context. The classical tradition stemming from Plato held both that it is rationality that fits human beings for social life and, therefore, that it is a society of a rationally organized kind that is fitted to human beings. This is to make justice a property of social relations that find their foundation in human rational nature. It is undoubtedly an attractive theory, and so it is unsurprising that it should have become so influential in western thought. In the modern era, it continued to attract adherents, most notably being reaffirmed in the founding work of modern international law, Hugo Grotius's *The Rights of War and Peace* (1625). Grotius held that human sociability depends directly on "the light of human understanding," which understanding is "the fountain of right," that is, of justice (Grotius 2005: Preliminary Discourse, VIII). He thus tied justice directly to a foundation in human rationality, making it a natural concern of human beings.

Grotius's work was composed in the middle of the ruinous Thirty Years' War (1618–48), and in this light can be understood to be an optimistic work, denying that the surrounding conflicts are natural to human beings. However, reflection on this conflict and the religious conflicts of his own day led Thomas Hobbes to reject Grotius's optimistic conviction of human sociability. His approach can be understood as the simple application of *modus tollens*: according to Grotius, human beings are sociable because they are rational; but they are not sociable; therefore they are not rational. He argued that human beings are not governed by reason, but by passions; passions which are so blindly acquisitive and competitive that they lead us inexorably into mutual conflict. The natural condition of human beings is thus not to live in a rationally organized society, but to find themselves in a "war of every man against every man," in which life is "solitary, poor, nasty, brutish, and short" (Hobbes 1994: 13.13, 9).

This bleak vision appalled most of Hobbes's readers, not least because it seemed to deny the Christian conception of the Deity. Certainly the Hobbesian natural condition

is no Garden of Eden, and so his view met with a string of attempted refutations. Typically, these responses sought to re-establish the human capacity for social life by reaffirming human rationality. They took it that, although Hobbes's inference was sound, his premises were false (see e.g. Locke 1975: II.2). In one case, however, Hobbes's argument was met with the opposite strategy—of accepting his central premise, that human beings are ruled by passion, but denying his inference from it.

In *An Inquiry into the Original of our Ideas of Beauty and Virtue* (first published in 1725), Francis Hutcheson argued that human beings are naturally sociable because they possess a natural moral feeling which disposes them to mutually beneficial relations with others. He called this natural feeling the "moral sense," meaning by it an instinctive tendency in human beings to approve of all actions which conduce to human happiness. (He is the originator of the expression, later made famous by Bentham, that moral principles aim at "the greatest happiness for the greatest numbers" (Hutcheson 2008: 125).) This instinctive feeling, he further argued, was the means by which human beings fitted themselves into the divine providential order. It was thus a divinely bestowed instinct for social and indeed cosmic harmony, and, therefore, the foundation on which all virtues are built—justice included (Hutcheson 2008: 197, 181–3).

Hutcheson did recognize, however, that the implications of justice do not in every instance conform to the principle of benevolence that is the moral sense. He called these non-conforming cases "external rights." They include such things as the necessity of accepting a rich miser's right to demand, whenever it suits him, the return of money lent by him to an industrious poor tradesman, and other such cases where justice plainly opposes the promptings of the benevolent regard for the greater happiness. Hutcheson's response to these cases was to point out that, although each instance is at odds with the moral sense, the universal system of regard for rules of justice was not. Therefore, even "external rights" could be fitted into his benevolent system (Hutcheson 2008: 185).

Hutcheson's argument depends on supposing that justice comes into being all at once: that it is, from the very beginning, a system. If it is not, then the particular instances cannot be justified by reference to the whole. If individual cases come first, if systems of justice are the result of social evolution—and the subordination of reason to feeling suggests they *must* be, since feelings do not plan ahead—then justice must arise quite independently of any benevolent promptings. Moreover, Hutcheson's appeal to the providential role of the moral sense commits him to a belief in final causes, and so puts his theory at odds with the emphasis on efficient causation alone that was the hallmark of the new natural science developed by Galileo, Descartes and Newton.

Hume takes off from these problematic features of Hutcheson's theory. He accepts Hutcheson's argument that morality begins in feelings, but rejects his identification of "moral sense" with a natural instinct geared to the greater good, a cog in a providential mechanism (Hume 1932: I, 33). Instead, Hume argues that this "moral sense" is no more than our sympathetic responsiveness to others. It is not an instinct directed at a goal, but a mechanism which responds blindly to whatever impinges on it. It thus works locally without regard for an overall good. These responses can then be modified by reason—which is to say that we learn from experience the likely consequences of our responses, modify their expression accordingly, and denominate these modifications "rational." This means that morality develops in response to experience of its utility (Hume 1998: 5.42ff.).

Justice, argues Hume, derives entirely from the discovery of its utility. The benevolent feelings play no role in its origin (Hume 1998: 3.1). Benevolence is limited by the scope

of our sympathetic responsiveness, and this rarely extends beyond our immediate circle—whereas justice is concerned with relations with strangers no less than with one's own. However, where benevolence has little effect, human beings are governed by the self-interested passions. This suggests that justice arises from self-interested exchanges between individuals who are under no ties of obligation to each other, and so only gradually develops into a system governing all, as its advantages become clear (Hume 2007: 3.2.2).

Hume builds his position on this thought, and on another, no less important. The individuals who engage in these exchanges are beings of limited rational capacity, with no clear sense of the longer-range consequences of their actions: they develop a sophisticated understanding of their situation and actions only as they go along. This aspect of Hume's theory shows the background influence of Epicurean accounts of human social and psychological evolution. The original of these was the Roman poet Lucretius's philosophical poem *Of the Nature of Things*; a more immediate example was Bernard Mandeville's notorious work, *The Fable of the Bees* (first published in 1714). These accounts, by emphasizing the hand-in-hand development of the human psyche and human society, imply that social outcomes are not the product of prior intentions, but are unintended (if unavoidable) consequences. The point of recognizing this background debt is that it brings out the significance of Hume's most striking claim: that justice arises not from a contract of any form, but by convention. Lucretius's story is of creatures which are human only in the most limited sense: they possess the same passions as we do, but are entirely lacking in social affections and rational foresight. Hume denies the strict application of the first point—he holds that human beings have never been solitary, having always lived in families (Hume 2007: 3.2.2.4–5, 14)—but he accepts that this makes little overall difference to the problem of forging social union. Such union depends on the establishment of general rules, but the lack of rational foresight means that a social contract—in which the causes of social disharmony are worked out in advance, and then agreed to by all parties—is impossible. The necessary rules can only grow up slowly, through individuals coming to a shared sense of what is to their mutual benefit, and this sense steadily extending its scope.

The mutual recognition, and mutual expression, of the usefulness of these rules is the convention on which the rules of justice are built. Initially it applies only locally, to those parties who have expressed this sense of mutual benefit to each other, but, because it is so useful, and so readily recognizable as such, it spreads over all, thereby evolving into the systems of justice that govern whole societies (Hume 2007: 3.2.2.9–10). As its influence spreads, so also does the sense of its worth for all, rather than just for the individuals. In this way, a convention originally approved for self-interested reasons comes to be approved for impartial reasons as well, and so the rules of justice come to be recognized as morally good: justice becomes a virtue (Hume 2007: 3.2.2.23–4).

Hume's account of justice is, then, a story of the evolution of informal local arrangements between self-interested individuals into a system of rules which regulate the whole society; rules which, because of their evident utility for all, come to win approval from an impartial point of view, and so become an object of moral approval. Justice is not, then, a natural human concern directly grounded in a rational nature, but an artificial virtue, the product of human artifice. It is the unintended consequence of useful local arrangements; arrangements so useful that they have become entrenched in human social consciousness, which entrenchment itself explains the mistaken supposition that it is a natural feature of the human psyche. The whole story is sufficiently confirmed by

human history, which shows that justice grows alongside the education of human nature: "History, experience, reason sufficiently instruct us in this natural progress of human sentiments, and in the gradual enlargement of our regards to justice, in proportion as we become acquainted with the extensive utility of that virtue" (Hume 1998: 3.21).

Hume's target in all of this is the idea that justice arose as the result of a promise or contract. Indeed, he argues that promising itself cannot be understood except as the result of the same kind of process that gives rise to rules of justice. Like justice, promising is a social institution that arises through conventions, as a sense of the utility of binding agreements, brought about by conventionally established public signs, steadily comes to be understood and widely shared. So it is a learned sense of the utility of the institution, discerned as the practice develops, that explains the institution of promising—not a natural awareness of the importance of giving one's word (Hume 2007: 3.2.5).

Justice and other related institutions are thus to be understood as consequences of social development. They are children of social utility. This means that the specific content of justice's rules will vary with different circumstances. Hume does not emphasize this in the *Treatise*, but in the *Enquiry concerning Morals* it is given a prominent role. Although respect for what is the property of another is non-negotiable, nonetheless there is, and must be, great variety in "municipal," that is, local laws. Such variety is not without an underlying uniformity, however, because in all cases "the purposes … are everywhere exactly similar"—that is, they all exist for the sake of public utility (Hume 1998: 3.45).

In thus emphasizing variety within an overall framework of utility, Hume shows his Epicurean sympathies. Epicurus and his followers had treated justice as simply the result of agreements made for the sake of mutual advantage, and Hume's conventionalist account, in giving a minimalist version of agreement, is true to the evolutionary spirit of the Epicurean story. He does not, however, stop there, instead digging deeper to consider the circumstances which must obtain for justice to exist at all. These circumstances, he argues, are the relative scarcity of goods and the natural partiality of human beings to themselves and their friends—a partiality nevertheless modifiable into accepting general rules that work for the overall good.

The former, the scarcity of goods combined with human partiality, creates the need for rules to prevent chaos; and the human ability to adapt to the requirements of useful rules shows the possibility of abiding by them. Where these circumstances do not obtain, there justice does not exist: "Produce extreme abundance or extreme necessity: Implant in the human breast perfect moderation and humanity, or perfect rapaciousness and malice: By rendering justice totally *useless*, you thereby totally destroy its essence, and suspend its obligation upon mankind" (Hume 1998: 3.12). Hume supports the point by drawing attention to two cases where the thought is presupposed. In the first place, the ancient poetic fiction of the "golden age" presents that idyllic time as completely without "the punctilious distinction of mine and thine" because human goodness and natural abundance made it unnecessary; and, at the other extreme, the modern philosophical fiction of the (Hobbesian) state of nature is represented as a state of "mutual war and violence, attended with the most extreme necessity," and therefore also as a situation in which "No law was heard of: No rules of justice known: No distinction of property regarded: Power … the only measure of right" (Hume 1998: 3.14, 15). These two fictitious states of affairs thus teach that justice exists only where human nature and its circumstances avoid these extremes; that is, only where it is useful.

Hume's emphasis that justice, both in its origin and its development, reflects utility, or human social needs, gives his account a markedly progressivist character: the story of

human social rules is a story of progressive sophistication or improvement as new challenges come to bear. This becomes clear in the political and economic essays, where he directly opposes the pessimism of Mandeville and Jean-Jacques Rousseau. The former argued that human activity is based in selfishness and greed, and that society would collapse without them; the latter argued, in the prize-winning essay that shot him to fame, the *Discourse on the Sciences and the Arts* (published in 1750), that "luxury" (or as we might now say, consumer society) "is seldom found without the sciences and the arts, and they are never found without it," going on to argue that "luxury is diametrically opposed to good morals," because it produces a feverish activity that obstructs happiness and undermines both the social and the martial virtues (Rousseau 1997: 16–28).

Hume directly countered these views in one of the most influential political essays, "Of Luxury" (published in 1752; later re-titled "Of Refinement in the Arts"). He argues there that "the ages of refinement are both the happiest and the most virtuous." Refinement is beneficial to private life because "industry, knowledge, and humanity, are linked together by an indissoluble chain." Activity is the life-blood of the human spirit, overcoming the "languor and lethargy, that destroys all enjoyment." The spirits, once roused, are turned to all departments of life, even enhancing social relations: People "flock into cities; love to receive and communicate knowledge; … Both sexes meet in an easy and sociable manner; and the tempers of men, as well as their behavior, refine apace." Refinement also benefits public life by creating "a kind of storehouse of labor, which, in the exigencies of state, may be turned to public service." Thus prosperous societies are both the best-governed and militarily the best-equipped; and even martial virtue is not threatened, since the active life in such a society means that there is no enervating effect on mind or body. Luxuries can become vicious; but, even when this occurs, it is still preferable to the "sloth and idleness" characteristic of "barbarous nations," vices which, by breeding crime and revolt, "are more hurtful both to private persons and to the public" (Hume 1987: 268–80).

Given these views, it is obvious that a necessary task for such a viewpoint is to determine how prosperity is promoted and protected. This explains why Hume turned to economic questions; it also explains why a young man who had closely followed Hume's philosophical project and accepted its main outlines should himself have turned increasingly to economic questions.

3. Smith

This is the lens through which to view the intellectual career of Adam Smith. If it is not immediately obvious, it is Smith's own fault. He left to posterity two main works, *The Theory of Moral Sentiments* and *The Wealth of Nations*: the former a work in a broadly Humean spirit, emphasizing the role of sympathy in the formation of moral judgments; the latter a "very violent attack … upon the whole commercial system of Great Britain" (Smith 1977: 251), containing not only a defense of free trade against ill-advised government restrictions, but also, as the famous remarks on the "invisible hand" show, a justification of the place of self-interest in economic life (Smith 1976b: IV.ii.9). The apparently contrasting themes of these two productions have caused such perplexity that it was given its own name: "The Adam Smith Problem." There is no doubt that the two books breathe rather different air; but the sense of intractable opposition between the two was exacerbated by the fact that, although Smith had, all along, planned a linking volume, its incomplete state at the time of his death led him to have the manuscript destroyed.

To see how the missing volume would have filled in Smith's project we have to turn to the testimony of John Millar, one of Smith's students at the University of Glasgow (and later the Professor of Civil Law at Glasgow). Millar gave this report to Dugald Stewart, who incorporated it into his "Account of the Life and Writings of Adam Smith, LL.D." (1794). It runs as follows:

> Mr Smith was elected to the chair of Moral Philosophy. His course of lectures on this subject was divided into four parts. The first contained Natural Theology … The second comprehended Ethics, strictly so called, and consisted chiefly of the doctrines which he afterwards published in his *Theory of Moral Sentiments*. In the third part, he treated at more length of that branch of morality which relates to justice, and which, being susceptible of precise and accurate rules, is for that reason capable of a full and particular explanation.
>
> Upon this subject he followed the plan that seems to be suggested by Montesquieu; endeavouring to trace the gradual progress of jurisprudence, both public and private, from the rudest to the most refined ages, and to point out the effects of those arts which contribute to subsistence, and to the accumulation of property, in producing correspondent improvements or alterations in law and government. This important branch of his labours he also intended to give to the public; but this intention, which is mentioned in the conclusion of the *Theory of Moral Sentiments*, he did not live to fulfil.
>
> In the last part of his lectures, he examined those political regulations which are founded, not upon the principle of justice, but that of expediency, and which are calculated to increase the riches, the power, and the prosperity of a State. Under this view, he considered the political institutions relating to commerce, to finances, to ecclesiastical and military establishments. What he delivered on these subjects contained the substance of the work he afterwards published under the title of *An Inquiry into the Nature and Causes of the Wealth of Nations*.
>
> (Smith 1980: 274–5)

Millar's summary brings out clearly the overall nature of Smith's project, and the crucial linking role to have been played by the never-completed book on justice and legal development. It is thus fortunate indeed that Smith's students did not all follow his example: some of them took very extensive notes during his Glasgow lectures. Two such copies have come to light, and have since been published under the title *Lectures on Jurisprudence*, thereby providing a window on the project Smith himself never completed.

Millar's description is sufficient to show the broadly Humean nature of the project: it is an evolutionary account of the development of social rules through which human beings themselves progress from a "rude" to a "refined" condition; and the concern with the origins of wealth falls into place against the background of Hume's argument in "Of Refinement in the Arts." Millar's reference to Montesquieu does not indicate a contrary influence, but a broadly complementary one. Montesquieu's *The Spirit of the Laws* (1989 [1748]) provided a detailed account of the development of social institutions in response to differing circumstances, and, similarly, Smith pursued the historical detail that the theory required for its completion, but which Hume did not attempt. In this light, Smith's ambition for his book on law and government can be described as the attempt to complete the Humean project, albeit modified by the influence of Montesquieu's detailed researches.

Keeping the Humean connection in mind shows how "the Adam Smith Problem" can be solved. As Millar's summary shows, the subject matter of *The Wealth of Nations* is consequent upon the theory of justice, and so belongs in its broad sphere. From a Humean point of view, this would mean that economic virtue is an artificial virtue, arising originally from self-interest, and winning the approval of the moral sentiments only once the overall benefits of the whole system of economic rules are recognized. Individual self-interested acts win no moral approval unless they are recognized to be part of a system of socially beneficial rules. For Hume, then, whereas the benevolent virtues can be compared to the bricks in a wall, each of which contributes to the whole independently of the others, justice and its associated principles, in contrast, "may be compared to the building of a vault, where each individual stone would, of itself, fall to the ground; nor is the whole fabric supported but by the mutual assistance and combination of its correspondent parts" (Hume 1998: Appendix 3.5). This means that Hume is not committed to regarding self-interest as good, except in those circumstances where self-interested acts are part of a larger framework, the tendency of which is socially beneficial.

Smith's theory is not identical, but allows a similar resolution. He takes up a point made by Hume, and places it at the centre of his account: justice, unlike the other virtues, is recognized because of the pains, and consequent resentment, caused by its violation (Smith 1976a: II.ii.1.3–5; cf. Hume 2007: 3.2.2.24). It is thus "a negative virtue," satisfied by abstinence from unjust acts, rather than by the positive acts required by the other virtues: "We may often fulfil all the rules of justice by sitting still and doing nothing." Moreover, because pain is "a more pungent sensation than pleasure," justice enjoys the pre-eminent position amongst the virtues (Smith 1976a: II.ii.1.9; I.iii.1.3). This leads Smith to offer his own variation on Hume's theme:

> Beneficence … is less essential to the existence of society than justice … It is the ornament which embellishes, not the foundation which supports the building … Justice, on the contrary, is the main pillar that upholds the whole edifice. If it is removed, the great, the immense fabric of human society … must in a moment crumble into atoms.
>
> (Smith 1976a: II.ii.3.3–4)

The debt to Hume is plain, and reflects the shared purpose of insisting on the over-riding importance of justice. Nevertheless, it is the divergence signaled by the modified metaphor that is the more significant. By grounding justice directly in our natural response to pain, he renders it independent of an overall system of useful rules. He also thereby restores its status as a natural virtue. These differences are not accidental. They show Smith recasting Hume's Epicurean theory along Stoic lines: justice is the fundamental virtue because it is the virtue most intimately linked with our self-preservation.

Smith's account thus provides a standard for assessing self-interested acts, such as those between economically rational agents. Such acts are just whenever there is no "injury," no "real and positive hurt" to the relevant parties (Smith 1976a: II.ii.1.1–5). Smith is not, therefore, committed to regarding self-interest as good. It is good *only* in those circumstances where self-interested acts arouse no resentment from an impartial observer. The factors that arouse such resentment may be manifold, but they will focus in particular on transactions that cause hurt to particular persons. In all such cases, then, Smith's theory implies disapproval of self-interested actions. This means that

he could be expected to regard with a suspicious eye all those transactions in which there are recognizable asymmetries of power—and therefore the possibility of coercive dealings. Despite his hostility in *The Wealth of Nations* to the irrationalities of the economic restrictions of his time, he is not committed to approve of every exercise of economic self-interest. He defends the system of free trade, but he does so because of the benefits it delivers, not because of the inviolability of individual self-interest. He is, in fact, fully aware of the ineradicability of conflicts of interest, and thus of the ever-present possibility of abuse of the system by those able to manipulate it—as his various acidic comments about the conspiratorial tendencies of merchants suffice to prove (Smith 1976b: I.x.c.27; IV.iii.c.10). Where such conspiracies are able to cause harm to others, Smith's principles imply that they earn the resentment of the impartial spectator, and so fail the test of justice. Hence his conspicuous sympathy for the interests of the poor (Smith 1976b: I.viii.13, 36). The supposition that Smith shifts from a sympathy theory to a self-interest theory is therefore misguided; "the Adam Smith Problem" dissolves.

These references to the impartial spectator serve to introduce the most distinctive methodological element in Smith's theory. Once again it shows his debts to, and revisions of, Hume. Hume had argued that our moral judgments are based in sympathetic feelings, and that they are impartial in the sense of reflecting a disinterested perspective. Smith combines these two elements in his theory of the impartial spectator, guaranteeing disinterestedness by transferring the sympathetic basis from ourselves to the impartial spectator, a fictional creature we must first construct before moral judgments are possible. This construction is forced upon us by the experience of trying to live with others: we learn that peaceful co-existence depends on taking up a point of view we can all share, and it is the disinterested perspective that satisfies this requirement. The development of this viewpoint is thus the development of moral consciousness (Smith 1976a: III.i.6; III.iii.22).

Smith applies this thought directly in the *Lectures on Jurisprudence*. He explains the origins of obligations arising from contracts by reference neither to private acts of will, nor to utility, but to what "an impartial spectator would readily go along with." He adds that this makes "the expectation and dependence" of the parties the central issue. Similarly, he explains rights in terms of causes of the spectator's resentment, and injustice, in its turn, as the violation of rights (Smith 1978: LJ(A), ii.42–5, 56; i.9). The key feature of this standard is disinterestedness and nothing more. The impartial spectator is not in possession of a philosophical formula, nor privileged in other ways: the spectator is not outside time or space, but a historically situated member of human society. Smith's standard is, then, ideally designed for elucidating a natural history of human social institutions and legal practices. It is also, thereby, a critical weapon, since this natural history can then be counter-posed to actual history, including to contemporary society, to expose injustice and other evidence of the effects of coercive interests.

The earlier of the two sets of student notes shows Smith pursuing the former task in great detail. He introduces there the "four stages theory" of human development: the view that human society must pass through four distinct forms of social organization: the age of hunters; of shepherds; of agriculture; and, finally, of commerce (Smith 1978: LJ(A), i.27). Each of these stages is characterized by distinct forms of property rights: hunting society respects property rights only in objects of immediate possession; shepherding or pastoral society respects rights also in flocks, pastoral territories, etc., and also recognizes rights arising from accession (such as the milk and offspring of legitimate

possessions); and agricultural society is characterized by its focus on property in land and its uses (Smith 1978: LJ(A), ii.97, i.64, 52–3). The critical side of Smith's account is shown by his acid comments on entailed property, which he denounces as a "foolish" form of property by succession (Smith 1978: LJ(A), i.144). The impartial spectator cannot approve of it because it prevents the rational transfer of land in response to changing circumstances. Needless to say, restrictions on legal transfer would be all the more "foolish" in the age of commerce.

The approach adopted in these earlier lectures is well suited to the task of elucidating the legal developments necessary in the course of human history. What it is not so well suited to is showing the interdependence of property law and forms of government; it also tends to push the critical task to the end of his long exposition. So, in the second set of lectures, Smith reversed his order of exposition, beginning with government and then moving on to property and other rights (Smith 1978: LJ(B), 11). This gave much greater prominence to the account of the origin of modern European legal and political forms.

This is the story of how the feudal system was destroyed by modern commerce, and it returns us to a familiar theme. Smith shows how the feudal lords' love of luxury led them to redirect their fortunes from their armies of retainers to indulgence in luxury goods. By doing so they freed their dependants and created the conditions for the rise of the merchants and their towns. Luxury thus destroyed the power of the nobility. In the short run, this led to arbitrary (absolute) monarchical rule; but the monarchs' dependence on the people for their income saw concessions extracted in return, so in the longer run it created "a rational system of liberty" in Britain (Smith 1978: LJ(B), 55–63).

Smith retells this story in Book III of *The Wealth of Nations*, there sharpening the point. He says: "A revolution of the greatest importance to the publick happiness, was in this manner brought about by two different orders of people, who had not the least intention to serve the publick." The lords acted from "the most childish vanity," the merchants from "their own pedlar principle of turning a penny wherever a penny was to be got." In this way "the folly of the one, and the industry of the other" meant that, in Europe, "commerce and the manufactures of cities" had resulted in a system of "order and good government, and, with them, the liberty and security of individuals" (Smith 1976b: III.iv.17, 18, 4).

Smith thus re-affirms Hume's story that luxury, or refinement in the arts, brought about a society that was not only more prosperous, but also freer and better-governed, than the mistakenly admired poorer societies it replaced. His more detailed account allows him to find another task for the invisible hand, in his account of the rise of commercial society and its liberties from the very differently directed actions of the medieval lords and merchants. But the links with the Humean account remain visible. Although he rejects Hume's argument that justice is an artificial virtue focused on utility, he endorses Hume's treatment of it as a human product, arising in response to human need and developing accordingly. For both it is also a story of progress: of the wider benefits for good order and freedom from arbitrary power that prosperity makes possible. As such it is certainly an optimistic story; a story which subsequent economic and political transformations would put to the test.

Related Topics

Hobbes, Rousseau, Social Evolution

References

Grotius, H. (2005) *The Rights of War and Peace*, ed. R. Tuck, Indianapolis: Liberty Fund.
Hobbes, T. (1994) *Leviathan*, ed. E. Curley, Indianapolis: Hackett Publishing.
Hume, D. (1932) *The Letters of David Hume*, 2 vols, ed. J. Y. T. Greig, Oxford: Clarendon Press.
—— (1987 [1985]) *Essays Moral, Political, and Literary*, ed. E. F. Miller, Indianapolis: Liberty Fund.
—— (1998) *An Enquiry concerning the Principles of Morals*, ed. T. L. Beauchamp, Oxford: Clarendon Press.
—— (2000) *An Enquiry concerning Human Understanding*, ed. T. L. Beauchamp, Oxford: Clarendon Press.
—— (2007) *A Treatise of Human Nature*, 2 vols, eds. D. F. Norton and M. J. Norton, Oxford: Clarendon Press.
Hutcheson, F. (2008 [2004]) *An Inquiry into the Original of Our Ideas of Beauty and Virtue*, ed. W. Leidhold, Indianapolis: Liberty Fund.
Locke, J. (1975) *Two Treatises of Government*, ed. P. Laslett, Cambridge: Cambridge University Press.
Lucretius (1975 [1924]) *On the Nature of Things* [*De Rerum Natura*], eds. W. H. D. Rouse and M. F. Smith, Cambridge, MA: Harvard University Press.
Mandeville, B. (1988 [1924]) *The Fable of the Bees*, 2 vols, ed. F. B. Kaye, Indianapolis: Liberty Fund.
Montesquieu, C.-L. de S. (1989) *The Spirit of the Laws*, eds. A. M. Cohler, B. C. Miller and H. S. Stone, Cambridge: Cambridge University Press.
Rousseau, J.-J. (1997) *The Discourses and Other Early Political Writings*, ed. V. Gourevitch, Cambridge: Cambridge University Press.
Russell, B. (1961 [1946]) *History of Western Philosophy*, London: George Allen and Unwin.
Smith, A. (1976a) *The Theory of Moral Sentiments*, eds. D. D. Raphael and A. L. Macfie, Oxford: Clarendon Press.
—— (1976b) *An Inquiry into the Nature and Causes of the Wealth of Nations*, 2 vols, eds. R. H. Campbell, A. S. Skinner and W. B. Todd, Oxford: Clarendon Press.
—— (1977) *The Correspondence of Adam Smith*, eds. E. C. Mossner and I. S. Ross, Oxford: Clarendon Press.
—— (1978) *Lectures on Jurisprudence*, eds. R. L. Meek, D. D. Raphael and P. G. Stein, Oxford: Clarendon Press.
—— (1980) *Essays on Philosophical Subjects*, eds. W. P. D. Wightman and J. C. Bryce, with Dugald Stewart's "Account of Adam Smith," ed. I. S. Ross, Oxford: Clarendon Press.

Further Reading

J. Moore, "Hume's Theory of Justice and Property," *Political Studies* 24 (1978): 103–19, is an excellent introduction to Hume's theory. S. Buckle, *Natural Law and the Theory of Property: Grotius to Hume* (Oxford: Clarendon Press, 1991) offers a detailed account of Hume's debts to Grotius and Hutcheson (although it over-emphasizes continuities). R. Hardin, *David Hume: Moral and Political Theorist* (Oxford: Oxford University Press, 2007), addresses Hume's theory from the different angle of modern game theory. E. J. Hundert, *The Enlightenment's* Fable*: Bernard Mandeville and the Discovery of Society* (Cambridge: Cambridge University Press, 1994), shows the impact of Mandeville's Epicurean project on eighteenth-century social theorizing. K. Haakonssen, *The Science of a Legislator: The Natural Jurisprudence of David Hume and Adam Smith* (Cambridge: Cambridge University Press, 1981), is a thorough treatment of Smith's project in the light of the *Lectures on Jurisprudence*. D. Lieberman, "Adam Smith on Justice, Rights and Law," in K. Haakonssen, *The Cambridge Companion to Adam Smith* (Cambridge: Cambridge University Press, 2006), is an excellent summary account in the same spirit. E. Rothschild, *Economic Sentiments: Adam Smith, Condorcet and the Enlightenment* (Cambridge, MA: Harvard University Press, 2001), sets Smith's work in its wider European context. Finally, N. Phillipson, *Adam Smith: An Enlightened Life* (London: Allen Lane, 2010) is a very readable account of Smith's life and work with an eye to his Humean debts.

10
KANT

Oliver Sensen

Immanuel Kant (1724–1804) is best known not for his political philosophy, but for his three *Critiques* of the powers of reason. Kant's first explicit work of political philosophy appeared when he was already 60 years old (1784); most of his writings on the subject are essays written for a wider audience, such as "Idea for a Universal History with a Cosmopolitan Purpose" (1784), "On the Common Saying: That May be Correct in Theory, but is of No Use in Practice" (1792), and "Toward Perpetual Peace" (1795), while his most systematic treatment of political philosophy is part of a treatise on legal philosophy, the *Doctrine of Right* (1797). (I refer to Kant's texts citing volume and page number of the Prussian Academy edition of *Kants gesammelte Schriften*, Berlin: de Gruyter, 1902–.)

Nonetheless, Kant's work has had a profound influence on contemporary political philosophers, such as John Rawls and Robert Nozick (Rawls 1971: 10, 121f., 221–7; Nozick 1974: 32). More importantly, Kant has crucially shaped the notions of autonomy, dignity, and respect for persons which are central to modern political life.

In his three *Critiques* Kant examines the powers the faculty of reason possesses independently of experience or a priori. In Kant's political philosophy his emphasis is likewise on the principles of politics that can be known a priori. Accordingly, his political philosophy consists primarily in a very general framework, which Kant presents as the necessary condition of what could count as a just political constitution. Kant does not paint a model constitution in great detail. His general framework leaves a lot of freedom for how particular groups can constitute themselves. Before spelling out this framework, I shall examine at some length the guiding idea behind it.

1. Kant's Question

Kant's political philosophy was shaped by both the social contract theories of Hobbes, Locke and Rousseau, and the natural law theories of Grotius and Pufendorf. These latter theories found expression in the writings of Achenwall, on whose work Kant lectured repeatedly, and they helped shape Kant's terminology as well as the structure of Kant's *Doctrine of Right* (cf. Byrd and Hruschka 2010: 15–19). Like both of these traditions, Kant presents social conflict as the basic problem of political philosophy and, like the natural law tradition, he sees law and obligation as the primary concepts for the solution, although he construes the source of obligation in a novel way (cf. Schneewind 1998: 518–22).

What then is the basic problem Kant seeks to address in his political philosophy? For articulating the initial problem Kant—like many thinkers before him—utilizes a state

of nature scenario (1797: 6:311f.). Kant is quick to say that he does not conceive of the state of nature as a historical fact (1793: 27:589; 1792: 8:297). Rather, it is a device to establish what the right political order should be (1797: 6:315; 1792: 8:302). In this scenario there is a multitude of human beings, living in proximity to one another, whose actions mutually affect one another. These human beings have unsocial tendencies. They have self-seeking inclinations, and "their inclinations make it impossible for them to exist side by side for long in a state of wild freedom" (1784: 8:22f.). It is not merely selfish inclinations but particularly the human striving for "honour, power, or property" that is the source of strife (Kant 1784: 8:21). In the state of nature there is no authority to enforce laws and keep the peace. The political problem is then "so to order this multitude … that in their public conduct the result is the same as if they had no such evil dispositions" (1795: 8:366).

The natural law tradition of Pufendorf saw God-given laws as the *solution* to the potential social conflict in the state of nature (cf. Schneewind 1998: 520). In contrast, Kant does not think that any external command, whether it be given by God, tradition or an external metaphysical order, can resolve the conflict. (I shall spell out the reason for this below.) For instance, Kant does not argue that in a state of nature one would immediately see that the other human being has value and is therefore deserving of respect: "the existence of man is not by itself a *factum* that produces any obligation" (1793: 27: 545). Accordingly, knowing that other human beings are free or have the faculty of reason is not the solution, but a reason for concern. One can anticipate how an animal will behave—its instincts are regular—but a free being is unpredictable and so more fearsome (1784a: 27: 1320). According to Kant, there is therefore no preexisting order one could rely on to establish order among human beings: "Here, then, we see philosophy put in fact in a precarious position, which is to be firm even though there is nothing in heaven or on earth from which it depends or on which it is based" (1785: 4:425).

Instead of looking up to heaven or down to earth, Kant finds the solution to the problem of strife within each human being. In his view human beings are not wholly unsocial—they are not entirely made of "warped wood"—but they have an "unsocial sociability" (1784: 8:23, 20). While the inclinations are a cause for strife, the conditions for a fair solution to conflict are already within the faculty of reason of each human being. Kant says that the condition of justice is known "a priori" (1792: 8:290; 1797: 6:355, 318), which means that the knowledge of this condition does not have to be gained by experience, but is necessarily possessed by all human beings (1787: B4). But what is this condition for justice?

Kant argues that the general condition for right actions and just solutions is universality: "Any action is *right* if it can coexist with everyone's freedom in accordance with a universal law" (1797: 6:230). This is Kant's "Universal Principle of Right" (*ibid.*). According to it, everyone is free to do what he or she wants, unless it interferes with the freedom of others. But this does not mean that any collision by itself makes one's action wrong, according to Kant. Rather it is a universal law that is the arbiter of justice.

Take a simple example. You are in a parking lot, and you find the last open space. In taking it you will restrict the freedom of the next person who has to wait. However, the collision of interests by itself does not make your taking the spot wrong (1784a: 27: 1320). There is a law, let us say, that parking spots can be taken on a first-come-first-served basis. Your freedom can coexist with the freedom of others in accordance with a universal law (in this case the first-come-first-served rule). The particular law

in this case could be something else (e.g., that parking spots are first given out to the person who is physically weaker). The Universal Principle of Right is a meta-principle. It requires that the more particular law *could* coherently be adopted by all in a specific domain. It does not by itself generate the particular law, but it would rule out proposals that fail to meet that requirement. The question is not so much whether others would actually or hypothetically consent to a proposed law, but whether the law could be held by all in the sense that a law to enslave others or severely injure others could not (cf. O'Neill 1996: 57–9). While only some people could put into practice the law to own slaves, one should act upon a universalizable law. The Universal Principle of Right is basically a requirement of fairness: A specific law should be valid for all in the relevant domain, and one should not make an exception for oneself (1785: 4:424; 1792: 8:297 note).

But the Universal Principle also lays the ground for justifying coercion. If one's action is right if it can coexist with the freedom of others in accordance with a universal rule, then whoever hinders one does one wrong. Coercion that restrains this wrong-doing is a "*hindering of a hindrance of freedom*" (1797: 6:231). Coercion supports one's rightful freedom, and is therefore right. Hence the authorization to coerce (for oneself in the state of nature, or for the police in a state) follows by the principle of non-contradiction (*ibid.*).

According to Kant, the Universal Principle of Right is already known in the state of nature. Accordingly, "in the state of nature, too, there can be societies compatible with rights," and a civil state "contains no further or other duties of human beings" that are not already known in the state of nature (1797: 6:306, 312). However, even if it can be known what is right (in contrast to what is by violence and force), there is no reason to expect that right will be upheld when it is not enforced. So even among "well disposed and law-abiding human beings" in the state of nature "it still lies a priori in the rational idea of such a condition" to leave the state of nature, "unless it wants to renounce any concepts of right" (1797: 6:312).

To secure right one should enter a civil state, that is, "a union of a multitude of human beings under laws of right" (1797: 6:313). This explains why Kant treats of his political philosophy in a treatise on legal philosophy, the *Doctrine of Right*. The solution to the initial problem of political philosophy, social strife, is in the first instance a civil state that can secure laws in accordance with the Universal Principle of Right. This condition for a rightful state is "politics cognizable a priori" (1795: 8:378).

In order to lay out more fully the a priori structure of Kant's political philosophy, one first must consider what the supreme principle of right entails, and what its status is.

2. The Universal Principle of Right and the Categorical Imperative

Kant's claim that the Universal Principle of Right has its source in human reason does not fully specify its basis, and the main debate in regard to Kant's political philosophy concerns the question of whether the Universal Principle of Right is the same as or different from the supreme ethical principle, Kant's famous Categorical Imperative: "*act only in accordance with that maxim through which you can at the same time will that it become a universal law*" (1785: 4:421). In particular, the debate is about whether the two principles share the same justification, or whether they are independent of each other. (For an excellent overview of the literature see Seel: 2009: 72f.)

What speaks in favor of the independence of the two principles? Systematically, the most important reason is the following: One can argue that the Universal Principle of

Right can be discovered by *prudential* reason, and that one can be motivated to follow it and enter a civil constitution simply out of self-interest (cf. Pogge 2002: 149; Geismann 2006). To leave the state of nature is "sheer necessity" (Kant 1784: 8:22). The state of nature is a state of war, and one "can never be secure against violence from another" (1797: 6:312). Life would be miserable, and Kant famously says that even a race of devils, i.e., a group of thoroughly selfish rational beings, could form a rightful constitution (1795: 8:366). So the Universal Principle of Right and the requirement to form a civil state might merely be a means to achieving peace. Kant's political conclusions would then amount to a hypothetical imperative ("enter a state *if* you want peace") in contrast to the Categorical Imperative, which commands unconditionally.

While this interpretation is plausible at first glance, it is harder to defend as an interpretation of Kant. For Kant says explicitly that prudence is *not* the basis of these requirements: "to unite into a state in accordance with freedom and equality as the sole concepts of right … is not based upon prudence but on duty" (1795: 8:378). Thus Kant clarifies that neither the Universal Principle of Right nor the requirement to enter a civil state is based on prudence, in his view. On the first point, Kant also says explicitly that the Universal Principle of Right is not a hypothetical but a "*categorical imperative*" (1797: 6:318, 252; 1793: 27:540). Kant argues that it is not just a law that is valid *if* one wants something else, but that it has "unconditional necessity" (1795: 8:377; cf. 1797: 6:354f.). And he specifies the political law in terms virtually indistinguishable from the Categorical Imperative: "So act that you can will that your maxim should become a universal law (whatever the end may be)" (1795: 8:377; 1797: 6:225f.).

But if the supreme principle of right is not based on self-interest, why does Kant say that "the problem of establishing a state … is *soluble* even for a nation of devils" (1795: 8:366)? The reason is that, when he makes this remark in "Toward Perpetual Peace," Kant is discussing the *motivation* to follow the Universal Principle of Right rather than how it can be *known* or its *justification* (similarly Seel 2009: 87; Oberer 2006). He addresses an objection that "human beings, with their self-seeking inclinations, would not be capable of such a sublime form of constitution" (1795: 8:366). Against this Kant argues that even selfish beings would be *motivated* to form a rightful state. (Below I shall address the issue of how the principle is known and justified.) In Kant's view politics or law—in contrast to ethics—does not demand that one do what is right from a moral motive. One can act on "self-seeking inclination" (*ibid.*) as long as one conforms to the principle of right. This is Kant's famous distinction between legality and morality (1788: 5:71f.; 1797: 6:214). In ethics "it is not enough that it *conform* with the moral law but it must also be done *for the sake of the law*" (1785 4:390). This motivational requirement to do the right thing simply because it is right is absent in legal and political philosophy (1797: 6:231). One should stop at a red traffic light, but one can do so to avoid punishment rather than because it is right.

While ethics and politics differ in their *motivational* requirement, their supreme laws share the same *content* (namely, universality), and—as I shall shortly argue—the same *justification*. There is therefore one "universal moral law" which "comprehend[s] both legal and ethical laws" (1793: 27:526, cf. 523f.).

In short, there is one supreme law for politics and ethics, which Kant calls the Moral Law. It requires universality, and it does so unconditionally. The Universal Principle of Right does not contain a requirement that one have a moral motive (in contrast to the Categorical Imperative, 1785: 4:431f.). But what exactly is the justification for the overarching Moral Law? Why is this law the supreme principle, and not another?

3. Kant's Justification of the Moral Law

Kant offers two lines of argument for the Moral Law. The first is that only his law could be a law expressing moral obligation; the second is that there really is such a law. Kant supports the first claim by arguing that no other principle could be valid for all and unconditionally valid. But why must the Moral Law be universal and unconditional? And why does the Moral Law alone fulfill this requirement? Kant claims that we commonly conceive morality to be universal and not conditioned by desires (1785: 4:389). But he also argues directly for the view that morality must be universal and necessary.

Universality: Why must there be one law that is valid for all, rather than many different laws for different peoples (or no laws at all)? This is the challenge of relativism. The way Kant sees the political problem (see above), a relativism without an overarching moral law would be a state of nature. If one group has a law to worship God A on Wednesday on threat of punishment, and another group has the law to worship God B on Thursdays or be punished, there is potential strife if these two groups come in contact with each other (1797: 6:350). The conflicting parties "need universal laws for their preservation" (1795: 8:366; 1784: 8:22). In some cases the solution might be a compromise between the two existing laws, in other cases it might be a law of toleration, but there needs to be a law covering both groups to avoid strife.

Necessity: The supreme Moral Law has to be unconditioned because otherwise it could not be universally binding. A law that would only be binding *if* something else holds, would not be obligatory for groups that do not accept the condition (1785: 4:389). If the supreme law were conditioned by a particular tradition or desire, those reasons "will seem at worst incomprehensible to those who do not grasp those norms and commitments and at best merely conditionally reasoned to those who grasp but do not share them; in either case they will seem arbitrary" (O'Neill 1996: 51). If it is to be binding for all, the supreme Moral Law cannot be conditioned upon contingent factors.

But why is only Kant's Moral Law universal and unconditioned? It is one thing to argue that there needs to be a law that is valid for all (against relativism), and another to say that the *content* of that supreme law must be universality (in contrast to a particular content like happiness). Kant's point is that any particular content other than universality would make the supreme law conditioned and thereby not universally valid. For instance, if one bases a law on a concrete desire (e.g., for happiness or peace), the law would only be valid *if* one has this desire. However, desires are relative and contingent, in Kant's view (1788: 5:26), and cannot ground a necessary and universal law. People's desires and conceptions of happiness differ, and there is therefore not one prescription for what to do to fulfill desires or bring about happiness (1792: 8:290; 1795: 8:377). But even if all people had the same overriding desire (such as peace), this would not solve the problem. For if two men want peace, one might attack the other to attain peace for himself (1797: 6:307; 1788: 5:28).

The supreme law would also be contingent if one were to base it on an external ground, such as commands from God, society, tradition, education or an order of value. Whenever a law is based on tradition, for instance, it is only binding for oneself *if one desires to* follow tradition (or desires to avoid punishment). The external ground would reduce itself to one's own desire. This led Kant to claim that "all previous" and more importantly "all possible" attempts (other than his own) to ground the supreme Moral Law failed (1788: 5:39f.; 1785: 4:441–4). Kant's point is that the "categorical imperative … as such only affirms what obligation is" (1797: 6:225), its nature so to speak, in

contrast to what it means to desire something. A maxim based on a desire by itself does not generate an obligation or moral ought: "the only way this maxim can be binding is through its qualification as a universal law" (1797a: 6:393; 1787: B575f.). Any particular content would make a law relative and contingent. Only Kant's Moral Law—in excluding all content—spells out what can be binding independently of desires. Or to put it in terms of reasons for action, the Moral Law spells out the necessary condition for what can count as a reason for all (cf. O'Neill 1996: 51; Hill 1992: 111f.).

This concludes the first line of defense, but it leaves open a problem I mentioned earlier. While Kant has made an argument as to what the supreme Moral Law must be like, it could be that there is no such unconditional law. Maybe Kant's Moral Law is only valid *if* one wants to avoid strife and violence. There is another argument needed to show that this law is also unconditional, valid independently of *any* desire. Accordingly, Kant offers a second line of defense in which he argues that there really is such a law. To recall: Kant's view is that the Moral Law "has unconditional necessity" (1795: 8:377), and rests a priori on reason (1792: 8:290). The law is an in-built principle of reason so to speak. It is an operating principle of how reason works, similar to the principle of non-contradiction, so that even if one has no desire to follow the Moral Law, it would still make itself known—like a bad conscience.

Kant's argument for the view that everyone knows the Moral Law is that it is given with one's capacity of free will. He conceives of free will as a form of causality, as starting a chain of events independently of natural determinants; and he argues that every kind of causality needs a law. If one abstracts from all of the matter that is provided by the natural desires, only the *form* of the law remains; this form is universality as expressed in the Moral Law (1785: 4:446f.; 1788: 5:29f.). Kant claims that even an eight-year-old child thinks that moral rules are unconditional, and judges morality in terms of fairness and universality (1792: 8:286; 1788: 5:155f.). These claims find support, incidentally, in recent empirical research (Nichols 2005, 2009).

4. The Emptiness Objection

Kant's Moral Law is basically an abstract requirement of fairness. But is this enough to yield a whole political philosophy? The most famous objection that has been brought forward against Kant's Moral Law is that it is empty and devoid of content (cf. Hegel 1820: §135). Is the universality requirement not too thin to yield any conclusions about what is right in politics?

Kant would admit that the formal law is empty (1793: 27:578), but he would not consider this an objection to the Moral Law. The Moral Law is not meant to generate specific laws on its own (Oberer 2006: 261f.; Seel 2009: 75). In Kant's view, the content or matter of specific laws is freely chosen based on inclinations or prudential considerations (1785: 4:421–4). Everyone is free to do whatever he or she wants—unless one's actions interfere with the freedom of others. Kant leaves individuals the freedom to arrange themselves according to their preferences. For instance, one society might think that the family should care for its elderly, while another society might prefer that the state assists with this. Kant wants to preserve this freedom (1792: 8:290f.); he is interested in the lowest common denominator or the necessary condition for a rightful society (1792: 8:289), i.e., what reason can say for all human beings prior to experience (1797: 6:216).

There is, however, more than just the Universal Principle of Right that Kant can say a priori about politics, and this comprises the structure of Kant's political philosophy.

5. The Structure of Kant's Political Philosophy

So far I have focused on the initial question that guides Kant's political philosophy, and its supreme principle. I have described the spirit of Kant's political philosophy, but not yet its specific structure—although he does not ever elaborate this structure in much detail. The structure is built under the norm of the supreme principle, but is on occasion supplemented by further arguments.

The structure of Kant's political philosophy can be described in four stages, from the initial state of nature to a rightful international order between nation states (cf. Byrd and Hruschka 2010: 51). First, even in the state of nature a human being would know the Universal Principle of Right. Kant argues that this entails further ideas. I shall introduce these ideas with reference to Kant's notions of autonomy, respect and dignity—notions that are popular in contemporary thought. Second, additional considerations about the possible interaction among human beings in the state of nature lead to a set of further rights. In the *Doctrine of Right* Kant argues for them under the title of "private right" (1797: 6:273ff.). Third, under the heading of "public right" (1797: 6:309ff.) Kant describes the conditions of forming a state that will secure what the supreme principle of right and what private rights require. Fourth, the security of one's rights in a nation state is incomplete without a rightful international order, according to Kant. I shall discuss each of the four stages in turn in briefly outlining the structure of Kant's political philosophy.

5.1 First Stage: Autonomy, Respect, Dignity

The Universal Principle of Right yields one innate right: One has a right not to be constrained by another's will, but to freedom "insofar as it can coexist with the freedom of every other in accordance with a universal law" (1797: 6:237). This right is justified with reference to the Moral Law, since "we know our own freedom (from which all moral laws, and so all rights … proceed) only through the *moral imperative*" (1797: 6:239). This right has three further aspects which Kant specifies in the political realm as freedom, equality and independence (1792: 8:290; 1795: 8:349; 1797: 6:237f., 314), and in the ethical realm as autonomy, respect and dignity. I shall briefly sketch the three latter notions, and connect each to its counterpart in the political realm.

Autonomy: Kant's notion of autonomy has many facets. Kant uses the term for the requirement that the Moral Law not be based on external or contingent factors, but must be given by one's own reason (1785: 4:441–4). Related to this is the capacity of a human being to be "free with respect to all laws of nature, obeying only those which he himself gives and in accordance with which his maxims can belong to a giving of universal law" (1785: 4:435). This form of autonomy has its counterpart in politics as one requirement for the establishment of a state. The first attribute of a citizen is "lawful *freedom*, the attribute of obeying no other law than that to which he [could have] given his consent" (1797: 6:314; 1792: 8:290).

Respect: A second notion that is already contained in the Moral Law is ethical respect for all human beings or the equality of all citizens in politics. The Moral Law is the "principle of equality" (1797a: 6:451). Kant explains the connection thus: In testing whether a proposed law could be universal, one rejects laws that another could not adopt. But in doing this one treats him as an equal (1785: 4:437f.; 1788: 5:87). Accordingly, ethical respect is the maxim "of not exalting oneself above others" (1797a: 6:449). In politics

this corresponds to equality as the second attribute of a citizen, for instance, that "each is entitled to one vote" (1792: 8:295f.), or that there are no hereditary privileges, but that everyone "must be allowed to attain any level of rank" (1792: 8:292; 1797: 6:314).

Dignity: A third concept that is entailed in the Moral Law is the dignity of human beings in ethics, and the dignity of citizens in politics. Kant conceives of dignity as an elevated status or "prerogative" (1785: 4:438). In ethics Kant argues that a human being has an elevated status or "dignity; for a human being cannot be used merely as a means by any human being" (1797: 6:462, 434f.), according to the principle of equality. In the context of political philosophy Kant emphasizes the independence or dignity of citizens (1797: 6:329f., 314; 1792: 8:290). As a citizen a human being "*serves* no one" (1792: 8:295); instead he is his own master and has an equal rank as colegislator (1792: 8:294–6; 1797: 6:314). That is, citizens "must always be regarded as colegislating members of a state (not merely as a means …)" (1797: 6:345).

In short, freedom, equality and independence are directly connected with the Moral Law. They are negative or protective rights in the state of nature, specifying what others may not rightfully do to a human being.

5.2 Second Stage: Private Right

In the state of nature the Universal Principle of Right entails the freedom and equality of each human being. But there are also further rights concerning the interaction of people: "For in the state of nature, too, there can be societies compatible with rights" (1797: 6:306). Following a traditional distinction made in Western legal systems, Kant discusses property rights, contract and status rights. These correspond to the different sorts of relations people can have to one another (cf. Ripstein 2009: 20): human beings as independent of each other (property rights), as consensually interacting (contract rights), and as non-consensually interacting, as in the case of parents and children (status rights). The emphasis in the Kant literature is on property rights, and I shall briefly sketch Kant's main points concerning them.

Kant's first concern in discussing property is to explain how rightful possession of property is possible. Kant's argument for the possibility is that the opposite principle—that rightful possession is not possible—would involve a contradiction. The idea, in short, is that while a free person necessarily wills the means for realizing his ends, "putting *usable* objects beyond any possibility of being *used*" would be a "contradiction of outer freedom with itself" (1797: 6:250; cf. Byrd and Hruschka 2010: 115). Kant sees this "permissive law" to own property as something with which "reason extends itself a priori" (1797: 6:247) beyond the Universal Principle of Right. But the Universal Principle is still important for property rights, since in taking possession of an object one also declares "that everyone else is under obligation to refrain from using that object" (1797: 6:255). Again, according to Kant obligation is only generated if one reciprocally grants possession to others, since "obligation … arises only from a universal law" (*ibid.*; cf. 1797a: 6:393). To have the assurance that others will respect one's property, however, one needs an authority to back it up. To secure one's rights, one therefore must leave the state of nature (1797: 6:256).

5.3 Third Stage: Public Right

In the state of nature there is no assurance that rights will be protected and laws will be enforced. There is no *executive* power—no police, so to speak (1797: 6:256). In addi-

tion there is no judge or *judicial* authority to administer justice if there is a dispute about contracts etc. (*ibid*. 6:312). Finally, there is "no law" (*ibid*. 6:306), no laws promulgated for all by a *legislative* authority. So unless one wants to renounce every concept of right, one morally should leave the state of nature, and form a commonwealth or state (1792: 8:310; 1797: 6:312). Kant calls the system of laws in a rightful state "public right" (1797: 6:311).

Kant discusses a rightful nation state in terms of three authorities: legislative, executive and judicial (1797: 6:313–5; Byrd and Hruschka 2010: 143–61). The legislative authority, Kant argues, must represent the united will of the people. This is because it should lay down what is right, and therefore cannot wrong anyone. And while one cannot wrong oneself in freely choosing a law for oneself, one can wrong someone else in choosing for him or her: "Therefore only the concurring and united will of all … can be legislative" (1797: 6:313f.). In an ideal state all citizens would agree upon the laws; the closest an actual state can come to approximating this idea is to form a representational system in which delegates decide in a parliament (1797: 6:341; 319). The representation is needed because otherwise a majority could outvote a minority, and this would be a "*despotism*" in contradiction with the general (united) will of all (1795: 8:352). The legislative does not rule by actual or hypothetical consent of the people, but by what it deems indispensable for a rightful state, and a law is just if consent is "*only possible*" (1792: 8:297). The principle of the legislative is therefore the Universal Principle of Right, while the executive makes conformity with this law obligatory, and the judicial branch determines what is right in accordance with the Universal Principle (1797: 6:313).

Kant emphasizes that the three authorities should be separate, that is, that the power should rest in different hands (1797: 6:316f.). Kant calls this ideal form of government a "republic" (1797: 6:340; 1795: 8:352). If the ruler (executive authority) were the same as the sovereign (legislative authority), the danger is that the government would become despotic, paternalistically overruling the autonomy of the people (1797: 316f.; 1792: 8:291). Instead the ruler should be under the law of the sovereign, the united will of the people. Here too, Kant confines himself to these very general lines, and leaves room for different implementations of a republican government.

In case of despotism, Kant famously denies the right to a revolution, but urges reform to bring the state in conformity with the Universal Principle of Right (1797: 6:321f.). This is a much-discussed topic in the Kant literature. It is not clear whether Kant adopts this position to avoid censure from the Prussian monarchy, or to prevent the horrors that followed the French Revolution. Some Kantians think that he is inconsistent in ruling out a revolution guided by the supreme law (Oberer 2006: 266). Others defend Kant on the basis that his view does not prevent resistance to the horrible regimes of the twentieth century (cf. Ripstein 2009: 325–52; Byrd and Hruschka 2010: 181–7).

5.4 Fourth Stage: International Right

The initial problem of strife between human beings is not fully resolved with the creation of a state under the Universal Principle of Right. The problem is that different states remain in a state of nature in relation to one another. The rights of the individual human being are endangered by a constant threat of war (1797: 6:344; 1784: 8:24–6). Kant sees the solution as being parallel to the solution for the initial state of nature (1797: 6:209; 1792: 8:312), but he does not favor a single world government.

The danger Kant sees is again despotism, in which cultural differences and disputes are suppressed by a single authority (1792: 8:310; Byrd and Hruschka 2010: 197f.). Kant's ideal solution would be a "state of nations" (1797: 6:350) in which states submit to public coercive laws (1795: 8:357), with an international court to decide disputes and a coercive executive power to enforce the decision (cf. Byrd and Hruschka 2010: 201–3). However, Kant considers the actual protection of each state over a vast territory "unrealizable" (1797: 6:350), and puts forth his solution only as something states should strive to approximate. Short of a state of nations, he favors a "league of nations" (1795: 8:356; 1797: 6:344) or "congress of states" (1797: 6:350) in which states aim to settle disputes peacefully without subordinating themselves to public coercive laws. Even if it is very difficult to achieve, the highest political good as set by reason is perpetual peace, understood as the just "condition in which what is mine and what is yours for a multitude of human beings is secured under *laws*" (1797: 6:355).

6. Conclusion

Kant's political philosophy is one of republican government, of ordered freedom under laws, and of the equal dignity of all citizens. Kant keeps his focus on the a priori framework of politics. While his practical philosophy has been criticized for its lack of determinate content, this is, paradoxically, exactly what should make his theory attractive in the modern world. Different traditions meet, the arrangements are not fixed. In Kant's political philosophy, the future is open, but reason tells us to shun violence, and strive for a fair and lawful peace.

Related Topics

Hobbes, Locke, Rousseau, Liberalism, Republicanism, Natural Law and Rights Theory, Human Rights and Cosmopolitanism, Equality, Freedom, Autonomy, Rights

References

Byrd, S. and Hruschka, J. (2010) *Kant's* Doctrine of Right. *A Commentary*, Cambridge: Cambridge University Press.

Geismann, G. (2006) "Recht und Moral in der Philosophie Kants," *Annual Review of Law and Ethics* 14: 3–124.

Hegel, G.W.F. (1820) *Elements of the Philosophy of Right*, trans. H.B. Nisbet, Cambridge: Cambridge University Press, 1991.

Hill, T. (1992) *Dignity and Practical Reason in Kant's Moral Theory*, Ithaca, NY: Cornell University Press.

Kant, I. (1784) "Idea for a Universal History with a Cosmopolitan Purpose," in I. Kant (1970).

—— (1784a) "Kants Naturrecht Feyerabend," in *Kant's Gesammelte Schriften*, vol. 27.2,2, Berlin: de Gruyter, 1979.

—— (1785) *Groundwork of the Metaphysics of Morals*, in I. Kant (1996).

—— (1787) *Critique of Pure Reason*, trans. P. Guyer and A. Wood, Cambridge: Cambridge University Press, 1998.

—— (1788) *Critique of Practical Reason*, in I. Kant (1996).

—— (1792) "On the Common Saying: That May be Correct in Theory, but is of No Use in Practice," in I. Kant (1996).

—— (1793) "Lectures on Ethics Vigilantius," in I. Kant (1997).

—— (1795) "Toward Perpetual Peace," in I. Kant (1996).

—— (1797) *Doctrine of Right*, in I. Kant (1996).

—— (1797a) *Doctrine of Virtue*, in I. Kant (1996).

—— (1970) *Political Writings*, trans. H.B. Nisbet, Cambridge: Cambridge University Press.
—— (1996) *Practical Philosophy*, trans. M. Gregor, Cambridge: Cambridge University Press.
—— (1997) *Lectures on Ethics*, trans. P. Heath, Cambridge: Cambridge University Press.
Nichols, S. (2005) "Innateness and Moral Psychology," in P. Carruthers *et al.* (eds.), *The Innate Mind*, Oxford: Oxford University Press, 353–69.
—— (2009) "Emotions, Norms, and the Genealogy of Fairness," *PPE* 9: 1–22.
Nozick, R. (1974) *Anarchy, State, and Utopia*, Oxford: Blackwell.
Oberer, H. (2006) "Sittlichkeit, Ethik und Recht bei Kant," *Annual Review of Law and Ethics* 14: 259–67.
O'Neill, O. (1996) *Towards Justice and Virtue*, Cambridge: Cambridge University Press.
Pogge, T. (2002) "Is Kant's Rechtslehre a 'Comprehensive Liberalism?'" in M. Timmons (ed.), *Kant's Metaphysics of Morals*, Oxford: Oxford University Press, 133–58.
Rawls, J. (1971) *A Theory of Justice*, Cambridge, MA: Harvard University Press [2003].
Ripstein, A. (2009) *Force and Freedom*, Cambridge, MA: Harvard University Press.
Schneewind, J.B. (1998) *The Invention of Autonomy*, Cambridge: Cambridge University Press.
Seel, G. (2009) "How Does Kant Justify the Universal Objective Validity of the Law of Right?" *International Journal of Philosophical Studies* 17: 71–94.

Further Reading

For Kant's main work, the *Doctrine of Right*, Byrd and Hruschka (2010) is very helpful as a historical commentary, while Ripstein (2009) unlocks the work systematically. Two collections of papers on the *Doctrine of Right* by leading scholars are M. Timmons (ed.), *Kant's* Metaphysics of Morals (Oxford: Oxford University Press, 2002) and L. Denis (ed.), *Kant's* "Metaphysics of Morals" (Cambridge: Cambridge University Press, 2010). H. Williams, *Kant's Political Philosophy* (Oxford: Blackwell, 1983) is a clear exposition of Kant's political philosophy more generally, and O'Neill (1996) a thoughtful defense of it. K. Flikschuh, *Kant and Modern Political Philosophy* (Cambridge: Cambridge University Press, 2000) and O. O'Neill, *The Bounds of Justice* (Cambridge: Cambridge University Press, 2002) enlighten the relationship between Kant and contemporary liberalism. T. Hill, "The Kingdom of Ends," in his (1992) brings out an important parallel between Rawls and Kant. P. Kleingeld, *Kant and Cosmopolitanism* (Cambridge: Cambridge University Press, 2011) discusses in depth Kant's views on international right and cosmopolitanism.

11
HEGEL

David Edward Rose

As is the case with most great philosophers, critics disagree on the worth of Georg Wilhelm Friedrich Hegel's (1770–1831) legacy, but he remains—either in spite of or due to this fact—one of the giants of the intellectual canon. The context of Hegel's political thought is perhaps best understood as the coincidence of the historical aftermath of the French Revolution and the theoretical aftermath of Kant's moral philosophy. Both, for him, were symptomatic of a deeper historical failure: that of the Enlightenment's investment in individual reasoning free from all authorities or external dictates. And it was the uncertain wavering of the once ambitious goals and projects of his immediate predecessors that determined Hegel's agenda and also underlies our own contemporary unease in interpreting the themes of his thought. He can simultaneously seem conservative and radical, an obscure metaphysician and, at the same time, a pragmatist, a liberal rationalist and, yet, a reactionary. His own writings are notoriously difficult to follow and the development of his thought travels from a youthful Romanticism to a more Kantian conservatism in the 1830s. It is the last phase of his thought which is perhaps most accessible and will form the basis for the interpretation offered in the following pages.

1. Political and Social Philosophy is a Philosophy of the Will

If one were forced to offer a general and formal characterization of the Enlightenment, it would not be too controversial to conceive of it as an effort to emancipate human beings from the dual tyranny of superstition and arbitrary power. Modernity in general can be characterized as the transition from authoritative sources of knowledge and moral dictates (the priest, the ruler, the philosopher) to legitimation through individual warrant (the individual him or herself must recognize the validity of the reasons offered independently of the source of who offers them). And the Enlightenment's grounding value of equality implied the claim that *all* human beings had access to these reasons, whether it be through some immediate feeling of right or wrong, a special moral faculty or through some prior principle of reason, such as Kant's categorical imperative. For Hegel, the Enlightenment's understanding of freedom is an historical advance, yet simultaneously undermines its own achievements; his political and social writings are an attempt to rectify this theoretical tension.

The ruminations on the centrality of free-will persist throughout Hegel's writings. The mature lecture courses on ethical and political philosophy in the 1820s and 1830s begin with the keystone to his whole theory of political philosophy: "The subject-matter of *the philosophical science of right* is the *Idea of right*—the concept of right and its actualization." (Hegel 1991: §1) which is then extrapolated to this claim:

> The basis of right is the *realm of spirit* in general and its precise location and point of departure is the *will*; the will is *free*, so that freedom constitutes its substance and destiny and the system of right is the realm of actualized freedom, the world of spirit produced from within itself as a second nature.
>
> (Hegel 1991: §4)

These two brief statements at the very beginning of his lectures embody much of the substance of Hegel's social philosophy. We are introduced to specific terms of art (actualization and Spirit or mind (*Geist*)), particularly Hegelian understandings of traditional philosophical concepts (right, will and freedom) and also given the connection between these and political science: political science (the realm of right) is the objectification of the freedom of the practical or acting will in social and economic institutions and also in cultural mores and attitudes such that the individual can be fully and rationally free. The relationship between the objectification of will and the famous (or notorious) *Doppelsatz* (double-saying) in the Preface to the *Elements of the Philosophy of Right*: "What is rational is actual/And what is actual is rational" rests upon the philosophical basis of Hegel's writings: absolute idealism.

A way, in so brief a space, to comprehend Hegel's absolute idealism is perhaps through an example that is at least implicit in his own thinking about the modern world: equality. Idealism holds that reality is constituted by the active work of reason in that the subject applies categories and concepts to what is encountered in order to make it meaningful. To offer an explanation of anything, I must apply the bare concept of "cause" to the relationships of events, but the concept belongs to my mind enabling me to produce knowledge statements. Where Hegel's absolute idealism differs from this simple form of idealism is that the activity of reason translates itself into external labor. So, if I believe that the concept of equality ought to be applied to the world, I may find that it is impossible to impose on reality (that what is rational is not yet actual). By imposing the idea of equality on the world, I may destroy certain economic and social institutions (such as slavery) that make human existence—at such and such a time in history—possible. However, the thinking mind has made a demand on the world and the world will be changed through the labor of communities and individuals: what must be the case for equality to be made possible is the abolition of slavery through the creation of alternative economic and social models and sometimes social worlds will be destroyed by the demands of reason. History becomes the playground of a reason imposing itself on the world to make its rationality actual. So, what does equality demand? It demands social institutions that respect persons as individuals and establish moral equality, as well as a system of justice based on rights. Only when these conditions are met (after a long historical process) will what is rational be actual and Hegel believes, in the second part of his statement, that in modern Europe—more or less—what is actual is in fact rational. The will of men is imposed on the world as "objective Spirit" or the political and moral institutions which make the social world rational and free.

The division of Hegel's lecture series robustly reflects the structure of his political thought and the shapes of "object spirit" in his contemporary world. The lectures begin with the concept of "right" intuitively understood as the expression of a will acting freely and rationally. The concept develops into an understanding of the "person" or an atomistic, desiring subject with his or her abstract rights and concomitant duties when interacting with other atomistic selves before becoming a consideration of the moral subject of modern Europe and, finally, concluding with the demonstration of the codes,

mores and norms of free behavior. For Hegel, the modern subject's self-understanding—which is made possible and also maintained by the existence of specific political and social institutions—guarantees the actual (that is rational and existent) freedom of that subject. The modern self understands itself as an atomistic, individual person, as a responsible, moral subject, as a family member, as a worker and as a citizen and these self-understandings harmonize to form the free, rational identity of the self.

2. Freedom and the Practical Will

Traditionally, philosophy has characterized free-will in two ways. One, as the freedom to satisfy desires: I am not free when I am inhibited from acting as I would have done had the other agent not been present. So, I am free when I can satisfy my desires with no external intervention and not free when I am unable to satisfy my desires. Two, as freedom of choice: I am free when I can transcend and negate any external determination. So, I am free when I can choose which amongst the set of my pressing desires I wish to act upon and am not free when suffering from addiction, psychological manipulation or when I am coerced. In the first case, desires and motivations are fixed and "givens" of the agent, in the second case motivations arise from either an act of will or through a deliberative process on those desires which constitute me as a person. Without delving too deeply into the intricacies of Hegel's discussion of the traditional understanding of free-will, it is enough to report that, for him, both accounts are "one-sided," which is to say that, although they do capture something of the truth of free-will, they also omit something equally necessary to the understanding. And to ground one's ethics or politics in a one-sided account will lead to damaging consequences.

To perceive the human being solely as a desiring-thing is to negate the very real instances when the agent sacrifices his or her strongest desire for something that, though not as keenly wanted, is valued over and above wants and preferences (acts of martyrdom, altruistic actions and so on) and also to negate the role of the will in a decision-making process. However, to prioritize the freedom "to act otherwise," that is the practical will's capacity to resist external determination, leads to the arbitrariness of an agent's rejecting authority for the sake of rejecting authority, to feel freedom in the power to say no. It is akin to a capricious teenager who begins to feel her independence and does not agree with her parents, not because of what they tell her to do, but because they are parents and agreeing with them, no matter how rational they are, is an admission of minimal dependence. When it is the individual and his or her own use of reason which is the ultimate arbiter of right and wrong, then "Whatever the will has decided to choose, it can likewise relinquish" (Hegel 1991: §16). And the arbitrary and unrestricted nature of human freedom is both the possibility of and also a danger to moral and political action. The immoral excesses of the Terror of the French Revolution and what Hegel calls the subjective moralities of his own time (culminating in Kantian ethics as he understood it) were symptomatic of the Enlightenment's one-sided understanding of the practical will and freedom.

For Hegel, the human being is a sensuous creature that lives in an actual world, and so must act with a consideration of desire-fulfillment and within the limits of physical and social constraints, but also a willing-thing that chooses to act upon certain motivations rather than others when these motivations matter more than simple satisfaction. Free-will is established by the triple axes of the agent's substantial, natural identity, the agent's social identity and also the projects, aspirations, values and preferences of the

particular agent and the "what I want" of the desiring-thing and the "what I think is right" of the reasoning mind are to be reconciled in a substantial, social identity that *actually* wills and wants what is *rational*.

Integral to such an understanding of free-will is, obviously, one of the most famous of the Hegelian concepts: recognition. In a theoretically bare sense, recognition is granted by the struggle to death between the would-be master and his would-be slave (Hegel 1977: §§178–96). Such a presentation of the role of recognition is extreme and it need not be such an all or nothing relationship between individuals, but what holds in that situation, holds for all relationships between individuals. As a human being, I want to be certain of my freedom. Therefore, I act upon motivations that I think are worthwhile and matter, but I cannot be sure that the actions I do are, in fact, objectively ones that matter and not mere arbitrary preferences of my will. Again, if we consider our teenager, she might want to start refusing to wear the clothes her mother chooses for her but is unsure whether her own taste is good or not. She makes an appeal to her peers (her classmates) to recognize her choice as worthwhile. The agent cannot rely on the will itself to inform him (or her), nor are there deontic values out-there to which he merely needs to conform. Rather, the certainty of freedom can only come from an other who recognizes him as a free agent through the value of the ends of his action and he himself *recognizing* that other as one whose judgment he trusts because the other is also a free being.

There is, of course, one action that proves my freedom beyond all others: the act of sacrificing trivial ends for a principle. In order to be recognized as free, I am prepared to sacrifice my life for the sake of freedom. No other animal, according to Hegel, can do such an act: to sacrifice its existence for the principle of recognition. However, I cannot do this before witnesses incapable of comprehending my risk since I will seem irrational, dangerous or diseased. Rather, I must perform the action before an other that recognizes it as a rational action (in the sense that it is acting on a deliberative outcome of the will). I can be certain that the other person is capable of judging me only if he or she is rational and also prepared to risk his or her life for a principle. Hence, Hegel offers his dialectic of recognition that is not immediately resolved, but for the present purposes underlines the need for me and the other to recognize one another as acting on principles rather than immediate drives.

The dialectic of recognition reappears in Hegel's account of free human action. His account is consistent with standard theories in that it describes action as a normative self-ascription: one of the causal conditions that bring about a specific event must be an intention and the intention must be recognized by the acting agent as his or her own. Yet, Hegel adds a further requirement for free, moral action: "it has an essential relation to the will of others" (Hegel 1991: §113). The intention has to be capable of reconstruction by others from the objectivity of the act itself because it is necessary that others recognize the action as one's own. A shared moral and social fabric is necessary so that one can be certain of being understood as *doing* rather than just *being* an event: I know that you have dived from the cliff if I see you pause, extend your hands over your head and spring from the rock. Such a judgment rests, though, on a shared pre-understanding of the possible actions and transgressions one can perform in such a position and how an agent can communicate his or her conformity to, or transgression of, these expectations (in short, his or her intention). Broadly understood, Hegel's postulation of "right" includes all the values, expectations, moral laws and laws of etiquette into which an agent has been thrown through pure social luck. And these structures of meanings and

the social and economic institutions that support them are objectified acts of my historical predecessors that allow me to be free. So, for example, the choice of the construction of an economic system of capitalism allows the inhabitants of such a system—through mass, cheap production—the subjective freedom to choose their vocation, their way of life and to consume those things that express their choices. Previous acts, deliberations and choices are frozen into social institutions that liberate the individual from his most pressing and immediate drives, from having to create the categories of rights and valuable action for himself and also free him from being caught in indecision of how to best express what he wants. His role in the family, the workplace, the economic market and the political and social framework of his culture are an objective freedom which makes his subjective freedom possible.

And so full, human freedom, for Hegel, has two interrelated aspects: subjectively, the will is free if the agent identifies an intention as his or her own and, objectively, the will's freedom is dependent on the rational order shared with others that allows the agent to articulate, order and prioritize his or her intentions as an expression of his or her identity. Free moral action, that is activity proper of the practical will, is only possible when a subject feels "at home" in his or her culture: the institutions and organization of the social realm are the product of past actions of individuals and communities and hence practical will that has become objectified. The fact that we in the United Kingdom drive on the left-hand side of the road is the result of wills pursuing goals in a world with other wills and recognizing the need for a simple convention to make the achievement of such goals easier, hence increasing the agent's liberty. A decision, agreement or an understanding by agents in the past becomes an institution that supports and maintains my individual freedom. In such institutions, I can feel at home. Certain institutions may, however, inhibit an agent's experience of homeliness: the institution of marriage, for example, might be structured so that the individual is prohibited from leading certain worthwhile lives that they might well otherwise have pursued. As such, critical reason can demand that the institution be changed to accommodate the will of the individual (reason is not yet actual). The substratum of meanings, norms, expectations and duties that guide and regulate behavior of individuals guarantees recognition between subjects. Having only just met you, your clothes, roles, relationships and vocation—even the ringtone on your cell phone—will supply a host of information about how I expect you to behave, the transgressions which are permissible (idiosyncrasy) and the transgressions that are forbidden (wrongness). I recognize you as an agent (and not an object, an animal or child) because I perceive from your actions that you are able to act rationally over and above the natural drives and desires of your being. In order to be recognized as a rational agent, one's intentions must address those expectations, values and motivations which an other would recognize as intelligible, appropriate, right or good as determined by my social identity (father, worker and so on).

In a society in which the rational values are contrary to or inhibiting of the agent's personal freedom (one in which individuals of a specific gender are not allowed to pursue an education and such an institutional prohibition is "backed up" by cultural attitudes about that gender and its potential and desirable roles, say), then the agent is not free. Similarly, an agent who rejects (as in the Terror or in asceticism) all claims and understandings of his or her fellow agents for the sake of some commitment to authenticity of individual self, then the agent is not rational nor free. The aim of social philosophy is not to distinguish rational agents from irrational ones (as many of the thinkers of the Enlightenment supposed), but to distinguish rational societies or cultures from irra-

tional ones. Practical will, the capacity to freely act on rational intentions, is dependent on the agent's participation in a specific culture of meanings, expectations and norms. The structures and axiologies of societies are the yardstick of the rightness and worth of possible actions and projects and, as such, supply the duties for the practical will to conform to and the grounding rationality of all personal motivations. The issue of freedom for Hegel is a political issue because the agent's social luck, where and when he is born, determines the extent of his freedom.

3. *Sittlichkeit*: Cultural, Social Fabric

The most innovative and unique aspect of Hegel's ethical thought consists in his postulation of an immanent doctrine of duty or *Sittlichkeit*. Implicit to the idea of *Sittlichkeit* is the belief that the drive to separate morality from sociology and from political science, to bracket off the individual's true identity from his or her social identity is erroneous and leads to normatively undesirable consequences. The agent is social through and through and what he ought to do is obey the imperatives and duties of the many varied roles he occupies and the social relationships in which he participates. But given Hegel's mistrust of unrestricted reason, the agent's social identity also supplies the rational standards for the evaluation of his or her society. And it is this commitment that grounds the conservatism inherent in Hegel's social philosophy: an agent cannot apply transcendental or universal standards of rightness (such as a doctrine of natural rights or universal reason) to the institutions and structures of the state because he is unsure whether his reason is arbitrary. The only critical resources for the evaluation of the state are the very concepts and meanings of one's own social identity. So, political philosophy should not be concerned with the theoretical conjectures of utopianism, nor the contractarian descriptions of possible political blueprints, but with the very real critical task of engaging with one's own social and political frameworks from within. One must always start from "what is" rather than imagine "what ought to be" and the task of philosophy is primarily a descriptive attempt to reveal the structures of rationality inherent in the historical emergence of the state since, as Hegel tells us, "the Owl of Minerva begins its flight only at dusk" (Hegel 1991: 23).

So, Hegel's political and social thought is concerned with the actual state of affairs in his Prussian surroundings and revealing that, as it stands in its most basic form, the state in which he lives is the most rational that history has ever produced. Hegel is claiming, in a formal way, that the institutions of modern, liberal culture, the capitalist economy and the subjective self-understandings these make possible are by far the most rational social existence for human beings. In the modern political state, the way in which we understand ourselves, the shapes of freedom as Hegel calls them, are as an individual with rights and liberties (abstract right), as a moral agent with a conception of a worthwhile life and a set of values approximating to the good which demand respect from others (Morality), and finally as a member of a society with a concrete and determinate identity as a family member, worker and as a citizen of a nation state (*Sittlichkeit*).

In the modern state, individuals first understand themselves as "persons" or as individual right-bearers who exist independently from, and have separate desires from, the tribe, clan and family. The desires, preferences and projects of the individual are, in a most basic sense, his or her own. However, Hegel should not be mistaken for a methodological individualist. He does not, as many liberals do, believe that political explanations should be reduced to the atomistic level of individuals. For him, the "person" is

an historical product of the development of Spirit: it is neither an independent, transcendental entity as liberals hold, nor an imaginary, methodological misrepresentation as communitarians hold. It is a representation of a real, modern self-understanding: I choose which foods I like, which clothes I should wear and how I pass my time and no other agent is allowed to inhibit me unless my action involves a violation of his or her rights. But such a self-understanding is inadequate because in cases of crime we want to distinguish between accidental violation of rights and intentional violation of rights. The person who kills his wife by cutting the brake cables on the car and the mechanic who is negligent and does not adequately check the brake cables both harm another person, but the status of their crimes is different. A difference which is gauged by the reconstruction of the intentions and projects of the agent and so we also understand ourselves as "subjects" or moral agents. As an individual, I demand respect for the choices I make and we distinguish between violations of rights that are trivial and accidental and those that are deliberate. We see ourselves and others as intentional agents who pursue goods and projects that are valuable. In order to understand the agent's intention, we need to be able to evaluate his identity, what he values and what he holds dear. So, we must also understand ourselves as a substantial identity and not just a formal rights-bearer or a universal, moral subject. Substantial identity is supplied through the objective institutions of the rational state which allow me to be rationally and freely myself:

> Ethical life is the *Idea of freedom* as the living good which has its knowledge and volition in self-consciousness, and its actuality through self-conscious action … Ethical life is accordingly the *concept of freedom which has become the existing world and the nature of self-consciousness*.
>
> (Hegel 1991: §142)

Ethical life (*Sittlichkeit*) is best understood as the cultural and moral fabric of a society coupled with its institutional structure. So we, like Hegel, exist in a society that predominantly promotes a monogamous, nuclear family, a civil society ruled by the distributive system of capitalism, private property and the free market. Each of these institutions supplies us with an understanding of ourselves that liberates us as particular individuals. As a family member, the individual subsumes his or her own ends to those of others. As a worker, he or she prioritizes his or her long-term welfare and that of his or her family over short-term benefits and sees the benefits of cooperative action. As a political participant, the individual reconciles the universal ends of society with his or her own (the good of all).

The family is the institutional structure that reflects the will's need to associate with others through relationships based on love and care. It provides an immanent doctrine of duty that liberates the individual from his immediate desires and helps him to exercise his reason over the claims of the situation, his identity and his drives. I understand myself immediately as a partner or lover and as a father and I put on hold my most pressing, egoistic desires for the sake of others and do so because I want to, that is, freely. The family allows me to act on motivations that have value and will be recognized as such by others: when I decline an invitation to a conference in order to see my child's nativity play, the possible harm done to my career is understood as both intelligible and justified and mitigated by rational family and employment legislation. The family is a form of liberation because it allows me to prioritize my desires, wants and preferences and acts as an educational environment for the raising of children into the social realm.

Most obviously, the family liberates us from slavish attendance to the sexual drive and turns the unstable reliance on emotion into a stable social form by, for example, rationalizing the rules for when love turns to hate (divorce).

However, the family has needs and these must be satisfied, protected and furthered. So, the head of the family is involved with others in a system of needs that generates and satisfies wants and preferences. Civil society is the realm of labor and consumption and the capitalist economy liberates us from immediate desires by supplying the satisfaction of basic needs in abundance. I recognize those with whom I exchange goods as desiring things as do I and, so, I view them as something to cooperate with to reach our mutual ends. Moreover, it is the economic system that best allows personal expression through the abundances of goods and choices. The fact I can own and exchange property allows me to express what I think is of value: cars, works of art, football games and so on. Recognition is fostered by civil society because I recognize you as an other pursuing your ends, which may or may not coincide with my own, and the system of needs divides labor leading to a unity of needs through the cooperation of specialist workers who indirectly work for the good of society as a whole. More significantly, the division of labor creates "estates" which determine a person's class, friends, lifestyle and values and, as with the family, the estate provides commitments through which persons can overcome the claims of the immediate situation. Therefore, civil society has to make provision for the correct and rational running of the system of needs. This is achieved through an administration of justice in which each person can appeal to his rights as a bearer of rational will. The rights of the system of needs become positive law and in positive law the equality and common aim of all members of the state is secured.

Civil society is built on the satisfaction of needs and because desire determines the codes of practice and behaviors of individuals, it is the extrapolation of the liberty to satisfy one's desires. As we know, such a conception of liberty is one-sided and Hegel is not blind to the normative consequences of using it as the basis for a political system (as most contract theorists do). He views both the creation of the disenfranchised unemployed and also the rampant, alienating egoism of all against all as two aspects of the capitalist economy that need to be controlled. For him, that control is political by which the individual sees the ends of the market as subservient to the end of the state proper. If freedom is conceived of as merely the satisfaction of desire, then the function of civil society can become an end-in-itself and the agent forgets that he is not merely a private person but also part of a larger, rational community. The individual must recognize himself as a member of the state and see its continued existence as his good and does so through political representation.

The "political state proper" limits the excesses of civil society which threaten to alienate people from one another and it also mediates between the interests of social classes and the unity of the state as a whole. It is divided into three parts: the legislative, the executive and the sovereign. The role of the sovereign is to enshrine decisions made by the former two parts into actual law. Hegel believes that the most rational form of such sovereignty is a hereditary monarch. The executive power comprises a civil service enforcing the laws and regulations of the state with the primary aim of curbing the excesses of civil society. Finally, the legislative power consists of the monarch and representatives of the major "estates" divided into two houses, respectively populated by the agricultural class and the confederation of business and industry.

Self-understanding and hence freedom is achieved and guaranteed on three levels: one, the family through unreflective, internal bonds of feeling; two, civil society through

external bonds of cooperation and identity in membership of an estate or social class (agriculture, trade and industry, or public service); and three, the state itself through internal, reflective bonds which allow the person to recognize the family and the estate bound in one concrete identity which is the rational life, inspired no doubt by the General Will of Rousseau where the "I" and the "We" are reconciled in social existence. These three aspects of an individual identity taken together constitute the "Rational state" (as opposed to the "purely political state").

4. The Justification of the State: History, Progress and Metaphysics

Such a brief description of the central elements of Hegel's account of the state does not do justice to the intricacy or subtleties of his extensive discussions, nor does it really pose the question of how far he is merely proposing a conservative agenda. It does seem that the account of the state is ultimately disappointing: we find that the estates override individual concerns, that equality is not yet actual as exemplified by the standing of women and the treatment of the poor, and that the principle of hereditary monarchy is counter to subjective freedom. Furthermore, the lack of any real democratic accountability or participation is also counter to subjective freedom. Surely, like capitalism, democracy is necessary for the expression of the subjective will: just as buying what I want is a rational expression of preference, so the most rational political system would be one in which subjective preferences could be expressed. Hegel does not think so; representation takes place through the medium of class and work, not through votes. If one were to be charitable, one could argue that Hegel does not think that the state in Prussia is history completed; only that it exists in the end of history phase. Just as the idea of equality took four hundred years to develop and is still extending its scope, what he may have believed is that Prussia's state had in place all the correct categories and concepts of rational, ethical life, but these still had to be developed fully. After all, in Hegel's lectures, the central values are liberty, equality and individual welfare and the discussions concern distributive justice, political freedom and rights; the very same values and discussions that dominate contemporary political philosophy.

And that turns us to the question of his liberal status, since it is obvious that much of his rational state, if charitably tweaked, has an affinity with the contemporary, liberal state. It is equally true that Hegel's thought forms the cornerstone for communitarian critiques of liberalism and has a direct and necessary connection with the Marxian rejection of capitalism and its ideological axiologies, such as liberalism. Depending on one's interpretation, Hegel can be apologist for libertarianism and the free market, critic of liberalism's atomism, flag waver for conservative defenders of the *status quo* and justifier of an objective, collectivized spirit of community.

The real debate, though, concerns the justification of Hegel's central values (liberty and equality). The accusation of conservatism made against Hegel can be extrapolated to a general philosophical worry about relativism and the realm of right: why should a particular, individual agent trust the categories and expectations of his or her social fabric if such norms and values have differed in the past? If, on the one hand, his political system, its norms and structure can be justified independently of his philosophical system and in terms of transcendental reason, then he could be considered to be a liberal, if a somewhat unique type. Such a reading would be supported by the *necessary* role of the modern, moral conscience in Hegel's system which guarantees the rationality of a state to the individual as well as the rights and responsibilities that maintains the existence of

the individual conscience. If, on the other, his political system is justified in terms of his overall philosophical system and the dialectical movement of ideas towards their own rationalization, then he rests on the communitarian, Marxian and conservative side of the political line because the individual, in claims against his or her state, has to cede all authority to history. And what would such a justification resemble? It is not enough to putatively assume that history progresses, that nations are more advanced than their predecessors, nor can it be shown to be empirically the case that progress has occurred through history (as one could possibly suppose in the conditional field of dentistry), but Hegel's political system is justified because history is a rational process by which the conceptual wheat is sorted from the contingent chaff and reason reconciles what ought to be with what is through the labor of civilizations and individuals. Yet, such reconciliation is the prerogative of History itself and not of thinking individuals who are (always and necessarily) temporally bound in their own specific ethical lives. And if this is the case, then the cogency of Hegel's political philosophy rests on his metaphysics and his absolute idealism.

Much of Hegel's thought, rather unsurprisingly given many writers' indebtedness to him, pre-empts the debates between liberalism and communitarianism which defined political philosophy for much of the last part of the twentieth century. And yet he himself seemed to offer a sensible, if unspectacular, understanding of the debate. Moreover, without Hegel, there would have been no Marx. His political and social philosophy is both unique and innovative and, yet, so integral to the tradition of Western thought, that whether one wishes to resist his conclusions or to reconstruct his arguments, one has to definitely engage with him. And such an engagement will always be worthwhile.

Related Topics

Kant, Marx, Continental Political Philosophy, Liberalism, Conservatism

References

Hegel, G. W. F. (1977) *The Phenomenology of Spirit*, trans. A. Miller, Oxford: Oxford University Press.

Hegel, G. W. F. (1991) *Elements of the Philosophy of Right*, trans. H. Nisbet, Cambridge: Cambridge University Press.

Further Reading

Hegel, G. W. F. (1971) *Philosophy of Mind: Part 3 of the Encyclopaedia of Philosophical Sciences with the Zusätze*, trans. W. Wallace and A. Miller, Oxford: Oxford University Press.

Hegel, G. W. F. (1979) *System of Ethical Life*, trans. T. Knox and H. Harris; and *First Philosophy of Spirit*, trans. H. Harris, New York: State University of New York Press.

Hegel, G. W. F. (1991) *Philosophy of History*, trans. J. Sibree, New York: Prometheus Books.

Hegel, G. W. F. (1999) *On the Scientific Ways of Treating Natural Law, on its Place in Moral Philosophy, and its Relation to the Positive Sciences of Right*, trans. H. Nisbet, in *Political Writings*, Cambridge: Cambridge University Press.

Knowles, D. (2002) *Hegel and the Philosophy of Right*, London: Routledge.

Neuhouser, F. (2000) *Foundations of Hegel's Social Theory: Actualizing Freedom*, London: Harvard University Press.

Rose, D. (2007) *Hegel's Philosophy of Right*, London: Continuum Press.

Taylor, C. (1975) *Hegel*, Cambridge: Cambridge University Press.

Wood, A. (1990) *Hegel's Ethical Thought*, Cambridge: Cambridge University Press.

12
MILL

C.L. Ten

1. The Principle of Liberty

In the essay *On Liberty* Mill seeks to defend individual liberty against the social tyranny of the majority who, through laws and the pressure of prevailing opinion, seek to suppress both opinions and conduct to which they are opposed. This is the source of the threat to liberty in the stage of progress that civilized societies have now entered. The majority are in power, and they may abuse that power. It is taken for granted that people are entitled to act on the basis of their feelings, even when these are unsupported by reason. The more progressive elements of society do not challenge this view that the mere preferences, "likings or dislikings," of society should be the standards of conduct of all. Instead, they have merely sought to replace one set of preferences for another, which they find more congenial (Mill 1977: 222).

It is only in the area of religious belief that a principled stand has been taken, and freedom of conscience been claimed as "an indefeasible right" (Mill 1977: 222). A principled basis for religious liberty denies absolutely that a person's religious belief is of concern to others. But even here, when the feelings of the majority are still intense, the case for toleration is repudiated. Toleration thrives only in matters of religious indifference.

Until we have a recognized principle for determining the proper scope of interference with the conduct and opinions of individuals, such interference is likely to occur in a haphazard manner. Mill then states the "one very simple principle, as entitled to govern absolutely the dealings of society with the individual in the way of compulsion and control, whether the means used be physical force in the form of legal penalties or the moral coercion of public opinion" (Mill 1977: 223). The principle, sometimes referred to as Mill's principle of liberty, is:

> [T]he sole end for which mankind are warranted, individually or collectively in interfering with the liberty of action of any of their number, is self-protection. That the only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others. His own good, either physical or moral, is not a sufficient warrant. He cannot rightfully be compelled to do or forbear because it will be better for him to do so, because it will make him happier, because, in the opinion of others, to do so would be wise, or even right.
>
> (Mill 1977: 223–4)

The principle of liberty, which applies to adult human beings living in advanced societies, identifies the prevention of harm to others as a relevant reason for interference

with the conduct of others. But it excludes paternalistic and moralistic reasons, which are good reasons for persuading or advising others, but are totally ruled out as the bases for interference in the sense of compulsion or coercion.

Mill enumerates three protected areas of liberty. First is "the inward domain of consciousness," consisting of liberty of conscience, liberty of thought and feeling, "absolute freedom of opinion and sentiment on all subjects, practical or speculative, scientific, moral, or theological," and liberty of expressing and publishing opinion. The second area is "liberty of tastes and pursuits." The third area is freedom to associate with others "for any purpose not involving harm to others" (Mill 1977: 226).

2. The Defense of Liberty

Mill then proceeds to defend his views, starting with a detailed account of liberty of thought and discussion. Freedom to express all opinions is needed, firstly, in order to help us in discovering the truth. The received opinion could be false, or partly false, and some elements of truth would be lost with the suppression of opinions that are opposed to the received opinions. However, Mill is more concerned about our having a deeper understanding and appreciation of the truth, as distinguished from merely having a superficial grasp of true beliefs. So even when the received opinions contain the whole truth, and the suppressed opinions are all false, much is lost when there is no collision between truth and error. We will have no clear understanding of why certain opinions are true or false. We will lose the meaning and ground of true opinions, which will be held as dead dogmas, without any proper influence on our lives, or on the way in which we see the world and frame our approach and perspectives on issues of general interest. In the case of the received opinion of Christianity, Mill maintains that some who regard themselves as Christians are not in fact influenced by it: "The sayings of Christ coexist passively in their minds, producing hardly any effect beyond what is caused by mere listening to words so amiable and bland" (Mill 1977: 249).

For Mill, it is the manner in which the truth is sought and held, when discovered, which is of vital significance. He describes this normative ideal in a number of related ways. People are to base their opinions on the "facts and arguments"; they should aim at "knowing the truth," which requires keeping an open mind to all the evidence and considerations which can be presented for or against an opinion, including alternative and competing views; they should seek a "rational assurance" of the truth of their beliefs (Mill 1977: 244, 231). Censorship is ruled out because, when alternative and opposing views are suppressed, we can have no rational assurance that our beliefs are true, even if they are in fact true. We will hold on to our beliefs without understanding them, or knowing the respective merits or weaknesses of various contending views. The right attitude to adopt to our beliefs is to give "a standing invitation to the whole world to prove them unfounded" (Mill 1977: 232).

Mill believes that "an intellectually active people" will be produced in an environment where there is the greatest liberty in professing and discussing any view of general interest. Freedom is required not only to produce "great thinkers." Rather, it is "even more indispensable to enable average human beings to attain the mental stature which they are capable of." It is in the atmosphere of freedom that "even persons of the most ordinary intellect" can be raised to "something of the dignity of thinking beings." Then we will have "an intellectually active people" (Mill 1977: 243).

Mill then defends the freedom to act on one's opinions. Such freedom, within the limits of not harming others, gives rise to different "experiments of living," in which new activities and ways of conducting one's life can be discovered and put into practice (Mill 1977: 261). But the central thesis is the free development of individuality. People must be free to choose the plan of life they regard as valuable. Blind conformity to custom subverts such choices, and is therefore incompatible with desirable forms of individuality. Choices are partly important in that, when they are regularly exercised, they develop important and distinctive human faculties. The exercise of choice calls into play the faculties of "perception, judgment, discriminative feeling, mental activity, and even moral preference" (Mill 1977: 262). These faculties enable us to make reflective and informed choices between alternative ways of life. We are thereby more likely to adopt ways of life that are better for us, in that they suit our nature or character better, than those that are merely customary. Customs often outlive their value in changed circumstances. In any case, one size does not fit all, and even if a customary way of life is good for some, it will not be suitable for all. By exercising choice, and by looking critically at customs, we are able to adapt to changing circumstances, and thereby thrive better. Mill is not rejecting conformity to customs as such. Such conformity should be reflective, at least at some point before the commitment to follow custom is made, and it should be open to further endorsement or change in light of new circumstances.

Implicit in Mill's view is the idea that some practices or ways of life are objectively better than others. However, given the variety of human characters, it is most unlikely that the same way of life is the best for all. Each would need to find the superior plan of life for herself. The discovery of superior life-styles requires the exercise of reflective choice. But Mill also regards the exercise of choice itself as an indispensable feature of a progressive and enlightened life-style. So such a life-style must not be imposed from above or from below. Each person must choose for herself: "If a person possesses any tolerable amount of common sense and experience, his own mode of laying out his existence is the best, not because it is best in itself, but because it is his own mode" (Mill 1977: 270).

There are therefore, for him, two dimensions along which a way of life is to be compared and evaluated. The first is suggested by what is said to be "best in itself," without reference to choice as a distinctive and identifying element of it. However, even here we need choice, and the human faculties it develops through its regular exercise, to help us to discover what is "best in itself." When Mill says that freedom and variety of situations are needed for the development of individuality, he is partly referring to the role of free choice to conduct experiments in living as a means to discover new and better ways of life. But the instrumental role of free choice is only one of its values. Freedom and individuality are also related more intimately, as a crucial part to a welcome whole. No way of life can reach the desired level of goodness unless it is freely chosen by the person whose life it is.

These two dimensions for assessing the goodness of ways of life are also indicated in Mill's remark, "[i]t really is of importance, not only what men do, but also what manner of men they are that do it" (Mill 1977: 262). He draws a contrast between human beings and machines. A good machine need only satisfy the first dimension of goodness, "best in itself." If something is a good machine for a certain job because it has certain features, then any other machine with exactly those features would also be just as good for the purpose at hand. The manner in which the machine is produced does not affect its goodness. With humans, it is different:

> Human nature is not a machine to be built after a model, and set to do exactly the work prescribed for it, but a tree, which requires to grow and develop itself on all sides, according to the tendency of the inward forces which make it a living thing.
>
> (Mill 1977: 263)

Mill's references to the importance of having "native pleasures," or "feelings of home growth," as opposed to those which are simply derived unreflectively and uncritically from prevailing customs, drive home the same point that human goodness depends in part on choice (Mill 1977: 265).

It is important to recognize the place that the two dimensions play in Mill's plea for freedom and individuality. Choice is one of the dimensions for evaluating a conception of the good life, a valuable and worthwhile human life. But there is a different level of evaluation with which Mill is concerned, and this has to do with the plans of life, or the activities or conduct of individuals or groups, which should be tolerated or permitted by society. Here Mill passionately pleads for the toleration of conduct and ways of life that do not harm others, but which fall below his own standards of goodness, or of which he strongly disapproves or dislikes. His insistence, which we noted earlier, that people's feelings and negative judgments about the conduct of others should cease to be the basis of coercive interference is part of his principle of liberty. The failure to distinguish, or to understand, Mill's different accounts of what is good, on the one hand, and what is to be tolerated or permitted, on the other, is the basis for much confusion about his view, and for unwarranted criticisms of it. This confusion is further compounded by a failure to understand that Mill's conception of toleration does not imply indifference to what is tolerated, and is compatible with, and sometimes even requires, active attempts to dissuade persons from certain conduct through non-coercive means (Ten 2002).

Mill clearly distinguishes between good reasons for compelling people to change their conduct, and good reasons for "remonstrating with him, or reasoning with him, or persuading him, but not for compelling him, or visiting him with evil in case he do otherwise" (Mill 1977: 224). Thus those who do not satisfy Mill's standard of individuality, or whose plan of life falls short of Mill's conception of the good because they blindly follow a custom-flawed practice, would still be tolerated by him if their conduct is not harmful to others. They would be subject to criticism, not coercion. Mill invokes "the right of each individual to act, in things indifferent, as seems good to his own judgment and inclinations" (Mill 1977: 271). He rejects any claims that "the strong man of genius" may impose his more enlightened views on others. The strong person can "point the way," but, "the power of compelling others into it is not only incompatible with the freedom and development of all the rest, but corrupting to the strong man himself" (Mill 1977: 269).

3. Liberty and Utility, and Rawls on Mill

I have outlined Mill's case for individual liberty, both of expression and of conduct. In the relevant areas, he seeks absolute liberty from coercive interference by others. But at the same time he rejects an appeal to "the idea of abstract right, as a thing independent of utility," proclaiming instead that "I regard utility as the ultimate appeal on all ethical questions, but it must be utility in the largest sense, grounded on the permanent interests of man as a progressive being" (Mill 1977: 224). The issue now is whether, or how,

he can show that absolute liberty in certain areas promotes "utility in the largest sense." Elsewhere, I have shown that various sophisticated attempts to derive his principle of liberty from utility are unsuccessful (Ten 1991). I want now to focus on Rawls's views on Mill, which will help us to better understand the character of his arguments for liberty. In fact, Rawls has two different accounts of Mill. In *A Theory of Justice*, he summarily dismisses Mill's defense of liberty as open to all the usual objections to a utilitarian defense of what should be a fundamental value (Rawls 1972: 209–11). But in his posthumously published lectures on Mill in *Lectures on the History of Political Philosophy*, he gives a highly sympathetic account of Mill, concluding that the content of the liberty that Mill defends is very similar to the non-utilitarian Rawlsian liberalism (Rawls 2007: 251–316). Rawls's more well-known attack on Mill in *A Theory of Justice* and elsewhere has itself been subjected to some critical analysis (Ten 1980: 19, 78–9; Amdur 2008). However, his more sympathetic treatment of Mill, including his detailed reconstruction of Mill's case for liberty, is relatively unknown as it only appeared in 2007, a few years after Rawls's death.

In *A Theory of Justice*, Rawls maintains that Mill applies the choice criterion of value to defend free institutions. According to the choice criterion, "one activity is better than another if it is preferred by those who are capable of both and who have experienced each of them under circumstances of liberty" (Rawls 1972: 209). Free institutions are required for the development of people's capacities and powers. They also allow people to have rational and informed preferences about different activities. In addition, free institutions have intrinsic value as historical experience shows that human beings rationally prefer to live under free institutions. But Rawls argues that these considerations in favor of liberty do not guarantee equal liberty for all unless one makes assumptions such as those about the equal capacity of individuals for the activities and interests of progressive beings, and the diminishing marginal value to persons of basic rights. Without these assumptions, the goal of maximizing the sum of intrinsic value, or the greatest net balance of the satisfaction of interests, can be achieved by denying the liberty of some. Rawls assumes that Mill's defense of liberty in terms of "utility in the largest sense" still shares the central features of the more usual forms of utilitarianism in that it seeks to maximize what is good. This is the feature that leads Rawls to maintain that Mill's argument for equal liberty for all "relies upon precarious calculations as well as controversial and uncertain premises" (Rawls 1972: 211). It is also the feature that could lead to the suppression of the liberty of unpopular minorities, as Rawls makes plain in his example of a majority in a society who abhor some of the religious and sexual practices of a minority. The abhorrence is so intense that the mere thought of the practices taking place is enough to arouse anger and hatred. Rawls then points out that "[s]eeking the greatest satisfaction of desire may, then, justify harsh repressive measures against actions that cause no social injury" (Rawls 1972: 450).

Later, in defending his political liberalism, Rawls argues that the maintenance of "one comprehensive religious, philosophical, or moral doctrine" can only be done by "the oppressive use of state power" (Rawls 1993: 37). He further claims that this is true even for Mill's reasonable liberalism: "A society united on a reasonable form of utilitarianism, or on the reasonable liberalisms of Kant and Mill, would likewise require the sanctions of state power to remain so" (Rawls 1993: 37). Presumably, it is the allegedly maximizing element in Mill's account of liberty that leads Rawls to claim that Mill has to resort to the oppressive use of state power in order to achieve his goals. For the mere replacement of orthodox Benthamite utility, in the form of undifferentiated pleasures and the absence

of pain, with Mill's sense of "utility in the largest sense," including the development of individuality, will not in itself remove the maximizing element, and hence the use of oppressive state power. As H.L.A. Hart points out, it is possible that "a greater realization of individuality summed over all would be secured if the individuality of some were suppressed so as to permit its greater realization by others" (Hart 1982: 97). Hart further maintains that this possibility is not merely theoretical, for it is arguable that the oppression of the individuality of minorities has been justified in this way in some great cultures which "have flourished on the dunghill of slavery" (Hart 1982: 97).

But my account of Mill's principle of liberty and his appeal to the "free development of individuality" provides sufficient evidence to show that the damaging maximizing element is not part of Mill's view. Even ordinary people in the type of society for which Mill was writing had the relevant liberty and the right to lead their lives in accordance with their own views and values. The limits of their liberties are set by harm to others, and not by the fact, if this should turn out to be a fact, that restrictions on their liberties, even when they are not harming others, will lead to the good of others, in, for example, enabling them to better cultivate their individuality. The relevant liberties for all are absolute, and not to be forcibly traded off for the greater benefit of others. As Hart rightly notes, Mill's principle of liberty is not an aggregative principle of general welfare, but a distributive principle (Hart 1982: 101). It sets out certain conditions that must be present before the principle applies. But once society has reached a certain stage of its development, and the person is a normal adult, the principle maintains: "The only part of the conduct of any one, for which he is amenable to society, is that which concerns others. In that part which merely concerns himself, his independence is, of right, absolute" (Mill 1977: 224). Mill explicitly rules out considerations of the general welfare as inadmissible unless one can show that the interests of others are harmed. So the trade-off of one person's welfare for another's in order to maximize the general welfare cannot even be considered unless the individual's conduct crosses the threshold of harming others. As he puts it so clearly and explicitly:

> As soon as any part of a person's conduct affects prejudicially the interests of others, society has jurisdiction over it, and the question whether the general welfare will or will not be promoted by interference with it, becomes open to discussion. But there is no room for entertaining any such question when a person's conduct affects the interests of no persons besides himself
>
> (Mill 1977: 276)

It is therefore clear that for Mill the right to individuality acts as a side-constraint on the maximization of the general welfare. So long as persons are seeking to promote their own individuality, or even simply following the traditional life-style, and refusing to exercise reflective choice between different plans of life, they may not be coerced for the greater overall good. It is only when they cross the threshold of harming others that their interests can be traded off against the greater interests of others in order to maximize the good. Their conduct within their absolutely protected sphere might not satisfy Mill's conception of a good and worthwhile life, but they can only be persuaded and advised to improve themselves, but cannot be compelled to do so, even by those who are abler or more enlightened.

Contrary to Rawls's claim, Mill was strongly opposed to using the oppressive power of the state to promote any comprehensive conception of the good, including his own.

He made this clear in his attacks on Comte's despotic system (Mill 1977: 227). As we know, he himself accepted Comte's Religion of Humanity, but strongly rejected the Comtian approach to its propagation, which he once described as *liberticide* (Mill 1972: 29). Mill wanted to ensure that there is a protected space for individuals to lead their lives in accordance with their own conceptions of the good. No conception of the good, even one that is "best in itself," may be imposed by the use of state or spiritual power. Rawls's failure to see this is a striking defect of his account of Mill in *A Theory of Justice* and elsewhere.

However, in his lectures on Mill in *Lectures in the History of Political Philosophy*, Rawls gives a more detailed and sympathetic interpretation of Mill's principle of liberty. He holds on to the view that the principle of liberty is not a supreme principle, but a principle that is subordinate to the Principle of Utility, although for Mill utility is reinterpreted as "the permanent interests of man as a progressive being." Rawls explains that the principle of liberty applies "absolutely" because, in order to advance the interests of man as a progressive being, it is better to affirm the principle without exception, even in the case where there is only a single dissenter to be silenced or suppressed (Rawls 2007: 294). Human societies can continue to improve until we reach the best society, which is a normal and natural state in which there is full equality, with equal basic justice and liberty for all. In order for such progress to be possible, certain necessary conditions must be present. One of the permanent interests is an interest in the necessary conditions for continued progress towards the state of equality, and an interest in remaining in that state once we reach it. Our nature as social beings is realized in the state as we exercise and develop our higher faculties, and satisfy our most important wants in a manner consistent with equal justice for all and the legitimate interests of others.

A second permanent interest is in the social conditions generating freedom of thought and liberty of conscience, which are required for the discovery of truth and for its application in all significant matters. Our third permanent interest is in the conditions of individuality, in securing our liberties of tastes and pursuits, our liberty in forming our mode of life to suit our character, and the liberty to associate with like-minded people. All these liberties are needed both for reaching the best state of society, as well as for sustaining that state when reached. It is only in the presence of free institutions that people can come to have the knowledge to make reasonable decisions about the plan of life that is most likely to give them happiness.

The fourth permanent interest is in just and free institutions and the attitudes required to bring about the normal and natural state of equality, in which the interests of all are considered equally, and in which we have the desire to live in unity with others.

So Rawls attributes to Mill the view that in order to maximize utility in the sense of our permanent interests as progressive beings, we need to set up just and free institutions. Without them society is unable to acquire the relevant knowledge and information about what will maximize utility. Only persons raised and educated under just and free institutions can develop their faculties in ways that are best suited to their own character and desires. According to Mill's decided preference criterion, only such persons can properly judge one pleasure or activity to be higher or better than another. All normal persons are equally capable of enjoying and exercising their higher faculties. They would prefer a range of activities as central to their lives, and the decided preferences of all are equally valuable. In arguing that equal rights to justice and liberty will maximize general social utility, Mill relies on psychological principles of human nature, such as that we have a powerful natural sentiment to be at unity with others. This desire

for unity is strengthened by certain features of civilization, an increasing equality and greater social cooperation for joint purposes. A society in which people regard each other as equals can only exist if equal consideration is shown to the interests of all.

Rawls concludes his lectures on Mill by saying that he is not "much concerned with the overall success of Mill's view" (Rawls 2007: 313). He wants to explain how Mill "managed to end up with principles of justice, liberty, and equality not all that far away from justice as fairness, so that his political and social doctrine—lifted from his overall moral view—could give us principles of a modern comprehensive liberalism" (Rawls 2007: 313). Rawls still sees Mill as giving a consistent and systematic utilitarian defense of liberty. But he has set aside all the trenchant objections he had raised against any utilitarian defense of liberty: its maximizing requirement which would allow the sacrifice of the liberty of some for the greater general utility, the resort to oppressive state power to achieve desired ends, and the precariousness and uncertainty of the utilitarian calculation in support of liberty. His view would be understandable if Rawls now acknowledges that there are some non-utilitarian features of Mill's defense of liberty. But this is not the case. In all his discussions of Mill on liberty, Rawls never deviated from his belief that Mill always treated liberty as a value derivative from the ultimate Principle of Utility. Thus in his last publication before his death, *Justice as Fairness*, Rawls summarizes the unity of Mill's views in a manner that is similar to the lectures. He again points out that this unity rests on "a few psychological principles," and that "Mill connects his conception of utility with the permanent interests of humankind as progressive beings" (Rawls 2001: 147). Mill's claim is that "in the conditions of the modern world, following his principles of justice and liberty is an effective if not the best way to realize these permanent interests" (Rawls 2001: 147). But now, unlike in the lectures, Rawls doubts whether the psychological principles will hold: "from commonsense knowledge and ordinary experience, Mill's principles may seem an excessively optimistic view of our nature" (Rawls 2001: 147).

Rawls's account of Mill in the lectures totally ignores the non-utilitarian elements in his case for various liberties. He includes these liberties as among the permanent interests, but he treats these permanent interests as merely essential means for arriving at the best society and for sustaining it, when reached. But for Mill the liberties are partly constitutive of "utility in the largest sense." Thus the end state that Mill seeks in arguing for freedom of thought and discussion is not simply the discovery of truth, but *knowing* the truth, which requires that the truth be acquired and retained in an atmosphere of freedom, with "a standing invitation to the whole world" to prove one's beliefs wrong (Mill 1977: 232). It is only in an atmosphere of freedom, rather than that of censorship, that we can have the society that Mill cherishes, where people have "the dignity of thinking beings" (Mill 1977: 243). Similarly, when it comes to conduct, it is the *free* development of individuality that is the end to be reached. All worthwhile forms of human life have to be freely chosen.

4. Mill on Individuality

The liberties that Mill identifies for protection and promotion are among the permanent interests of man as a progressive being, and as such they are important constituents of his notion of utility. But to those who do not as yet appreciate their intrinsic value, Mill points to their instrumental value in promoting values that they associate with utility. Eventually these people will come to value the liberties as forming a part of

utility. As Mill explains in *Utilitarianism*, what was initially regarded as a means to an end, can by constant association with that end, become a part of the end (Mill 1969: 235–7). This is not how Mill himself sees the value of the liberties he defends, but he has to convince skeptical minds. He has to educate and persuade them about the value of liberty by showing that it enables persons to discover plans of life that best suit them. Mill hopes, and perhaps expects, that these people will subsequently also adopt his normative ideal of individuality, and understand that part of what makes a way of life the best for them is that they themselves have freely chosen it, and continue to endorse it under conditions of freedom.

Rawls thinks that Mill ensures equal basic liberties for all by treating all normal persons as equally capable of enjoying and exercising their faculties. However, this does not seem to be Mill's view. His account of individuality acknowledges that some, "the strong man," for example, or "a more highly gifted and instructed One or Few," are more capable of developing their individuality than others (Mill 1977: 269). He also laments the unwillingness of some to respond to his plea for individuality, being satisfied with the use of no other faculty than "the ape-like one of imitation" (Mill 1977: 262). Nonetheless, as we have already noted, the development of individuality is "the right of each individual to act in things indifferent, as seems good to his own judgment and inclinations" (Mill 1977: 271n). *Each* individual has this right, whether she chooses to exercise it wisely or foolishly. Mill's principle of liberty sets the threshold of harming others that has to be met before one's freedom of action may be restricted. Similarly, there is to be the "fullest liberty of professing and discussing, as a matter of ethical conviction, any doctrine, however immoral it may be considered" (Mill 1977: 228n). No interference with liberty could be justified unless the circumstances in which opinions are expressed are such as to make their expression "a positive instigation to some mischievous act" (Mill 1977: 260).

Mill does not regard certain kinds of adverse effects that our opinions or conduct have on others as harmful to them. Specifically, their intense disapproval or dislike as such could never be used as a proper basis for restricting liberty. These are relevant effects, considered from the usual utilitarian perspective, and Rawls relies on the widespread and intense abhorrence of the majority to state the utilitarian case for suppressing certain religious and sexual practices of the minority. But by confining the utilitarian calculation to the harmful effects, as he defines and confines them, Mill insulates liberty from the precarious calculations to which Rawls fears that it is exposed. With Mill, although not with more traditional utilitarians, liberty is in safe hands in the changing and uncertain circumstances of the future. The closest Rawls comes to recognizing this is when he maintains that Mill's principle of liberty excludes three kinds of reasons from public political discussions, namely, paternalistic reasons, reasons of excellence and ideals of human perfection, and reasons of dislike or disgust or of preference, when these are not supported by reasons of right or wrong (Rawls 2007: 291). He points out that the result is that "only certain kinds of reasons—only certain kinds of utilities—are appropriate to invoke in Mill's form of public reason" (Rawls 2007: 293). But he does not acknowledge that the exclusion of these reasons removes potentially anti-liberal elements in the utilitarian framework, thereby undermining his objections that Mill's case for liberty is dependent on precarious utilitarian calculations, and that it allows for the use of oppressive state power. Whereas Mill's exclusion of these reasons is peremptory, unqualified, and fundamental, Rawls believes that it is derivative from utilitarianism, together with some uncertain psychological assumptions. Mill's conception of utility is liberal-friendly precisely because it excludes certain utilities.

It is true that Mill emphasizes the growth of the social feelings and the increasing desires that people will have, as civilization advances, to be at unity with their fellow beings, to share interests with them, and to treat them as equals. But the growth of these natural feelings is intended to provide support in human nature for the general welfare as against the pure pursuit of self-interest. The social tendencies do not settle the issue of when it would be just to expect uniform action from all, and when the diversity resulting from individual liberty should prevail. As Mill maintains:

> To be held to rigid rules of justice for the sake of others, develops the feelings and capacities which have the good of others for their object. But to be restrained in things not affecting their good, by their mere displeasure, develops nothing valuable … To give any fair play to the nature of each, it is essential that different persons should be allowed to lead different lives.
>
> (Mill 1977: 266)

Mill does not take for granted that the desire for liberty will grow naturally, nor does he assume that those who have themselves enjoyed liberty will automatically respect the liberty of others. He believes that there is also a disposition in human nature for mankind "to impose their own opinions and inclinations as a rule of conduct on others," and that we have to resist this by erecting "a strong barrier of moral conviction" (Mill 1977: 227). His principle of liberty seeks to guarantee that each and every person has, in the appropriate region, freedom of thought, discussion, and action.

Related Topics

Liberalism, Contractualism and Political Liberalism, Utilitarianism and Consequentialism, Perfectionism, Pluralism, Natural Law and Rights Theory, Freedom, Autonomy, Rights, Toleration, Paternalism, Moralism, and Markets, Religion in Public Life

References

Amdur, R. (2008) "Rawls's Critique of *On Liberty*," in C.L. Ten, ed., *Mill's* On Liberty*: A Critical Guide*, Cambridge: Cambridge University Press.

Hart, H.L.A. (1982) "Natural Rights: Bentham and John Stuart Mill," in H.L.A. Hart, *Essays on Bentham*, Oxford: Clarendon Press.

Mill, J.S. (1969) "Utilitarianism," in J.M. Robson, ed., *The Collected Works of John Stuart Mill*, vol. 10, Toronto & Buffalo: University of Toronto Press, and London: Routledge & Kegan Paul.

—— (1972) "Letter 213. To Harriet Mill (Jan. 15, 1855)," in F.E. Mineka and D.N. Lindley, eds., *The Collected Works of John Stuart Mill*, vol. 14, Toronto & Buffalo: University of Toronto Press, and London: Routledge & Kegan Paul.

—— (1977) "On Liberty," in J.M. Robson, ed., *The Collected Works of John Stuart Mill*, vol. 18, Toronto & Buffalo: University of Toronto Press, and London: Routledge & Kegan Paul.

Rawls, J. (1972) *A Theory of Justice*, Oxford: Clarendon Press.

—— (1993) *Political Liberalism*, New York: Columbia University Press.

—— (2001) *Justice as Fairness: A Restatement*, Cambridge, MA and London: The Belknap Press.

—— (2007) *Lectures on the History of Political Philosophy*, Cambridge, MA and London: The Belknap Press.

Ten, C.L. (1980) *Mill on Liberty*, Oxford: Clarendon Press.

—— (1991) "Mill's Defense of Liberty," in J. Gray and G.W. Smith, eds., *J.S. Mill* On Liberty *in Focus*, London and New York: Routledge.

—— (2002) "Was Mill a Liberal?" *Politics, Philosophy & Economics* 1: 355–70.

Further Reading

Berger, F.R. (1984) *Happiness, Justice, and Freedom*, Berkeley: University of California Press.
Donner, W. (1984) *The Liberal Self*, Ithaca, NY: Cornell University Press.
Feinberg, J. (1984) *Harm to Others*, New York: Oxford University Press.
Gray, J. (1996) *Mill on Liberty: A Defense*, 2nd edn., London: Routledge.
Rees, J.C. (1985) *John Stuart Mill's* On Liberty, Oxford: Clarendon Press.
Riley, J. (1998) *Mill on Liberty*, London: Routledge.

13
MARX

David Leopold

1. Life and Work

Karl Marx was born in 1818 to a Jewish family, living in the Rhineland region of Prussia, with a distinguished rabbinical ancestry. His father was a lawyer whose conversion to Christianity had enabled him to continue his legal career. Marx studied at the universities of Bonn and Berlin, writing a doctoral thesis on the philosophies of nature of Democritus and Epicurus (containing much on the "swerve" of the atom). Marx's subsequent life combined independent scholarship, radical political activity, and financial insecurity, in varying proportions. In order to associate and publish as he wished, given contemporary political conditions, Marx had to spend most of his adult life living outside of Germany. There were three successive periods of exile. Between late 1843 and early 1845, he lived in Paris, a cosmopolitan city full of émigrés and radical artisans (he was expelled by the government following Prussian pressure). Between early 1845 and early 1848 he lived in Brussels, the capital of a rapidly industrializing Belgium (he was expelled by Royal decree following demonstrations involving foreigners). Finally, from late 1849, he made his family home in London, which provided both a stable haven for political exiles and a superb vantage point from which to study the most advanced economy in the world. Marx devoted much of his time and energy to the latter task, although only the first volume of *Capital*—his epic examination of the capitalist mode of production—was published during his lifetime. Together with his close friend and collaborator Friedrich Engels (1820–95), Marx was politically active throughout his life. The most intense periods of activity centered on the revolutions of 1848 (when he returned to Paris and Cologne), and the International Working Men's Association (the First International) between 1864 and 1874 (especially around the Paris Commune of 1871). His later years were marked by illness, the death of his wife Jenny (in 1881), and travel aimed at improving his health—with Algiers, the Isle of Wight, Jersey (in the Channel Islands), and Karlsbad, amongst the convalescent destinations. Marx died in March 1883, two months after the death of his eldest daughter, also called Jenny. His estate was valued at £250.

Marx was a prolific writer. The authoritative edition of his writings, the *Marx-Engels-Gesamtausgabe*, is still a work in progress, but in current form—already much scaled-down from the original plan—will contain 114 large volumes (of which roughly half have been published as of 2012). The edition was started in Moscow and (East) Berlin in the 1960s but is now produced under the auspices of an international panel of experts, the Internationale Marx-Engels-Stiftung (IMES). In addition to his various published and unpublished works (including much-studied texts such as the *Communist Manifesto*

and *The Eighteenth Brumaire of Louis Napoleon*), the edition contains Marx's journalism, correspondence, drafts, and (some) notebooks.

There is a lot of Marx to read, and not all of it is easy. There is plenty of clear and powerful prose, but also much that is opaque and less certain. In addition to the complexity and unfamiliarity of some of his ideas, the interpretative difficulties facing the reader include texts: which are written in a number of languages (English, French, and German); which proceed via a close critical engagement with long-forgotten contemporaries; which were published under conditions of censorship; which use an unfamiliar intellectual idiom; which presuppose unfamiliar assumptions; which were written primarily for self-clarification; and so on.

One might well wonder whether it is worth the effort required to wrestle with these texts. It used to be said that one should read Marx because of his historical influence. What was meant by "historical influence" was typically that his work had played a central part in the emergence, character, and development, of the Soviet Union. I was never entirely convinced by these claims. That his photograph had hung in government offices in Moscow did not seem a good reason for being interested in his writings (at least, for anyone other than a historian of Soviet Russia). I do not mean to suggest that Marx had no intellectual influence on those historical developments (although that influence was, in my view, more indirect and contestable than often imagined). However, I do think that this indirect and contestable historical influence constitutes a poor reason for most of us bothering to read him. (Similarly, I do not think that we should read Locke *because* he had an indirect and contestable historical impact on the American revolution, nor read Rousseau *because* he had an indirect and contestable historical impact on 1789.) Rather, we should be interested in Marx (like Locke and Rousseau) because they have interesting, insightful, and relevant things to say on important issues of political and social philosophy.

Specifying what these things are in Marx's case is not uncontroversial, but any thorough survey of candidates would surely include: his theory of history; his account of capitalism; his philosophical anthropology (less grandly, his view of human nature); and his vision of socialism. I will say something about each of these four subjects in turn.

2. Theory of History

Perhaps the first point to emphasize about Marx's theory of history is that he did think that we could talk meaningfully about the history of humankind. (My account here follows Marx (2000: 424–427) and Cohen (1978).) That is, despite the enormous diversity of human activities and arrangements (across time and across cultures), Marx holds both that some overarching narrative does emerge, and that it can be explained.

For Marx, the overarching narrative consists in the growth of human productive capacity. Humankind depends on the natural world for its survival, but nature is initially stingy with its blessings. As a result, individuals have to overcome this scarcity by the expenditure of effort and ingenuity. The capacity of humankind to produce the things required to meet their various wants and needs tends to improve constantly. The historical results include not only a growth in productivity, but also a consequent transformation of both the natural and social worlds.

For Marx, the overarching increase in productivity takes place within a series of discrete forms of society. He sees each of these successive forms of society as organized around a particular economic structure. Examples of forms of society organized around

economic structures include what he calls "Asiatic society" (that is, pre-classical societies such as Sumerian Mesopotamia and Pharaonic Egypt), "ancient society" (that is, the slave economies of Greece and Rome), feudalism, and capitalism.

The relation between the different forms of society and the overall narrative is crucial. Marx maintains that these societies rise and fall according to whether they facilitate or frustrate, respectively, the growth of human productive power. For a time (it might be a long time), a particular economic structure will further the development of productive power. However, a limit to the growth of productive power that can occur with that economic structure will eventually be reached. And at that point—when it is restricting the growth in human productive power—that economic structure is liable to be replaced by another form of society. And so the process continues (giving the theory a predictive dimension). Or, in another (more contestable) version of the story: so the process continues, until we reach the form of society known as communism.

The relation between the economic structure and (at least, some of) the legal and political institutions of a society is also central to Marx's account. He portrays the economic structure as providing the "base" on which a legal and political "superstructure" arises. By this architectural image he seems to mean that the economic structure (which itself corresponds to the level of development of the productive power) explains the character of (at least, some of) the legal and political institutions that exist alongside it. This is not to deny that the latter have any effect on the former. Indeed, it is a central part of Marx's account that the legal and political superstructure helps to stabilize and protect the economic base. As a result, particular superstructures are said to rise and fall along with the bases that they help stabilize and protect.

This basic sketch of Marx's account might provoke numerous questions and doubts. Some of these queries might constitute requests for clarification.

Thus, one might wonder how this basic sketch relates to the more technical vocabulary found, for example, in the 1859 "Preface" to *A Contribution to the Critique of Political Economy*. A couple of examples of "translation" might be given here. Thus, Marx often talks about the growth in productive power in terms of "the development of the forces of production," where these forces include both means of production (land, raw materials, tools, and so on) and labor power (knowledge as well as brute force). One suggested measure for the consequent growth in productivity is the decline in the amount of labor that it takes to produce what is required to satisfy the basic physical needs of the immediate producers (Cohen 1978: 55–62). Similarly, Marx often talks about the economic structure in terms of the "relations of production," namely the set of relations of power that individuals have over the productive forces (means of production and labor power). These relations divide the population into economic roles that form the basis of class divisions. And he often uses "ownership" as a kind of shorthand for the economic power that these groups have over those forces; thus talk of proletarians "owning" their own labor power is equivalent to the claim that they have effective control over it.

Similarly, one might wonder how this basic sketch relates to other claims that Marx makes about history, and perhaps especially to the *Communist Manifesto* claim that the history of all previously existing societies is the "history of class struggle." Some see a tension here, but the suggestion that major historical changes are brought about by class struggle is compatible with the claim (in the basic sketch) that the fundamental explanation of history is located elsewhere. For instance, class struggle might have brought about capitalism, and yet one could still maintain that it is necessary to look beyond class struggle (to the more fundamental relationship between forces and relations of

production) in order to understand why it brought about this particular change rather than some other (a reversion to slave society, perhaps). On Marx's account, it seems that the class which either rules throughout a period of stability or emerges triumphant from a period of epochal conflict, does so because it is best equipped to preside over the development of the productive forces at the given time (Cohen 1988: 14–18).

Other queries about this basic sketch might be critical rather than clarificatory in character. They include concerns about the coherence and truth of Marx's theory.

As an example of a coherence worry, consider the consistency of various claims about the interaction of the forces and relations of production. In particular, Marx's view that the relations of production have a significant impact on the development of the forces of production is sometimes thought to be in tension with his insistence that the relations of production are determined by the forces of production. G.A. Cohen proposes that "functional explanation"—whereby something which has a certain effect is explained by the fact that it has that effect—can tie these various claims together in an intellectually satisfactory manner. On this account, the impact of the relations on the forces is no embarrassment to the purported primacy of the forces because of the manner in which the development of these forces explains the relations: economic structures are said to rise and fall according to whether they sustain or frustrate the development of those productive forces (Cohen 1978: 278–296).

As an example of a truth worry, consider the claim that the productive forces tend to develop throughout history. On a strong reading, Marx is claiming that the productive forces tend to develop over the history of *each* discrete kind of society (and therefore also over the course of human history as a whole). In other words, every kind of society is said to have an internal tendency to innovate economically. Some historians see that view as an overhasty, and implausible, generalization from the case of capitalism. It might be suggested that a more cautious reading of the historical evidence supports only a weaker reading of the development claim: namely, that whilst forms of society tend to preserve the productive accomplishments of their predecessors, they do not all contain an internal tendency guaranteeing the continual development of productive power.

3. Assessment of Capitalism

"Capitalism" (especially used as a noun) was a relatively new word (which Marx helped popularize) for a relatively new form of society (whose dynamics he worked hard to understand). Two usages predominate in his work: the first defines capitalism in a structural manner, as a society in which the immediate producers own no productive force other than their labor power (and so, in order to survive, are forced to sell that labor power to those others who do own means of production); the second defines capitalism in terms of its driving purpose, namely the accumulation of capital (that is, the exchange of money for commodities in order to make more money). Putting the two together, we might say that Marx saw capitalism as a form of society organized around an economic structure in which immediate producers own only their own labor power, and the goal of production is the remorseless pursuit of profit.

The central point about Marx's assessment of capitalism that I want to emphasize here is that it is not an unremittingly negative one. He obviously (and appropriately) considered himself an opponent of capitalism, but his enmity was always colored by an appreciation of its achievements. The result is a more even-handed, and more interesting, assessment of capitalism than is sometimes realized.

Here I can only give a flavor of that account. Rather than simply listing numerous achievements and failings, I will discuss one example of each. (Those concerned about this paucity might note that a further failing—namely, alienation—appears in the discussion of Marx's philosophical anthropology, and that a further achievement—namely, bringing about the objective and subjective conditions for socialism—appears in the discussion of his vision of socialism.)

My example of a failing of capitalism is exploitation. Discussions of Marx's account of exploitation are often embedded in discussions of his labor theory of value and his attitude towards what we might (cautiously) call the "wrongfulness" of capitalism. I will say a little about the latter but nothing about the former (not least, because I am persuaded that—whatever Marx might himself have thought—his account of exploitation does not depend on the labor theory of value).

The proper characterization of Marx's account of exploitation is much debated. The essential point looks to be that capitalism operates in such a way that the capitalist retains (without recompense) a portion of the value of the product of the worker's toil. The worker creates the product (that which has value), but receives in payment less than the value of that product. The capitalist, in turn, appropriates some of the value of the product which (qua capitalist) they have not produced. However, some maintain that the performance of unpaid labor is only exploitative when it results from a pertinent vulnerability (Wood 2004: 246–253). In the capitalist case, the obvious candidate—a vulnerability in which Marx shows great interest (not least, historical interest in how it came about)—is the worker's lack of access to means of production (and their consequent need to sell their own labor power in order to survive). Adopting less dry and static language, we might say that it is in the interests of the capitalist to extract as much productive effort from the worker at as little cost as possible (and that this results, primarily, from the dynamic of the economy, rather than from the moral character of the individual capitalist).

Exploitation is at the heart of a lively debate about Marx's views on the wrongfulness, or otherwise, of capitalism. There are few positions in this debate which have not been occupied by commentators: some suggest that Marx did not condemn capitalism (but only analyzed or explained it); others insist that he condemned capitalism but only on "non-moral grounds" (such as its frustration of community); others claim that he (also) offered a moral condemnation of capitalism but on grounds other than injustice (such as its frustration of self-realization); and others maintain that he (also) morally condemned capitalism in terms of injustice. Not the least difficulty for those who would deny the latter claim is Marx's repeated characterization, in *Capital* and elsewhere, of the capitalist's appropriation of unpaid labor as "robbery," "theft," and "embezzlement" (claims that seem to invoke the idea of injustice in a clear and emphatic manner).

Note that even those who agree that exploitation constitutes an injustice for Marx, often disagree about adjacent issues. For example, they might disagree about the precise nature of the injustice here; whether exploitation violates: a "reciprocity" claim linking income entitlement with a requirement to work; a "Lockean" right to one's labor product; a prohibition on certain kinds of coercive interference with others; or some other principle. In addition, they might disagree about whether, and to what extent, Marx recognized the character of his own views about this injustice (or whether, and to what extent, he suffered from some kind of deceptive self-understanding on this issue).

My example of an achievement of capitalism is perhaps less familiar. It has been called the freedom of "detachment," involving, as it does, a liberation from the "engulfment" which characterized pre-capitalist societies.

Marx had little time for what is sometimes called the "romantic" critique of capitalism. On this "romantic" view it is modern industry and technology (rather than, for example, the social relations within which they are utilized) that are the primary evil, responsible in particular for undermining the communitarian idyll of pre-capitalist societies. In contrast, Marx celebrates the bourgeoisie's destruction of the parochial, customary, and deferential aspects of the feudal world, and insists that technological growth and human liberation (ultimately) go hand-in-hand.

There is an important complexity in this celebration. One way of trying to convey that complexity utilizes a "dialectical" presentation of the developing relationship between a "self" (the individual) and its "other" (say their social position and the wider community). The developmental pattern here is a triadic one: starting from a pre-capitalist class society (a stage of "undifferentiated unity") where the self is "engulfed" in its other and has no identity apart from it; progressing through capitalist society (a stage of "differentiated disunity") where the self is freed from engulfment but now experiences its other as something alien and oppressive; and ending in communism (a stage of "differentiated unity") where the self and other remain distinct but harmony between them is restored (Cohen 1988: 183–196). You might think of (one simplified aspect of) this triadic development as involving successive stages of: community without individuality; individuality without community; and community with individuality.

What that "dialectical" presentation illustrates is that there is a gain as well as a loss in the second stage of the development (capitalism). The loss is the disappearance of harmony (which confirms that our emancipation is as yet incomplete), but the gain is that the individual breaks free of "engulfment" in their social position and community. It is this gain which helps explain Marx's claim that capitalism has "rescued" the majority of the population from the "idiocy" of pre-capitalist "rural" life, a life of limited appetites, local prejudices, and unthinking deference (Marx 2000: 246–249). There is, of course, something insatiable and problematic in the capitalist pressure for growth and progress, but Marx undoubtedly values the ways in which, as a result, human aspirations have been expanded beyond the parochial, hierarchical, and customary (Berman 1999: 100–110).

4. Philosophical Anthropology

Discussions of Marx's philosophical anthropology, his understanding of human nature, are often dominated by two issues which I will mention only briefly here.

The first is Marx's attitude towards the universality of human nature; that is, towards aspects of human nature that are constant across history and between cultures. There is a long-lived but problematic suggestion that he denies that human nature has any universal elements. The suggestion is "problematic," not least, because the evidence against it is overwhelming (Geras 1983). Marx clearly and explicitly affirms that human nature has both constant and mutable elements; that is, that human beings are characterized by universal qualities, constant across history and culture, and variable qualities, reflecting historical and cultural diversity (McMurtry 1978: 19–53). In *Capital*, he refers to them as "human nature in general" and "human nature as modified in each epoch," respectively.

The second is Marx's account of the *differentia specifica*; that is, those features that distinguish human beings from other animals. Commentators identify a variety of candidates in his writings for this role, but most often select the claim that humankind alone

engages in productive activity (although explaining precisely what is distinctive here is not easy). Marx might well be mistaken in thinking that other views that he held—for example, about the value of self-realization in work—depended on the truth of these claims about uniqueness.

I am more interested here in Marx's account of human flourishing, his account of what is required for beings like us to do well. However, I introduce that subject somewhat obliquely via some remarks about alienation. The latter concept has already appeared in the discussion above—in the second stage ("differentiated disunity") of the dialectical triad discussed under the achievements of capitalism—but without a proper introduction and not under its own name.

At the most general level, the concept of alienation suggests a kind of dysfunctional relation (for example, an improper separation or hostility) between entities. As such, alienation seems always to involve the loss, or lack, of something of value (in the present example, the loss, or lack, of the proper connectedness or harmony between the relevant entities).

Accounts of this dysfunctional relation are sometimes divided into "subjective" and "objective" variants. Alienation is *subjective* when it is said to consist in the presence (or absence) of certain beliefs or feelings (for example, when individuals do not feel "at home" in the world, or when they do feel "estranged" from the world). In contrast, alienation is *objective* when it is described in terms that make no reference to the beliefs or feelings of individuals (for example, when individuals fail to develop and deploy their essential human characteristics, whether or not they experience that lack of self-realization as a loss). It will be apparent that, as described, these two forms of alienation can be exemplified separately or conjointly in the lives of particular individuals or societies (Hardimon 1994: 119–122). Marx sees both subjective and objective forms as prevalent in capitalist society, but typically attributes some kind of primacy to the latter. He seems to hold, for example, that, whilst subjective alienation undoubtedly exists, it usually (if not always) tracks its objective counterpart.

As characterized, accounts of alienation have to provide criteria of "impropriety"; that is, they have to offer some explanation of what makes any putative separation or hostility dysfunctional (since some hostilities and separations are presumably entirely appropriate). The suggestion here is that, for Marx, relations between individuals and their social and political arrangements are appropriately dysfunctional when they involve a separation that frustrates human flourishing.

Reconstructing Marx's account of human flourishing is somewhat fiddly, but his scattered remarks about non-volitional human needs in the early writings are especially rich in this regard. (Significantly, this part of his early writings is, I would maintain, consistent with later works.) In order to flourish, the essential capacities of the individual must develop in a healthy and vigorous manner. For Marx, this is possible only in a society that satisfies not only basic needs (for sustenance, warmth and shelter, certain climatic conditions, physical exercise, basic hygiene, procreation and sexual activity), but also less basic social needs, both those that are not always seen as a distinctive part of his account (for recreation, culture, intellectual stimulation, artistic expression, emotional satisfaction, and aesthetic pleasure), and those that are often seen in this way (for fulfilling work and meaningful community) (Leopold 2007: 227–245).

This account of the various needs requiring satisfaction in a society in which human beings would genuinely flourish raises many questions. Typical worries concern the possible omissions, emphasis, and extravagance of Marx's account. Does he exclude any

important needs? Does he exaggerate or underestimate the importance of certain needs? And does he set a standard for human flourishing which no society could ever reach?

Those questions won't be tackled here, but I will say a little more about one of the distinctive items on Marx's list, namely fulfilling work. There is a familiar view of work as a necessary evil, as something unpleasant which we are bound to do. It is sometimes called the "Christian" view of work as a curse, of work as the result of the Fall which required us, thenceforth, to survive by the sweat of our brow. In contrast, Marx denies that the negative character of work is part of our fate, a fact about the human condition that no amount of social change can remedy. He maintains, rather, that work has the potential to be something creative and fulfilling. Productive activity, on his account, is a central element in what it is to be a human being, and any self-realization worth the name will involve self-realization through work.

That Marx does think that—in a different form of society—work could be creative and fulfilling, partly explains the intensity and scale of his condemnation of the experience of work under capitalism. The contrast between what labor is and what labor might be helps generate his anger at contemporary economic arrangements. In his account of alienation in work, Marx describes productive activity in capitalist society as inappropriately separating the individual from their product, from their productive activity, from other individuals, and from their own nature. In treating essential human characteristics solely as means to profit or survival, the productive activity characteristic of capitalism transforms workers into mentally and physically dehumanized beings.

5. Vision of Socialism

Marx's work also contains a vision of what the socialist epoch beyond capitalism (which I will call communism) will look like, together with an account of how the transition to that society might be effected.

Marx's views on this transition are reasonably clear. He thinks of communism as having "objective" and "subjective" preconditions, and of the evolution of capitalism as providing for both of these. "Objectively," communism requires that the productive forces have reached a considerable level of development, and it is capitalism's historical vocation to develop productivity up to that point. "Subjectively," communism requires the existence of a historical agent with the power and motivation to bring it about, and capitalism creates and organizes the proletariat in that role. It is in the struggle against capital that organized workers, who form an increasing majority of the population, come to favor communism, and to develop the organization and skills needed to take political power and transform society. Some strategic issues remain, but Marx's various views are usually readily discernable. (To give some examples: he opposes communists adopting conspiratorial modes of organization; he encourages their engagement with everyday economic and political struggles; he maintains that violence is not an inevitable part of revolutionary change; and so on.)

Marx's vision of what communism might look like is much less clear. He says little on this topic, and that reticence is deliberate—he did not believe that socialists should devote time and energy to delineating the social and political arrangements of the future. I will say something about both the little that he does venture and the reasons for his reluctance.

Marx's infrequent and insubstantial remarks on this subject suggest that communist society will consist of (at least) two broad historical stages. (These two stages of

communist society are sometimes preceded by a temporary, extra-constitutional but democratic, provisional government that he calls the "dictatorship of the proletariat.") These two historical stages might be distinguished according to both their economic and political arrangements. Communism will initially be marked "economically, morally, and intellectually" by its emergence from capitalist society, but in the fullness of time its arrangements will no longer bear the "birth marks of the old society from whose womb it emerges" (Marx 2000: 614). The movement from one to the other of these communist stages involves an evolutionary rather than a ruptural transformation.

Rather schematically, we might distinguish the economic arrangements of the two stages by their dominant distributive principles. The first stage of communism is a form of society in which productive activity would be rewarded according to "the contribution principle," whereby those who contribute more to production receive a proportionately greater income than those who contribute less. In the second "higher" stage of communism distribution will be according to "the needs principle" which breaks that motivational linkage between contribution and reward. In this stage, individuals are sufficiently motivated by something other than their own pecuniary reward (perhaps moral duty or pleasure in self-realization) such that, although they contribute according to their "ability," they are content to be rewarded according to their "needs."

Rather schematically (and less authoritatively), the political arrangements of the two stages might be distinguished as follows. In the first stage of communism political arrangements would be (further) democratized, decentralized, and "deprofessionalized" (with bureaucrats and professional politicians being replaced by ordinary citizens). "Freedom," Marx suggests, "consists in converting the state from an organ superimposed upon society into one completely subordinate to it" (Marx 2000: 610). That polity would, however, still utilize organized coercion (not least, to defend communist society against its internal and external enemies). In the second "higher" stage of communism, however, it seems that this organized coercion would eventually prove unnecessary. Precisely how this might be possible—at least, how it might be possible without implausibly assuming some spontaneous harmony of desires—is not obvious. One tentative suggestion emphasizes the role of democratic commitment of some sort, making the use (or even threat) of coercion redundant.

In short, communist society will be marked by an evolution whereby: a democratic, decentralized, and "deprofessionalized" polity gradually ceases to play any coercive role; and distribution according to the contribution principle gradually gives way to distribution according to the needs principle. Sketched in that brief and abstract manner it is hard to know what to think about the desirability and feasibility of these predicted developments. Yet Marx never attempts to elaborate this picture, to put serious sociological and institutional flesh on these brief and abstract bones.

This looks puzzling. Marx accepts the need for socialists to possess a detailed and persuasive account of what is wrong with modern society, but denies the need for similarly detailed and persuasive accounts of what its socialist replacement might look like. For example, he criticizes utopian socialists (such as Charles Fourier), less for the inadequacy and implausibility of their blueprints of the future, than for supposing that we need blueprints at all (Leopold 2005).

I can identify three main lines of Marxian argument against describing (in any serious detail) the social and political arrangements of the future. None of them strikes me as persuasive. First, utopian blueprints are said to be *undemocratic*; that is, they somehow pre-empt the freedom of future generations to decide their own arrangements.

This seems doubtful. My describing an ideal society does not prevent others from building a different one (indeed, it might help them decide what it is that they want to build). Second, utopian blueprints are said to be *impossible* to draw with complete accuracy; the impact of contingent and unforeseen events means that we cannot now plan precisely what we will need to do in the future. This might be true, but scarcely entails that we should reject all plans. After all, flexible and less than perfectly accurate plans might still offer useful guidance. Third, utopian blueprints are said to be *unnecessary*; that is, socialists do not need to design solutions to social problems but should rather facilitate the delivery of solutions already developing within the womb of the old society (to adopt Marx's often obstetric language). This confidence in the right solutions being generated by the historical process (unaided by individuals engaging with questions of socialist design) strikes me as implausible (not least, given the twentieth century) and unwarranted (not least, by Marx's own theory). My own view is that Marx's position risks deleterious practical consequences by discouraging serious reflection on our ends, and on the social and political arrangements that might best embody them.

6. Concluding Remarks

Having sketched Marx's views on these four topics—theory of history, assessment of capitalism, philosophical anthropology, and vision of socialism—I conclude with comments on some absences from that list, and on the relation between its constituent items.

Marx writes about many subjects not discussed here. Some of these omissions are due to space; otherwise much more might have been said, for example, about class or ideology. Other omissions are due to the remit of this volume; otherwise much more might have been said, for example, about the cyclical analysis of capitalist crisis or the falling rate of profit. However, some subjects were omitted—perhaps controversially—because I doubt that they are to be counted amongst Marx's most important legacies. For example, some will regret the omission of the "dialectical" method of inquiry from the list. Yet I consider Marx's investigative tools—as opposed to his substantive concerns and normative commitments—as not amongst the more distinctive elements of his work. He thought of his own work as open to evaluation by conventional standards of scientific adequacy, and an examination of his actual historical and empirical methods of inquiry (as distinct from overt methodological comments) suggests that these are less "esoteric" than often thought (Little 1986: 92–126). This need not entail a rejection of all "dialectical" talk. Dialectics can also refer to: a form of presentation; a pattern of change (in nature, society, and consciousness); and a theory of social contradictions (including the "real contradiction," which Jon Elster calls "counterfinality," and which involves unintended and deleterious consequences resulting from something like the fallacy of composition) (Elster 1985 : 36–48).

Finally, there is the relation between Marx's views on these various topics. In particular, I want to resist one familiar way of thinking about the systematic character of Marx's work (which can take positive and negative forms). Consider Lenin's (positive) description of Marxism as "a block of steel," a metaphor surely suggesting that Marx's ideas as a whole form a unified system impervious to criticism. This strikes me as entirely the wrong way to think about his work. With each of the subjects discussed above there remain questions about the meaning, coherence, and truth of Marx's views. Those concerns scarcely diminish, let alone disappear, when these views are put together. This is

not to deny that there might be some important and interesting links here. It is to deny that when put together these views come to form a larger theoretical unity which stands and falls with any of its component elements. There is, for example, no incoherence in doubting the desirability and feasibility of Marx's vision of socialism, and yet seeing him as an acute and profound critic of capitalism (Wolff 2002: 126).

Marx has many interesting, insightful, and relevant, things to say about significant issues in political and social philosophy. These include, but are not exhausted by, his theory of history, his account of capitalism, his philosophical anthropology, and his vision of socialism. Whilst there might be some (complex and contestable) connections between these various views, it is a mistake to think of them together as forming a theoretical unity which has to be either swallowed or rejected in its entirety.

Related Topics

Hegel, Marxism and Contemporary Political Thought

References

Berman, Marshall (1999) *Adventures in Marxism*, London: Verso.
Cohen, G.A. (1978) *Karl Marx's Theory of History. A Defence*, Oxford: Clarendon Press.
Cohen, G.A. (1988) *History Labour and Freedom. Themes From Marx*, Oxford: Clarendon Press.
Elster, Jon (1985) *Making Sense of Marx*, Cambridge: Cambridge University Press.
Geras, Norman (1983) *Marx and Human Nature: Refutation of a Legend*, London: Verso.
Hardimon, Michael O. (1994) *Hegel's Social Philosophy*, Cambridge: Cambridge University Press.
Leopold, David (2005) "The Structure of Marx and Engels' Considered Account of Utopian Socialism," *History of Political Thought* 26/3: 443–466.
Leopold, David (2007) *The Young Karl Marx. German Philosophy, Modern Politics, and Human Flourishing*, Cambridge: Cambridge University Press.
Little, Daniel (1986) *The Scientific Marx*, Minneapolis: University of Minnesota Press.
Marx, Karl (2000) *Selected Writings*, edited by David McLellan, second edition, Oxford: Oxford University Press.
McMurtry, John (1978) *The Structure of Marx's World View*, Princeton, NJ: Princeton University Press.
Wolff, Jonathan (2002) *Why Read Marx Today?* Oxford: Oxford University Press.
Wood, Allen W. (2004) *Karl Marx*, second edition, London: Routledge.

Further Reading

A stimulating intellectual biography of Marx is provided in Jerrold Seigel, *Marx's Fate: The Shape of a Life* (Princeton: Princeton University Press, 1978). A highly illuminating account of Marx's literary and cultural interests can be found in S.S. Prawer, *Karl Marx and World Literature* (Oxford: Clarendon Press, 1976). For some sense of the economic dimension of Marx's thought, the reader might profitably consult Samuel Hollander, *The Economics of Karl Marx. Analysis and Application* (Cambridge: Cambridge University Press, 2008). G.A. Cohen's *Karl Marx's Theory of History. A Defence* (Oxford: Clarendon Press, 1978) is a brilliant work which illuminates a wealth of issues beyond its ostensible subject matter. There is much interesting material about Marx's political thought in the two-volume study by Richard N. Hunt, *The Political Ideals of Marx and Engels* (Pittsburgh: University of Pittsburgh Press, 1974, 1984). The political dimension of Marx's earlier writings is covered in David Leopold, *The Young Karl Marx. German Philosophy, Modern Politics, and Human Flourishing* (Cambridge: Cambridge University Press, 2007). Finally, for a wide-ranging, thoughtful, and sometimes controversial, account of Marx's philosophical views, Allen W. Wood, *Karl Marx*, second edition (London: Routledge, 2004) can be recommended.

14
LATE NINETEENTH- AND EARLY TWENTIETH-CENTURY BRITISH THOUGHT

Maria Dimova-Cookson

Key features of late nineteenth- and early twentieth-century British thought include analysis of the nature of liberty, keen interest in the role of the state in creating conditions for personal development, and belief either in perfectibility of human beings or in social progress, often in both. In spite of significant disagreements on these issues, the main representative thinkers of this period—here we focus on Herbert Spencer, T.H. Green, Bernard Bosanquet and L.T. Hobhouse—believed in systematic studies of human nature and society, drawing on a range of disciplines in humanities and natural sciences. Also, all of them believed in a link between morality and politics. Unlike the liberal political theorists of the second half of the twentieth century, these Victorian and Edwardian thinkers, all passionate in defending liberty, were not moral pluralists.

1. Intellectual Background

British political thought of this period developed against the backdrop of Adam Smith's political economy and Bentham's utilitarianism, Darwinian and Lamarckian evolutionism, J.S. Mill's moral and political philosophy and the influx as well as domestic generation of socialist ideas. Smith's and Bentham's belief in free trade and free interplay of individual activity was representative of late eighteenth- and early nineteenth-century liberalism that battled against the vested interest of the landed aristocracy and the remnants of feudalism. The main political expression of this battle was the condemnation of governmental interference. Economic liberalism as displayed in the politics of *laissez-faire* was at the heart of the process of political liberalization and it was this kind of liberalism that found its most prominent philosophical defense in the works of Herbert Spencer (1820–1903). However, the late nineteenth century witnessed a significant departure from the liberalism aiming at limited government towards a "new liberalism" advocating substantial state involvement in social and economic life. Most late-Victorian liberals fought, rather successfully, to dissociate liberalism from its "ethos of egoistic possessive individualism" (Bellamy 1990: 2).

The ideas of the British thinkers of this period were shaped to a remarkable extent by developments in biology: Darwin's theory of natural selection and Lamarck's theory of evolution. The impact of these advancements of biological science was twofold. On the one hand, they provided political thinkers such as Spencer with a scientific method in justifying their vision of a perfect society (Gray 1990). In the case of Spencer, the theory of evolution was put to the service of defending individualism and limited state intervention. On the other hand, biological and evolutionary theory led to the grounding of diametrically opposed ideas, in favor of state intervention, in the case of the new liberals L.T. Hobhouse (1864–1929), J.A. Hobson (1858–1940) and D.G. Ritchie (1853–1903), who used the idea of "organicism" as a model of social growth and thus managed to reconcile Idealism with biological evolution (Freeden 1976). During the late nineteenth and early twentieth centuries a number of disciplines, related to social philosophy, developed rapidly: not just political economy and biology but also jurisprudence, psychology, anthropology and history. All these disciplines provided quite diverging, yet previously unutilized, methods for the study of morality and social progress.

For all British thinkers of this period J.S. Mill was a man of a previous generation but his influence on them could hardly be overstated. Not only Mill's theory of liberty but also his understanding of individuality as an agency of talent, development, experimentation and thirst for knowledge and social utility, had a great impact on their metaphysics and political visions. Ernest Barker claims that the wide spectrum of meanings inherent in Mill's understanding of liberty anticipated the movement from old to new liberalism.

> From a conception of liberty as external freedom of action, necessary for the discovery and pursuit of his material interest by each individual, Mill rose to the conception of liberty as free play for that spiritual originality, with all its results in "individual rigour and manifold diversity," which alone can constitute a rich, balanced and developed society.
>
> (Barker 1951: 3)

Well before Green (1836–82) and Bosanquet (1848–1923), who were the most eloquent exponents of positive liberty, Mill outlined the spiritual and moral personality on which this concept was based. Mill, however, did not pursue the political implications of the different aspects of his concept of liberty: he did not develop a notion of rights, like Spencer; did not ascribe significant function to the state in providing the conditions for liberty, like Green and Hobhouse; and, arguably, he did not give due weight to the role of society in the constitution of personal flourishing, like Bosanquet. Interestingly, Mill's utilitarianism was not well received in this period. Spencer rejected it as it ascribed too much importance to the general wellbeing and thus carried the potential to undermine personal freedom. Green and Bosanquet believed that it was based on a false metaphysical understanding of the person: according to them human beings were not fundamentally motivated by desires but by ideas of moral wellbeing. During the late nineteenth century Henry Sidgwick (1838–1900) was the only high-profile exponent of Mill's utilitarianism.

The influx of socialist ideas in Britain after 1880 went hand in hand with the turning of the liberal tide away from *laissez-faire* towards advocacy of ever-increasing state intervention in internal affairs. Two types of socialism made their way: the revolutionary variant of Henry Hyndman (1842–1921) that took on board the ideas of class-war

and a socialist regime, and the reformist socialism of the Fabian Society, founded in 1884, that preached gradual change within existing institutions, through constitutional and democratic means. Aiming to influence university audiences, the Fabians (Sidney and Beatrice Webb, H.G. Wells and G.B. Shaw) succeeded in making socialism intellectually and politically acceptable. Their ideology played a key role in the formation of the UK Labour Party during the twentieth century. Socialist ideas impacted significantly, although with varying success, the political thinkers of this period. Bosanquet, who could easily qualify as the most anti-individualist British thinker of all times, had certain reservations about the economic and cultural claims of the working class. L.T. Hobhouse, on the other hand, embraced ideas of equality and redistribution to an extent that raised the question why the "socialist" label has not been more firmly attached to him. Yet as a whole, however receptive of socialist ideas, British Victorian and Edwardian thinkers resisted allowing the state to interfere in any way with the spontaneous process of formation of one's moral and political ideals.

2. Spencer

Herbert Spencer stood out from the other main late-nineteenth-century British thinkers in a number of ways. He did not have a conventional schooling and he did not pursue a university degree because of his outright rejection of school discipline and his refusal to study Greek and Latin under the private tutorship of his academically distinguished uncle. In his youth he worked for a number of railway companies taking part in station design, track building and general surveying—an experience that gave him direct insight into the nature of the Industrial Revolution (Harris 2004). His writing career started with his political "letters" to the *Nonconformist*, subsequently published together as a pamphlet (*The Proper Sphere of Government*, 1843) and consolidated when he became a sub-editor of the *Economist*. During this period he wrote and published *Social Statics* (1851), where he spelled out the moral principles from which the right role of the state could be derived. The inheritance he received after his uncle's death in 1853 allowed him to devote all his time to his own writing and in subsequent years he produced his *System of Synthetic Philosophy* which contained ten volumes on the principles of biology, psychology and sociology. In the late 1860s he was invited to apply for two major professorial chairs: in mental philosophy in University College, London and moral philosophy in Edinburgh. These offers were declined but they demonstrated the recognition he had achieved in academic circles.

There are two key elements in Spencer's social philosophy that can be seen as inconsistent with each other, but which can ultimately be understood only in each other's context: his individualistic understanding of liberty and his concern with social perfection. His reading of liberty as minimization of any external intrusion went hand in hand with belief in society as a perfectly functioning organism. In his early career he argued for extending the franchise to working-class men and for equal treatment of women, although in his later life he changed his views. The fusion of his political individualism and social perfectionism can be explained against the backdrop of the mid-nineteenth century, where the introduction of free trade in 1846 and the collapse of organized Chartism in 1848 bore witness to the benefits of economic prosperity based on capitalist competition.

Spencer's reputation as a "social Darwinian" also has to be understood in the context of his search for a theory of the perfect society. While Darwin was only marginally inter-

ested in an evolutionary process, focusing more narrowly on the process of natural selection, Spencer, following Lamarck, was intrigued by the inheritance of selected characteristics. Spencer's coined phrase "survival of the fittest" had an evaluative meaning: unlike the Darwinian natural selection, it had moral implications. Spencer's evolutionary mechanism appears to guarantee "directional, indeed progressive change" (Offer 2001: xviii). What is more, Spencer's ideas of social evolution were published prior to 1859, when Darwin's *Origin of the Species* appeared, and Darwin acknowledged the influence of Spencer on the formation of his own theory. However, a look into some of Spencer's more specific political recommendations will help us appreciate the extent to which the parallel with Darwin's natural scientific approach carries certain weight. By 1891, Spencer had started to argue that not only the state but also charitable organizations should abstain from alleviating poverty as any interference would produce false incentives, dependence, lack of responsibility and would interfere with the survival of the fittest. In Spencer's view, a key feature of the members of every species, including human beings, was their "fitness for life" which developed only in a natural process of spontaneous interaction (2001 [1884]: 128).

Spencer's arguments and "sentiments" on liberty were outdated by 1884 when *The Man Versus the State* came out. He was well aware of the changing nature of liberalism during the second half of the nineteenth century and described the contents of these changes as well as their advocates, T.H. Green and Hobhouse, though in rather resentful terms. He pitched his understanding of liberty against two evils: present and past. For him the bygone era of feudalism, war and protected aristocratic interest was as bad as the coming era of the "socialistic" state with its ever-increasing legislation, regulation and state bureaucracy. He used two related distinctions—voluntary versus compulsory cooperation and "regime of contract" versus "regime of status"—where voluntary cooperation and "regime of contract" explained the nature of authentic liberty, while their opposites were symbolic of the past and present threats to liberty (1997 [1891]: 142–3). The type of freedom most conducive to a well-ordered society is personal freedom mindful of the personal freedom of others. Growing state power is counterproductive to liberty for several reasons. First, it is based on the false philosophy that "by due skill an ill-working humanity can be framed into well-working institutions" (2001 [1884]: 105). Second, its advocates held the wrong belief that only illegitimate authorities violate liberty. For Spencer, neither the legitimacy of the political powers nor the nature of their intentions—altruistic or evil—mitigated the claim that state authority violates freedom. Spencer also developed a link between "socialism" and "slavery" in an argument anticipating Nozick's parallel between taxation and forced labor. Spencer claimed that not only a full subordination of one man to another amounted to slavery, but that even a partial coercion had to be seen as a form of enslavement. In a state with increased social regulation people had to contribute some of their time or labor to the welfare of others. And even if this represented a small proportion of their time or income, the intrusion into personal liberty implied in this process was as unjustifiable as slavery. While the mid- and late-twentieth-century libertarians strongly rejected any metaphysical foundations to theories of individuality or society, Spencer took very seriously natural law theories and made a case for the ontological status of natural rights. Unlike Bentham, he thought that they were not fiction but actually existed, and unlike Green and the new liberals, he thought that their existence was independent of social recognition.

Some of Spencer's arguments carried more persuasive power than others. He drew a distinction between family ethics and state ethics, where the first was based on the

principle of generosity and the second on the principle of merit, and observed that the wider community would not be able to function if it was run on the basis of family ethics. He also pointed out that both individual and social life "imply maintenance of the natural relation between efforts and benefits" (2001 [1884]: 164). His argument against free general education, though, was less impressive. He claimed that education in syntax, geography or history would have no benefit for the political maturity of citizens, in the same way that skills in cricket could not help someone to play the violin. To Spencer's credit, we should say that he was well acquainted with the arguments of his opponents. As with most key ideological debates, it all came down to weighing up two evils against each other. In this case the evil he feared more was the state's interference with what he believed was a natural process of developing individuals' economic and moral stamina, and the evil he was aware of but feared less was the poverty and economic destitution of the larger part of the population. One of his ways of expressing this dilemma was through the question "is it not cruel to increase the suffering of the better that the suffering of the worse may be decreased?" (2001 [1884]: 135). This was a tough question but one that T.H. Green was uniquely positioned to tackle well. The British idealist's understanding of human nature would provide the resources for explaining why the suffering of the better differed categorically from the suffering of the worse, as the former had the capacity to produce social benefits, while the latter had not.

3. Green

Green was a leading if not *the* leading figure of New Liberalism—his theory of "true" or "positive" freedom offered the philosophical and ethical foundations of the late-Victorian liberal critique of mass poverty and social destitution that accompanied the nineteenth-century industrialization and urbanization. Although the level of state intervention he recommended was rather qualified, a consistent application of his ideas, as demonstrated by Hobhouse, led to the justification of a much wider state involvement in economic and social infrastructures. Green founded the school of British idealism in Oxford and was one of the most influential philosophers in the period between 1880 and 1914. He is acclaimed for having succeeded in utilizing the resources of German idealism, typically associated with conservative politics, in the service of a left-wing Liberal Party program.

Green was educated at Rugby and then studied classics, philosophy, law and modern history as an undergraduate at Balliol College, Oxford. After graduation, Green was elected to a College Fellowship and then appointed as a tutor at Balliol, teaching ancient and modern history and philosophy. In 1878 he became White's Professor of Moral Philosophy. He was also an active member of the Liberal Party and campaigned for legal protection of workers and land tenants' rights, for extending education and the franchise, and for temperance reform.

Green's moral philosophy, as developed in his main, posthumously published, philosophical work *Prolegomena to Ethics* (1883), had strong Aristotelian and Kantian influences. He combined belief in the developmental and social nature of human agency with an account of morality that centered on the "good will" as opposed to either pleasure or outcomes. Like Kant's, Green's moral theory built on an opposition between a natural and a moral world. For this reason he differed from Spencer and idealists, such as D.G. Ritchie, who pursued a constructive link between philosophy and natural sciences. For Green, human agency had to be explained not in terms of desires but motives which

are constituted by an act of self-consciousness. Human conduct consists of rational self-direction towards an object in which the person seeks self-satisfaction. Self-satisfaction, in turn, can be found in the joint pursuit of fulfillment of one's capabilities and engagement with socially beneficial activities.

Green's theory of the common good builds on and substantiates his understanding of agency and his moral theory. The moral ideal in which self-realization can be found has, as mentioned, two significant aspects: the ever-fuller deployment of one's capabilities and the commitment to the common good. Each of these is linked to a fundamental trait of human agency: the first being its developmental nature and the second, its social nature. Human nature is fully revealed only when its intrinsic sociability comes to some form of fruition. Practically, service to the common good could manifest itself in various ways, ranging from raising a family, to being a good citizen or becoming a social reformer. A key insight of Green's common good theory is that the personal good (personal flourishing), properly understood, always incorporated the common good (service to community).

Green's theory of the common good was a key ingredient of his political philosophy as developed in his *Lectures on the Principles of Political Obligation* (1886), *Lecture on "Liberal Legislation and Freedom of Contract"* (1881), and *On the Different Senses of "Freedom" as Applied to the Will and to the Moral Progress of Man* (1886). It allowed him to make significant advances on Hobbes, Locke and Rousseau in the analysis of political obligation, rights and freedom. His critiques of natural rights and freedom of contract anticipated late-twentieth-century liberal theory that was critical of Hayek's and Berlin's negative libertarianism. Perfectionist/autonomy-based accounts of liberty and the rights recognition theory of recent decades build on Green's scholarship.

For Green, political obligation was justified on the grounds that it created the conditions for the fulfillment of man's vocation as a moral being. We owe obedience to a sovereign, the state or a political superior because their rule should contribute to establishing the institutions that enable the performance of our moral duties. However, the duty to obey the law is not equivalent to a moral duty. Moral duties cannot be imposed by the law because they are "duties to act *from certain dispositions and with certain motives*" (1986a [1886]: §10) which cannot be enforced as a matter of principle. The laws can control only our outward or external actions and thus indirectly encourage moral behavior, by promoting the conditions favorable to moral life. Therefore, the state should not regulate religious faith, create conditions that hinder self-reliance or moral autonomy, or set up institutions that take away the opportunity for the exercise of moral duty, such as the Poor Law, for example. Green seemed to have softened his stance on the latter, when in the context of his discussion of the reform in liberal legislation he recommended restrictions in the sale of alcohol. There he advanced the idea that a good law could act as "a powerful friend" (1986b [1881]: 203).

Green was critical of Hobbes and Locke who saw the justification of political obligation in the protection of natural rights. This critique prompted him to develop a new theory of rights that had been acclaimed as one of the "finest" theories of rights to date (Martin 2001: 49). Green challenged natural rights theories by arguing that rights existed as given to us by society. The reason society recognized the rights of its individuals was that thus it granted them powers that were necessary to their development as moral agents. Only by possessing certain rights could individuals become free contributors to the common good. The process of recognizing rights was a two-way process. Those deprived of rights—Green gave slaves as an example—had to demonstrate their

ability to adopt the values of the existing community of freemen and show a promise that, once they had rights, they would contribute to the wellbeing of this community. The freemen, in turn, by recognizing rights to former slaves made a step forward in the moral progress of mankind, progress marked by opening the existing civil institutions for the inclusion of those who were hitherto excluded from them.

The idea that rights were significant not only to the rights claimants (for example, the slaves), but also to those by whom the recognition of rights was actually fulfilled (in this case, the freemen), also found expression in Green's theory of positive freedom. Like Mill, Green saw liberty as something over and above "lack of external restraints." However, unlike him, Green did not conflate a variety of aspects of freedom in one all-encompassing concept but detected a principal difference between two types of freedom: "juristic" and "true." Juristic freedom was linked to the desire to act according to preference while true freedom could be found in the types of action where one lived up to his moral nature. Here Green aligned himself with thinkers like Plato, the Stoics, St Paul, Kant and Hegel who had understood freedom differently from "exemption of compulsion": either as a form of "fulfilment of the law of our being" or as "the condition of a citizen of a civilised state" (1986c [1886]: §§1, 17). Green demonstrated the practical application of true freedom in a political context when in his *Lecture on "Liberal Legislation and Freedom of Contract"* he explained the nature of "freedom" that should guide decisions on the legal reform of free contract. There Green introduced the term "positive" freedom, which was the freedom sought and found in activities that first, carried certain value, and second, were beneficial to everybody concerned. We should not seek freedom in activities that are detrimental to others. Free contract is notorious for its adverse effect on those with impeded bargaining powers, therefore its "freedom" should not be paradigmatic: a different, socially minded type of freedom, that is, positive freedom, should inform our political reasoning.

4. Intellectual Background of British Idealism

The rise of British idealism at the end of the nineteenth century, strongly connected to Green's academic influence, was not an isolated event. It was facilitated by an earlier development of cultural studies (S.T. Coleridge, 1772–1834), a concurrent rise of biblical studies through the influence of Benjamin Jowett (1817–13) among others, who taught both T.H. Green and Edward Caird, and the increasing impact of German idealism through the ideas of Kant, Hegel, Fichte and Lotze (Gaus and Sweet 2001). Richter (1964) argued that the main reason British idealism succeeded in occupying such a prominent position was that the philosophy of its key figures managed to soothe the crisis of the Christian faith—a crisis caused by the undermining impact of empiricism and Darwin's theory of evolution. The British idealism of T.H. Green and Bosanquet offered a secular foundation for the Christian moral values. There is more to that, however. Through their concept of true or positive liberty, and through Green's theory of rights based on social inclusion (Dimova-Cookson 2011) they found ways to address the problems of poverty and social deprivation that were the by-product of the Industrial Revolution and its underlying liberal individualism.

The two *Fin de Siécle* thinkers who most fully internalized Green's ideas and took them to further heights, even though in a rather divergent fashion, were Bernard Bosanquet and L.T. Hobhouse. Bosanquet developed further the personal metaphysics and moral and social philosophy of his idealist predecessor, while Hobhouse spelled out

in rather specific terms the political implications of Green's ideas on liberty, rights and political obligation.

5. Bosanquet

Bosanquet studied at Balliol College Oxford between 1867 and 1870. Upon graduation he was elected to a Fellowship at University College, Oxford. In 1881, with the inheritance from his father, he went to London where he got involved in adult education and social work through the Charity Organisation Society. At the age of 55 he became Professor in Moral Philosophy at the University of St Andrews but resigned this position after five years due to ill health. He was elected Gifford Lecturer for 1911–12 at the University of Edinburgh. Most of his philosophical research, however, was done outside the academic establishment, while his involvement with social work and social policy remained one of his priorities. Bosanquet's main work in social and political theory, *The Philosophical Theory of the State*, was published in 1899 and was quickly acknowledged as a classic statement of the British idealist view of politics (Nicholson 1990), while his Gifford lectures *The Principle of Individuality and Value* and *The Value and Destiny of the Individual* provided the most developed statement of his metaphysical views. Bosanquet's best known political ideas include his reading of social institutions as "ethical ideas," his particular theory of the general will (see Nicholson 1990), and his recommendation of "moral socialism" as against "economic socialism." He has been both criticized and acclaimed for having pushed Green's idealism away from its liberal and radical implications towards a very unusual mixture of collectivist, socially conservative and economically individualist ideas (Gaus 2001). However, Bosanquet had the time and the philosophical skills to elaborate further Green's personal metaphysics and moral philosophy and thus made its applications in social analysis more visible. Although the younger idealist made the collectivist aspects of idealism more explicit and thus made idealism in general less palatable throughout the greater part of the twentieth century, uncovering the inner mechanism of Green's morality has significant explanatory benefits. Also Green's untimely death in 1882 did not allow him to engage with the increasingly prominent socialist ideas in the period thereafter. The rise of socialist ideas triggered discussions on the nature of justice. It was Bosanquet who applied the idealist metaphysics on the analysis of justice and thus made the idealist legacy in this field more distinct.

Bosanquet's theory of justice as developed in his *Social and International Ideals* (1917) articulated moral dilemmas that anticipated Rawls's pivotal opposition between the "right" and the "good." Late-twentieth-century liberal thought proclaimed the distinction between personal "inviolability founded on justice," on the one hand, and the "welfare of society," on the other hand, as crucial for the process for protecting human rights (Rawls 1999: 3). Although the British idealists were acclaimed for overcoming dualities, such as that between justice and virtue, for example (Boucher 1997), Bosanquet, in many ways, defied this characterization. Like Green, Bosanquet saw continuity between the personal and the moral good, between the individual and social wellbeing, but he spelled out their points of conflict more clearly. As a result, he explained more precisely than Green did, the limits of state intervention: not only on the grounds that the state should not interfere with the spontaneous process of personal moral growth, but also because demands for justice can and do conflict with demands for public service.

Bosanquet's more specific argument was that the working-class claims for "their own function, their own culture, their own duty, their own ambitions for a worthy and

influential life" (1917: 195) were, on analysis, controversial. As a matter of principle, claims were demands made by individuals or groups for certain services that had to be provided to them by the state. And these claims represented the exact opposite of moral duties where individuals sacrificed aspects of their personal wellbeing in the name of the common good. Moral action and personal flourishing—strongly intertwined in the context of idealist metaphysics—crucially depended on one's preparedness to subordinate the narrowly personal to the wellbeing of the community. When we make claims for justice we are not acting as moral agents—we are not subordinating the individual to the social but using society to support the individual. However, the social ideal of justice is "justified" indirectly—the satisfaction of people's demands for justice helps them as individuals, but they will subsequently sacrifice their individuality in the name of public welfare. For Bosanquet, as was the case with the British idealists in general, every justification boiled down either to the assertion of moral action, or to creating the conditions for moral action, the key ingredient of which was the subordination of the narrowly conceived individual interest to the public good. Late-twentieth-century liberalism adamantly opposed such subordination. Yet nowadays, many political thinkers, critical of the negative libertarianism of the cold war era, turn back to the political ideas of New Liberalism in their attempt to resolve tensions in contemporary political theory (Simhony and Weinstein 2001). The contemporary interest in New Liberalism breathes new life into the British idealist metaphysics and thus into Bosanquet's analysis of the particular dynamics between the individual and the social in the context of moral action.

6. Hobhouse

British Idealism and New Liberalism are acclaimed by many for their success in reconciling personal flourishing with social welfare, liberalism with state intervention. While Green and Bosanquet balanced these on a metaphysical level, Hobhouse, "the most sophisticated intellectual exponent" of New Liberalism, made a case for their interdependence on a political level (Meadowcroft 1994). Hobhouse studied at Marlborough and at Corpus Christi College at Oxford, graduating in 1887. He stayed at Oxford as a Fellow subsequently in Merton College and in Corpus Christi College, but after his critique of idealism in his first major philosophical work, *The Theory of Knowledge* (1896) was not received well, he moved to the career of journalist, writing for the *Manchester Guardian* (1897–1903) and for *The Tribune* (1905–07). During some of this period he also worked as a secretary to the Free Trade Union. In 1907 he accepted the newly created chair in sociology at London School of Economics and held this chair until his death in 1929, continuing to make substantial contributions to the *Manchester Guardian*. His book *Liberalism*, written in the winter of 1910–11 has been celebrated as "the best twentieth-century statement of Liberal ideals" (Collini 1979: 121).

For Hobhouse, liberty, properly understood, allowed one to exercise the full spectrum of his faculties and reflected the capacity for personal growth. The concept of liberty "rests on the nature of the 'good' itself" and is compatible with "discipline," "organisation" and "strenuous conviction as to what is true and just" (1994 [1911]: 63, 56). The belief that liberty can only be achieved through the process of personal development, implied that it was impossible to exercise it unless certain social and economic conditions were already available. Therefore, a well-specified type of state intervention was necessary to liberty. A careful distinction had to be made between state control that

"cramps the personal life and spiritual order, and the control that is aimed at securing the external and material conditions of their free and unimpeded development" (1994 [1911]: 71). On these grounds Hobhouse advocated the "right to work" and the right to a "living wage," arguing that they were just as valid as "the rights of person or property" (1994 [1911]: 76). He also made the case for redistribution of resources on the grounds of distinguishing between the social and individual factors of wealth. While everyone should receive the income associated with the individual factors, the wealth based on social factors should go into the public coffers. In addition, Hobhouse argued that relative inequalities of wealth are unacceptable, remaining true to his conviction that there was "a close connection in each instance between freedom and equality" (1994 [1911]: 68). Hobhouse's radicalism in terms of redistribution of wealth and his misgivings about social inequality surpassed Green's own recommendations for reform and came into conflict with Bosanquet's beliefs in economic self-independence. He was among very few political thinkers who went so deeply into a socialist terrain yet retained their "liberal" qualification, although some critics doubt his entitlement to it (Collini 1979: 96).

Hobhouse's belief in the continuity between liberalism and socialism explains two of his most deep-seated political convictions. He repeatedly stated that liberalism had to adapt its political expression in order to protect its fundamental values. Only by increasing the role of the state in the provision of economic conditions, that is, only by embracing collectivist and socialist policies, could society foster the realization of the liberal values of personal flourishing, independence and voluntary collaboration. But this in turn provided the parameters within which socialism was justifiable: socialism that lost sight of liberal values was misguided. Hence New Liberalism collided on many points with Fabianism. Hobhouse opposed the materialist socialism of Hyndman's Social Democratic Federation (1881–1911) and the allegedly bureaucratic socialism of the Fabians, promoting instead socialism that fostered liberty and was democratic in nature.

As one of the first British sociologists, Hobhouse attempted to bridge the gap between idealism and positivism. He embraced an organic view of social life based on the idea of an "orthogenic evolution" where progress was corollary to the work of mind and reason, was measured on ethical grounds, and reflected a process in which cooperation superseded competition (Freeden 2004). His liberalism and ethical socialism found expression in his anti-imperialist views. In *Democracy and Reaction* he argued for "forbearance in international affairs" (1909 [1904]: 12), and he praised liberal reformers like Cobden who lagged behind him on the level of the state intervention in domestic life but nonetheless firmly resisted Imperial aggrandizement (Meadowcroft 1994: 12).

7. The Matter of Gender

All four thinkers covered here embraced J.S. Mill's belief in gender equality. And although the late-nineteenth and early twentieth-century British thought was dominated by male thinkers, a number of women earned their place in the history of ideas. Beatrice Webb (1858–1943) developed under the tutelage of Spencer but rejected his individualism in order to lay the foundations, together with her husband Sidney Webb (1859–1947), of Fabianism: the most distinct brand of British socialism (Greenleaf 1983: 381). She wrote on the cooperative movement and trade unionism in Britain. Helen Bosanquet (1860–1925), Bernard Bosanquet's wife, had an active career as a theorist and publicist of the Charity Organisation Society. She wrote a number of

sociological works on poverty, the standard of life and the family. Her recommended strategies for alleviating poverty were based on strengthening the rational will through private charity organizations and not centralized state policies. After publishing her *Defence of Idealism* in 1917, Mary Sinclair (1863–1946) became the first woman member of the Aristotelian Society. She was also among the first British philosophers who took interest and wrote on psychoanalysis. Sylvia Pankhurst (1882–1960) set up the Suffragettes movement in 1914 which, over the years, evolved into the Workers' Socialist Federation. However, female participation in public life more often took the form of political activism rather than generation of political thought. This slow but visible process of women's involvement in politics was an expression of the moral and liberal values deeply held during the previous *Fin de Siécle*.

Related Topics

Aristotle's Social and Political Philosophy, Hobbes, Locke, Rousseau, Kant, Mill, Liberalism, Natural Law and Rights Theory, Libertarianism, Equality, Freedom, Rights

References

Barker, E. (1951) *Political Thought in England 1848 to 1914*, London: Oxford University Press.

Bellamy, R. (ed.) (1990) *Victorian Liberalism: Nineteenth-century Political Thought and Practice*, London: Routledge.

Bosanquet, B. (2001 [1899]) *The Philosophical Theory of the State and Related Essays*, South Bend, IN: St Augustine's Press.

Bosanquet, B. (1917) *Social and International Ideals*, London: Macmillan.

Boucher, D. (1997) "British Idealism and the just society," in Boucher, D. and Kelly, P. (eds.), *Social Justice: From Hume to Hayek*, London: Routledge.

Collini, S. (1979) *Liberalism & Sociology. L.T. Hobhouse and Political Argument in England 1880–1914*, Cambridge: Cambridge University Press.

Dimova-Cookson, M. (2011) "Justice as a secondary moral ideal: The British idealists and the personal ethics perspective in understanding social justice," *European Journal of Political Theory* 10: 46–70.

Freeden, M. (1976) "Biological and evolutionary roots of New Liberalism in England," *Political Theory* 4: 471–90.

—— (2004) "Hobhouse, Leonard Trelawny (1864–1929)," *Oxford Dictionary of National Biography*, Oxford University Press; online edn, May 2006 [http://www.oxforddnb.com/view/article/33906, accessed 11 Jan 2011]

Gaus, G. (2001) "Bosanquet's communitarian defence of economic individualism: a lesson in the complexities of political theory," in Simhony, A. and Weinstein, D. (eds.), *The New Liberalism: Reconciling Liberty and Community*, Cambridge: Cambridge University Press.

—— and Sweet, W. (2001) "Editors' Introduction," in Bosanquet, B., *The Philosophical Theory of the State and Related Essays*, South Bend, IN: St Augustine's Press.

Gray, T. (1990) "Herbert Spencer's Liberalism—from social statics to social dynamics," in Bellamy, R. (ed.), *Victorian Liberalism. Nineteenth-century Political Thought and Practice*, London: Routledge.

Green, T.H. (1890 [1883]) *Prolegomena to Ethics*, Oxford: Clarendon Press.

—— (1986a [1886]) *Lectures on the Principles of Political Obligation*, in Harris, P. and Morrow, J. (eds.), *Lectures on the Principles of Political Obligation and Other Writings*, Cambridge: Cambridge University Press.

—— (1986b [1881]) *Lecture on "Liberal Legislation and Freedom of Contract,"* in Harris, P. and Morrow, J. (eds.), *Lectures on the Principles of Political Obligation and Other Writings*, Cambridge: Cambridge University Press.

—— (1986c [1886]) *On the Different Senses of "Freedom" as Applied to the Will and to the Moral Progress of Man*, in Harris, P. and Morrow, J. (eds.), *Lectures on the Principles of Political Obligation and Other Writings*, Cambridge: Cambridge University Press.

Greenleaf, W.H. (1983) *The British Political Tradition, Volume Two: The Ideological Heritage*, London: Routledge.

Harris, J. (2004) "Spencer, Herbert (1820–1903)," *Oxford Dictionary of National Biography*, Oxford University Press; online edn, Sept. 2010 [http://www.oxforddnb.com/view/article/36208, accessed 13 Dec 2010].
Hobhouse, L.T. (1906 [1886]) *The Theory of Knowledge: A Contribution to Some Problems of Logic and Metaphysics*, London: Methuen and Co Ltd.
—— (1909 [1904]) *Democracy and Reaction*, London: T. Fisher Unwin.
—— (1994 [1911]) *Liberalism and Other Writings*, Cambridge: Cambridge University Press.
Martin, R. (2001) "T.H. Green on individual rights and the common good," in Simhony, A. and Weinstein, D. (eds.), *The New Liberalism: Reconciling Liberty with Community*, Cambridge: Cambridge University Press.
Meadowcroft, J. (1994) "Introduction," in Hobhouse, L.T., *Liberalism and Other Writings*, Cambridge: Cambridge University Press.
Nicholson, P. (1990) *The Political Philosophy of the British Idealists*, Cambridge: Cambridge University Press.
Offer, J. (2001) "Introduction," in Spencer, H., *Political Writings*, Cambridge: Cambridge University Press.
Rawls, J. (1999) *A Theory of Justice*, Oxford: Oxford University Press.
Richter, M. (1964) *The Politics of Conscience: T. H. Green and His Age*, London: Weidenfeld and Nicolson.
Simhony, A. and Weinstein, D. (eds.) (2001) *The New Liberalism: Reconciling Liberty and Community*, Cambridge: Cambridge University Press.
Spencer, H. (2001 [1884]) *The Man Versus the State* in *Political Writings*, Cambridge: Cambridge University Press.
—— (1997 [1891]) "From Freedom to Bondage," in Stapleton, J. (ed.), *Liberalism, Democracy and the State in Britain: Five Essays, 1862–1891*, Bristol: Thoemmes Press.

Further Reading

Barker, R. (1997) *Political Ideas in Modern Britain: In and After the Twentieth Century*, London: Routledge.
Carter, M. (2003) *T.H. Green and the Development of Ethical Socialism*, Exeter: Imprint Academic.
den Otter, S. (1996) *British Idealism and Social Explanation: A Study of Late Victorian Thought*, Oxford: Oxford University Press.
Dimova-Cookson, M. and Mander, W.J. (eds.) (2006) *T. H. Green Ethics, Metaphysics and Political Philosophy*, Oxford: Oxford University Press.
Stapleton, J. (1997) "Introduction," in Stapleton, J. (ed.), *Liberalism, Democracy and the State in Britain: Five Essays, 1862–1891*, Bristol: Thoemmes Press.
Sweet, W. (1997) *Idealism and Rights: The Social Ontology of Human Rights in the Political Thought of Bernard Bosanquet*, Lanham, MD: University Press of America.
Taylor, M.W. (1992) *Men Versus the State: Herbert Spencer and Late Victorian Individualism*, Oxford: Oxford University Press.
Tyler, C. (2010) *The Metaphysics of Self-realisation and Freedom. Part 1 of The Liberal Socialism of T.H. Green*, Exeter: Imprint Academic.
Vincent, A. and Plant, R. (1984) *Philosophy, Politics and Citizenship: The Life and Thought of the British Idealists*, Oxford: Oxford University Press.

15
CONTINENTAL POLITICAL PHILOSOPHY

James Bohman

1. Introduction

While Continental philosophy has said much about the nature of politics and modern political institutions, it has not produced much political philosophy in the standard sense of evaluating policies or justifying institutions. That is, it has not been concerned with elaborating the principles necessary to justify the existing political and legal institutions, so much as examining and criticizing the very presuppositions of modern politics and society. This stance derives from the fact that the decisive influences have been two philosophers whose politics must be considered radical: Karl Marx and Martin Heidegger. Both rejected modern political institutions as currently constituted and sought to provide the basis for a new kind of politics. For Marx it would be a politics freed from the dynamics of class domination and rule, while for Heidegger it would be a new way of "dwelling on the earth," freed from the unthought presuppositions of Western metaphysics and rationalism. Such a critique of politics and political thinking is radical in the sense that it "goes to the root" not just of political institutions and their underlying norms, but to the very way in which modern society is organized, to its deepest structures of thought and its dynamic historical process. In both philosophers, there is a sense that something is deeply wrong with a whole complex set of ideas and practices that we might more generally call "modernity," so that, under their influence, political thinking in Continental philosophy concerns the question of what is to be negated and preserved in modern ideas of reason and politics. A range of different positions on "modernity" as a philosophical problem still characterizes Continental political theory to this day, from Marx to Heidegger and Habermas to Derrida. For many different reasons, however, this radicalness about politics has become increasingly difficult to sustain, and much of Continental philosophy's recent history represents a challenge to the Marxist and Heideggerian influence on political philosophy.

The highest political expression of modernity is found in the Enlightenment, so much so that challenges to the constitutive elements of modern reason and politics are challenges to Enlightenment thinking. The Enlightenment proclaimed modernity to be a new time (as the German word for modernity, *Neuzeit*, expresses), in which the superstition, authority and biases of the past would be replaced by the rational practices of science and democracy (Habermas 1987: ch. 1). Just as the sciences disenchanted nature and thus gave human beings power over it, democracy would permit them to overcome submission to external authority and thus to take control over their own lives. In the

eyes of the Enlightenment, modernity is a project, with a negative and destructive side as well as a positive and constructive side, both of which are found in its underlying critical conception of rationality. The destructive side of critical reason is aimed at those practices and beliefs that enslave human beings. Indeed, this side is so much emphasized by the nineteenth century that Marx considered the proper program of philosophy to be "the ruthless criticism of everything existing." Even so, the constructive and not merely critical dimension of this Enlightenment philosophy is equally important, and in politics it is concerned with the creation and justification of new emergent practices and moral and political ideals, including the justification of basic ideals of democracy, such as self-rule, human rights and the standard of unforced agreement and consensus as the sole basis for legitimate institutions. The main problem for Enlightenment philosophy was to maintain the balance between the constructive and destructive sides of critical reason. Already by the early nineteenth century, some philosophers began to argue that the destructive side of reason had overwhelmed its constructive side, leading to a new level of reflection on the logic of modernity that Max Weber called the "disenchantment of disenchantment." This argument was particularly relevant to the reflections on the nature and scope of reason found in German Idealism.

For Kant, and even more so for Hegel, the Enlightenment was insufficiently self-critical, especially about its own conception of reason and its connection to human freedom. Ordinary human understanding (*Verstand*) is sufficient for the sciences to attempt to understand nature and for acting out of self-interest, but it is reason (*Vernunft*) that brings about unity in the sciences and supplies the laws of the moral domain, and it is in the public use of reason that "free and equal" citizens give themselves laws that could be agreed upon by everyone even when they are not in their self-interest. It is through the public use of reason alone that politics can become a medium for Enlightenment and can overcome what Kant saw as the potentially destructive "unsocial sociability" of mankind. Hegel took this notion of reason further and turned it into the basis for a more thorough critique of modern society. The dominant institutions of modern civil society and the liberal state divided modern society, based as they were on the aggregation of interests in markets and in political institutions that defend the "arbitrary freedom" of individuals. Hegel saw modern societies as unbalanced; while they had achieved a form of society based on the principle of universal freedom, they had done so in a one-sided and inadequate way. The ultimate role of the state in modern society was to provide the context for higher forms of freedom and unity needed to sustain a rational form of ethical life, a community based on freedom. However different they were in political vision, both Kant and Hegel saw politics as the means for realizing reason in history, in institutions that would restore the unity and integrity of social life while guaranteeing the greatest extent and highest form of freedom. Opposed to both of them, Marx would challenge this conception of the unity of modern society and with it the very idea of a political integration of modern society. Heidegger, too, would seek a more radical solution beyond modernity, seeing politics, briefly, as a means to go beyond, rather than fulfill, the promise of a rational society. Both call for a broader and more comprehensive form of politics.

This challenge to claims for a non-coercive and rational politics under current conditions, whether radically democratic in the case of Marx or anti-democratic in the case of Heidegger, shaped Continental political theory for some time and continues today in debates between modern and post-modern rejections of liberalism. In this chapter, I shall trace the development of this argument in Continental political theory, showing

how the examination of the rejection of Idealism has led to the return of politics into Continental philosophy. In the next section, I shall examine the two different reasons for rejecting politics advanced by Marx and Heidegger which led to radically modern and anti-modern conceptions. Both aim to break with existing modern institutions. On the one hand, Marx rejected modern political institutions precisely because they failed to live up to the modern principles of freedom and justice that they required for their legitimacy; on the other hand, Heidegger rejected both such institutions and their supporting principles, seeking instead a radical break with modern society and thought. Heidegger demanded radically new criteria by which to judge the correctness of institutions and forms of life, criteria that are outside of modern society and its legitimating framework. In the third section, I show why both the Marxist and Heideggerian rejection of politics collapse on internal grounds. Finally, in the fourth section, I turn to another source of skepticism about rational and non-coercive politics in Continental political theory: the sociological nightmare of Weber's "iron cage" and the Frankfurt School's "dialectic of Enlightenment." Habermas's work on democracy in complex societies is examined as offering the basis of an answer to such skepticism.

2. The Rejection of Modern Politics: Marx and Heidegger

Modernity begins with an optimistic political philosophy of progress towards human freedom, with new institutions and forms of knowledge overcoming past barriers of superstition and domination. German Idealism represents a certain watershed in this conception, by critically examining the presuppositions of these institutions and forms of knowledge and showing that their internal contradictions led to the deepening distress and fragmentation of modern social life (Habermas 1987: ch. 2). Like Kant and Hegel, Marx, too, retained the progressive philosophy of history and modernity while critically examining the consequences of modern institutions; but unlike his Idealist predecessors he found in them even deeper contradictions and irrationalities that could be solved only by a revolutionary transformation of society as a whole. Progress towards human emancipation that was promised by the unleashing of human powers over nature could only be achieved in a developed form of society that would fulfill the promise of democratic self-organization. It would lack the typically modern forms of political organization, including the state and all forms of the "hierarchical division of labor" (Marx 1977: 539). Indeed, politics based on conflict and the struggle for power between classes would disappear and be replaced by decentralized and highly participatory forms of self-rule, the goal of which is to transform politics from oppressive class rule into the "rational medium for social life" (Marx 1977: 577). "Real" democracy as "the political form of full human emancipation" brings about the "end" of politics.

Unlike the previous Enlightenment conception of progress, Marx emphasized acute discontinuities between an emancipated form of life and modern society and its institutions; politics is a medium for realizing freedom only in a new social totality. Although he left only hints of this self-organized society to come, it is clear that his intentions are democratic in a radical sense: radical because democracy can flourish in the absence of coercive background institutions whose sole purpose is to maintain various forms of hierarchy and domination. This normative conception of radical democracy leads to a "ruthless" criticism of all existing institutions. However, it is often unclear whether Marx thinks that all such institutions will simply disappear in this radically democratic form of society. In any case, his conception of a fully democratic society supports criticisms of

all aspects of modern institutions, including ones that we might think bear more normative weight than he allocates them. The criticism of politics as the medium of domination adopts all aspects of liberal, and hence bourgeois, institutions without remainder, including normative ideals, as so much ideology to be dispelled. In a manner similar to some post-modernists today, Marx shows how even such laudable notions as civil rights are mixed with relations of power and domination and can be put into their service. All such social relations are supposed to disappear in the future socialist society, in which egalitarian political relations will be so transparent that they no longer need resort to the disguise of ideology or appeal to the false claim that modern society is reintegrated in an imaginary community of political equals in a non-coerced public sphere.

For all this "ruthless criticism," Marx's radical democratic rejection of politics is not ultimately skeptical of the modern ideals of equality and democracy. Conversely, according to Heidegger, democracy itself and its ideal of public reason are implicated in a broader "onto-theological framework" of subjectivity and agency. Indeed, they are trapped within it and can never escape. Politics is simply one activity contained within it, and certainly the one most governed by it. "Publicness," for Heidegger, contrasts sharply with truth; it is compared with the ordinary, the familiar and the commonsensical, whose correctness can never be challenged since it appeals to the unquestioned character of our everyday understanding of the world (Heidegger 1993: 137). Neither reflection nor criticism is supposedly sufficient to extricate us from this understanding of the world. The only way out is through radical new possibilities, defined in terms of *authenticity*, which open up novel practical understandings (Heidegger 1993: 187). Politics, too, can become a location for much as disclosure of new possibilities, but only if it goes beyond politics in the ordinary sense and becomes something more akin to art, to the poetic "saying of the unsayable," that opens up a whole new understanding and framework for things. Before Heidegger, Nietzsche similarly contrasted the ordinary and mundane politics of everyday democratic life with the creative politics of the politician-artist which would create a world beyond modern subjectivism. Crucial to both is that non-ordinary politics provides new criteria, much as an innovative work of art is self-validating to the extent that it creates the standards by which it is evaluated.

Since democracy is implicated in the modern subjectivist framework that must be overcome, the standards of any such creative act of politics would no longer be recognizable as such within liberal democracy. As in a revolutionary political act, this new sort of disclosure cannot be judged by current standards. Instead, Heidegger offers the analogy between art and politics, filled out in "The Origin of the Work of Art" essay, as the asymmetrical relation between the law-giver and his people and between "creators" and "preservers" of the new truth (Heidegger 1993: 186, 193). This is where the analogy to Marx's radical democratic "end" to politics breaks down, since here politics in the ordinary sense is not replaced by some transparent social relations or rational medium. Nietzschean aestheticism and post-modern anti-institutionalism is the risk internal to most anti-modern politics, and it can be found particularly in radical right-wing thinkers such as Carl Schmitt, who contrasted "the political" as the moment of such normative self-definition with "politics" as the result of settled institutions of law and bureaucracy (Schmitt 1985). It is "the political" which is outside of modern politics and modern rationalism. While in the case of Marx we might wonder if politics can become such a transparent social medium, in the case of Heidegger and his followers we may wonder if there is any way to understand such a politics of non-public innovation except as an absolute act of rupture and even violence.

Whatever their differences on the political spectrum—and they are significant—Marx and Heidegger share with Hegel the attempt to think of politics through the critique of modernity and thus beyond the categories of liberalism. Where they differ is the extent to which this attempt is an immanent or external critique, whether they simply reject liberalism in what Hegel would call an "abstract negation" or try to go beyond it by also incorporating its truths. It is clear that Heidegger seeks an increasingly external criticism of liberalism, ultimately rejecting any attempt (even modeled on aesthetics) to change the predominant understanding of politics by human agency. Marx's enterprise is more traditional because he argued that democratic ideals do express the goal of human freedom through self-organization. The debate about post-modernism in post-war Continental political theory asks whether or not this immanent critique of modern institutions and liberal political norms goes far enough—and whether the external critique must always be anti-political. The question remains how such a thoroughly external critique can return to politics without endorsing some of these norms and institutions, especially if this return is for the sake of some version of radical democracy.

3. The Return of Politics: Civil Society, the Public Sphere and Democracy

The resolutely critical stance towards political norms in Marxism (and ultimately its anti-political view of an emancipated society) proved unsustainable in the development of critical thinking about politics in Continental philosophy. Like nature as the ground of human rights typical of the eighteenth century, the catastrophes of the twentieth century destroyed all faith that history could serve a similar justificatory function. This critique of the metaphysical basis for political norms could lead to skepticism about the rationality of modern politics, as it did for Weber, who saw politics as the simple struggle between opposing "gods and demons." The problem was that political life could not, even in modernity, be reduced to the mere exercise of power; there were spaces in which both the contestation of power and deliberation on norms and opinions could take place. Already in the work of Antonio Gramsci, the conception of politics as merely supporting the dominant ideology of the dominant class comes under critical scrutiny. Jürgen Habermas also tried to show a space for democratic political life in the public sphere, where citizens can meet as free and equal. Similarly, Claude Lefort and Cornelius Castoridis developed a conception of politics as a public space, the contours of which are shaped and determined by rights. These same arguments can be turned against both Heidegerrian anti-politics and anti-modernism.

In order to discover a basis for the integrity of modern politics, Gramsci returned to Hegel's idea that civil society contained more than merely the market and its "system of needs" but also "corporations" or distinctively modern forms of association (Gramsci 1977). Thus, Gramsci found in civil society the possibility of political life that did not merely collapse into market forces or state power. Indeed, in civil society we find specifically modern forms of plurality and associations, in the form of unions, cultural institutions, churches, clubs and political parties. Such criticisms of civil society that collapse it into the state made it impossible for the question of the transition to a genuinely democratic society to be posed within Marxism; the absorption of the state into civil society is not a step toward democracy. Similarly, Arendt also argued that political life could not be reduced to economic management without the disappearance of "the political" into "the social." Thus, Gramsci not only opposed the economic reductionism of the orthodox Marxist conception of politics, he also opposed reducing cultural

processes to mere coercion and power. This reduction of cultural and social integration to political coercion is common enough today among post-modernists, inspired now by Michel Foucault's conception of the pan-optical state (Foucault 1977). Gramsci, however, did not see abolishing the state as leading to the emergence of an autonomous and pluralistic civil society. Besides the administrative role of the state, its modern structure of rights and liberties created the autonomous spaces for politics. Hannah Arendt also sought to keep the social and political spheres separate, governing each with radically different principles of the household (*oikos*) and the polity (*polis*) (Arendt 1958). The domain of politics is not reducible to the logic of other domains (particularly economic or bureaucratic rationality), nor is it simply a matter of coercion. Thus, both Gramsci and Arendt in their neo-Marxist and neo-Aristotelian arguments seek to establish an independent realm for politics in modern society between the forces of the market and the state. But instead of Gramsci's civil society, Arendt calls the space for politics "the public sphere"; a space protected by republican institutions and threatened by the invasion of "the social."

Against Arendt's Aristotelian and agonistic model of politics, Habermas argues for the legitimacy of a distinctly modern and democratic public sphere emerging in the late eighteenth century (Habermas 1983). Unlike the early Enlightenment's emphasis on science and its ability to control nature, Habermas finds the "rational content of modernity" in the new possibilities afforded for moral and political identity. In the salons and coffeehouses of the late eighteenth century, according to Habermas's historical analysis, social status and power "were left at the door." Only when such a public sphere exists could citizens compel authority to legitimate itself before the tribunal of public opinion. For this to be possible, social spaces had to be created in the forms of clubs, associations, literary societies and union halls where the influences of wealth and power were limited, if not suspended, creating the forms of dialogue and interaction necessary for political relations among equals. This "public of private persons" emerged around such spaces (and in new forms of private life in the modern family) and was extended by the new communication media of newspapers and publishing houses. What unites these diverse phenomena is that, in them, publicity takes on a normative rather than functional significance.

Such a public sphere permits the formation of "public opinion" in a normative rather than merely factual sense, in which free and open communication makes possible not only the improvement of the quality of reasons for decisions and beliefs, but also the absence of direct relations of political coercion and social power. Consensus reached under such conditions of the absence of all force "except the unforced force of the better argument" has, Habermas supposes, more than merely *de facto* significance. These effects of publicity are due to the "reflexive quality" of such public communication: the public is concerned about itself as a public, about maintaining the public character of its communication. In concrete historical terms, this concern is found in the struggle of the emergent public with state authority and power, which sought to limit such communication through censorship. While Habermas is skeptical whether any actual public spheres (or the modern family structure and private sphere that supports it) lived up to these claims for their publicity, they nonetheless establish the foothold for a new kind of rationality in political life and for the social relations needed for self-rule by public opinion. Moreover, these claims are not merely ideological; their reflexive character of constant testing and self-examination make any false claim to publicity challengeable by the very norm of publicity it invokes. Publicity has thus acquired a different meaning

from the one it has in Heideggerian anti-politics, and the solutions to the problems of power and coercion are to be found in the public sphere itself.

The struggle against state censorship transformed the literary and discursive public sphere and the public sphere of associations in civil society into what Habermas calls "the political public sphere," the public sphere of citizens interacting with parliamentary and representative institutions. With the creation of this type of public sphere, non-coercive politics is no longer simply a check on independent state power (as are Hegel's "corporations" in civil society), but the democratic limitation of coercive political power through the agreement of citizens who govern themselves by laws within a constitutional framework. Not only must democratic institutions themselves create public spheres (of parliamentary debate, free and fair elections), but they must also be open to, and influenced by, the public sphere as a whole. The democratic state, too, helps to create and maintain the existence of the public sphere, primarily by way of constitutionally guaranteed basic rights, including civil rights of political participation and free expression. Thus, the public sphere is not merely a space for discussion, but one that is shaped by the political status granted to each citizen as free and equal. The interaction between the state and the public sphere raises questions about the status of institutions and the norms that constitute them and that are required to structure and constrain the interaction within them. In the absence of institutions and their normative structures, it is unclear whether consensus on issues of moral conflict could ever be reached.

In her analysis of the catastrophic events of the twentieth century, Arendt takes the status of "stateless people" to be paradigmatic of the failures of modern politics (Arendt 1953: ch. 9). Such people are reduced to their bare humanity and lack not only of basic rights, but "the right to have rights," the right to participate and shape a common world where one's speech and actions have meaning and significance. Thus, the return to politics, especially a non-coercive kind, requires rethinking basic modern institutions and constructing an immanent critique of their underlying conceptions. As carried out by thinkers as diverse as Arendt and Habermas this immanent critique pushes liberal institutions in the direction of a more participatory and deliberative form of democracy that goes beyond the aggregative voting and mere self-interest of most liberal interpretations of democracy and its representative institutions.

Like Habermas, Claude Lefort offers an immanent, rather than external, critique of liberalism, emphasizing the role of the public sphere. The difference is that, for Lefort, the public sphere is a space for resolving conflicts rather than reaching consensus, disconnecting the reinterpretation of liberalism and democratic institutions even more thoroughly from the excesses of the general will and political rationalism. In democratic institutions the place of power is unoccupied and hence belongs to no one (Lefort 1988: 27). If this is the case, then the main form of politics in democracy is contestation, the institutionalization of legitimate conflicts. This means that the public space, too, cannot belong to anyone and can thus accommodate everyone. It is not an agnostic public sphere as Arendt (following Aristotle) would have it, since it is public only because no one has to compete for recognition or acclaim. Indeed, as persons claim rights for themselves and change the meanings of existing rights this space becomes more expansive. Thus, as new rights claims become recognized as legitimate, the contours of the public sphere itself change, since the public space is structured around what everyone can claim as legitimate and by the ongoing reinterpretation of such claims through speaking and listening to others. Rather than excluding the plurality of opinions and identities,

a democratic public sphere makes pluralism possible through the constant contestation of the very basis of political justification and legitimacy.

As opposed to Marx and Heidegger, Gramsci, Habermas, Arendt and Lefort have developed conceptions of democracy that serve as a corrective to the coercive character of modern society. Civil society, the public sphere and democratic institutions all open up the possibility of a form of politics that is governed neither by the invisible hand nor sovereign state power. To the extent that these arguments are sound, they provide the basis for an immanent critique of modern reason, discovering in political practices a foothold for the unifying and reconciling power of reason in uncoerced consensus, democratic contestation of power and domination, and inter-subjective structure of mutual recognition. Thus, such an excessively one-sided picture of modern society leaves no room for possibilities for rationality.

This sort of thinking is, however, subject to a more empirical and sociological challenge, embodied in the sociology of Max Weber and the increasingly pessimistic diagnosis of the Frankfurt School. This challenge recognizes that there are other possibilities for the realization of reason, but see them as overwhelmed by the powerful institutional forces behind a one-sided rationalization of modern capitalist and bureaucratic society. In the next section I will explore Habermas's attempt to provide a theoretical and empirical antidote for such political skepticism, by showing that the forces of modernity can be balanced by the rationalization of the life-world.

4. Overcoming Sociological Skepticism about Democratic Politics

Max Weber's analysis of "legal domination" uncovers a certain dialectic at work in the increasing complexity of modern society (Weber 1958). On the one hand, through democracy, citizens make increasing demands on the state, creating the need for ever-greater bureaucracy. The responsiveness to such demands is a direct result of the increasing access of the people to powerful state institutions. On the other hand, modern societies are becoming more and more complex, with a greater and greater division of labor, the introduction of expert knowledge, social technologies and the functional differentiation of distinct spheres of social life. This dynamic produces an "iron law of oligarchy," whereby democracy is undermined by the very conditions that created it, including the disenchantment of the world that leads ultimately to the irrationality of choices among basic values and a vision of politics where "gods and demons" fight it out without any possibility of rational adjudication. Weber's tragic vision of the fate of modern democracy means that we must abandon any greater hopes for democracy than "competent administration" and live disconsolately with irreconcilable value conflict in public life. While in the 1930s the Frankfurt School opposed this diagnosis and attempted to develop a conception of "real democracy" that would bring all social conditions under the control of rational consensus, gradually its leading philosophers (Max Horkheimer and Theodor Adorno) adopted their own version of this same dialectic analysis of modern rationalization. Arguing that pressures towards uniformity and conformity emerge out of the contradictions of liberal society, these anti-democratic trends culminate not in Weber's "iron cage" but in the Fascist state and the authoritarian personality. Under these conditions, Horkheimer argued, the autonomous liberal individual is a "hopeless fiction" (Horkheimer 1982: 211), to the extent that the subjective conditions for the exercise of freedom were gradually being eroded by increasingly totalizing social reification (Horkheimer and Adorno 1972). Unlike the radical

critique of liberalism, democracy and its political institutions are not directly implicated in the problems of modern society. Rather, it is increasingly less feasible in the fragmented, instrumental, and ultimately anonymous social order that works against it and the autonomy of the modern individual.

Built into this modified critique of liberalism and its reifying form of social rationalization is a two-sided conception of totality. In many respects, reference to modern society as a totality becomes the defining feature of the Frankfurt School analysis (see Jay 1984). On the one hand, they use totality in a negative sense as an explanatory basis for criticism. Ultimately, the "real totality" of liberal society is "false," an antagonistic and inegalitarian class society rather than harmonious whole. On the other hand, they use totality in a positive sense as well; as a normative ideal and a desirable goal for social change. A "true" totality contrasts with the "real" totality of capitalist society and the "false" totality of Fascism. In contrast, a "true" totality is rational and consensual: the social process is brought back under free rational control as an expressive whole. In a true totality, the needs and interests of each will be expressed in the whole society; such a society will be the adequate expression of the social praxis of its members. Democracy does not figure directly in this positive totality, except as a means of consensual expression and moral transformation. Besides this expressivist social ideal, the guiding descriptive idea of the critique of the totalizing spread of instrumental reason is that this false totality becomes more and more the reality of modern society, leaving the true totality as an external, and ultimately distant, ideal preserved only in art and subjective experiences, and then only in a highly fragmentary form.

In an attempt to open the possibility of a democratic politics enlivened by the vibrant and non-coercive public sphere, Habermas challenges the descriptive adequacy of this account of modern rationalization as purely instrumental. He does so in two ways. First, he argues that this order is not a totality, but crisis-ridden (Habermas 1975: ch. 1). In "late" capitalism the free market does not prove to be self-regulating, but in need of constant state intervention. This intervention, in turn, puts heavy burdens on the state, transforming the economic crisis of market failure and collapse into the legitimation crisis of the modern state as the problems of economic growth exceed the capacities of the state. Second, he argues that modern rationalization is two-sided, including within it not only instrumental control but also the cultural rationalization of modern art, politics and morality on the basis of which further democratization is still possible. Such rationalization reveals a different form of practical reason already operative in modern society: communicative rationality, the use of which makes it possible for practical questions to be more than merely irreconcilable conflicts of ultimate values. Such rationality is grounded not only in modern institutions but the everyday practices of communication, practices of giving and testing reasons in dialogue with others, which in turn can become the basis for discursive procedures of justification.

Both the formal and informal achievements of communicative rationality help Habermas to rethink the relation between democratic politics and social complexity, and thus to blunt the force of both Weber's skepticism and Adorno's and Horkheimer's dialectics of rationalization. But, even for Habermas, the space for such politics is increasingly limited by the demands of social complexity, which make it impossible for everything to be governed by communicative rationality. Even if polycentric, modern states are not now organized solely by anonymous relations of interdependence.

Habermas agrees with Marx that advanced capitalism limits the scope and significance of democratic institutions and norms; it produces an "over-complexity," or path-

ological version, of the differentiation of modern society (see Bohman 1996: ch. 4). Much like the early Frankfurt School's criticism of majority rule, in the 1970s Habermas targets the "formal" character of current democratic practice (Habermas 1975). To this reduced version of democracy, Habermas opposes "substantive" democracy, which emphasizes the "genuine participation of all citizens" in political will formation. Such a notion of will formation requires more than a purely formal or self-interested analysis of rationality, and Habermas attempts to ground practical reason on the inter-subjective structure of communication exhibited in the special reflexive and reciprocal form of communication he calls "discourse." The validity of a political decision is now related to rational consensus and procedural norms. Accordingly, a norm is justified only if all those affected could agree to it in a discourse under the conditions of what Habermas calls an "ideal speech situation."

Because this ideal of consensus is primarily epistemological rather than political, its purpose is to establish a procedural, discursive and inter-subjective notion of rationality. Habermas's argument here is against the value skeptic (such as Weber) who sees politics as reducible to irrational struggle and conflict. And, because of the epistemological character of this notion of rationality, Habermas has always been suspicious of attempts to apply it (or its counterfactual "ideal speech situation") *directly* to the structure of political institutions (see, for example, Habermas 1979: 186). He maintains that the adequate mediation of these ideals with social facts requires an adequate theoretical account of the rationalization of institutions and culture. For Habermas, however, the fundamental limitation on the direct application of rational norms and ideals of communicative association to social reality is not the pervasiveness of reification, but of complexity. Along with the pluralism of forms of life, this complexity changes the conditions of popular sovereignty, so much so that "the people" is now a fiction. Complex societies are now polycentric and this changes the character of political participation that must be institutionally mediated in order to be feasible. However, complexity by itself is not reifying; indeed, it permits new possibilities of private and public autonomy, as well as the communicative and democratic structuring of many areas of social life. But it does eliminate one possibility: that it is possible for the sovereign will of the people to constitute all of society in a conscious sort of way (Habermas 1996: ch. 7).

This focus on social complexity has important consequences for the critique of liberalism as an ideology that was central to Marxism and the early Frankfurt School. Rather than reject liberal ideals, Habermas now attempts to appropriate liberal constitutionalism and rights and to combine them with radical democracy and its emphasis on consensus and public reason. He endorses a "two-track" model for democracy that makes its ideals practical in a complex and pluralistic society (Habermas 1996: ch. 8). On the one hand, decision-making goes on in institutions whose rules are characterized by a modern constitution with its principles of rights to basic liberties, political participation and the social goods necessary for both. On the other hand, such institutions must remain open to the public sphere, which he sees as an "anonymous network of communication" that creates the pool of reasons by which decisions are made. Such public opinion does not rule, but it does influence and direct the power constituted in institutions where the public is represented in majority rule which, in turn, is the outcome of informed opinion and fair and open deliberation. Rather than full participation in every decision, what is left of radical democracy is the ideal of "deliberative politics." Even if it does not return to radical democracy, the two-track model answers the Weberian political skeptics, by showing how the ideals of democracy are feasible under current social conditions and

by providing the workable model of deliberative democracy, in a variety of domains. Having taken a "cosmopolitan turn," Habermas's recent work now seeks to extend and decenter his conception of democracy beyond the limits of the state (Habermas 2006; Bohman 2007).

Related Topics

Kant, Hegel, Marxism and Contemporary Political Thought, Human Rights and Cosmopolitanism, Democracy, Discourse Theory

References

Arendt, Hannah (1953) *The Origins of Totalitarianism*, New York: Harcourt Brace.
—— (1958) *The Human Condition*, Chicago, IL: University of Chicago Press.
Bohman, James (1996) *Public Deliberation: Pluralism, Complexity and Democracy*, Cambridge, MA: MIT Press.
—— (2007) *Democracy Across Borders*, Cambridge, MA: MIT Press.
Foucault, Michel (1977) *Discipline and Punish*, trans. A. Sheridan, New York: Pantheon.
Gramsci, Antonio (1977) *Prison Notebooks*, trans. Q. Hoare and G. Nowell-Smith, New York: International Publishers.
Habermas, Jürgen (1975) *Legitimation Crisis*, trans. T. McCarthy, Boston, MA: Beacon Press.
—— (1979) *Communication and the Evolution of Society*, trans. T. McCarthy, Boston, MA: Beacon Press.
—— (1983) "Hannah Arendt: On the Concept of Power," in *Philosophical Profiles*, Cambridge, MA: MIT Press.
—— (1987) *The Philosophical Discourse of Modernity*, trans. F. Lawrence, Cambridge, MA: MIT Press.
—— (1989) *The Structural Transformation of the Public Sphere*, trans. T. Berger, Cambridge, MA: MIT Press.
—— (1996) *Between Facts and Norms*, trans. W. Rehg, Cambridge, MA: MIT Press.
—— (2006) *The Divided West*, trans. C. Cronen, Cambridge: Polity Press.
Heidegger, Martin (1993) *Basic Writings*, ed. D. Krell, New York: Harper and Row.
Horkheimer, Max (1982) *Critical Theory*, trans. M. O'Connell *et al.*, New York: Seabury Press.
Horkheimer, Max and Theodor W. Adorno (1972) *Dialectic of Enlightenment*, trans. J. Cumming, New York: Seabury Press.
Jay, Martin (1984) *Marxism and Totality*, Berkeley: University of California Press.
Lefort, Claude (1988) *Democracy and Political Theory*, trans. D. Macey, Minneapolis: University of Minnesota Press.
Marx, Karl (1977) *Selected Writings*, ed. D. McLellan, Oxford: Oxford University Press.
Schmitt, Carl (1985) *Political Theology*, trans. G. Schwab, Cambridge, MA: MIT Press.
Weber, Max (1958) *From Max Weber*, ed. H. Gerth and C. W. Mills, Oxford: Oxford University Press.

Further Reading

Dews, Peter (1987) *Logics of Disintegration*, London: Verso.
Honneth, Axel (1988) *The Critique of Power*, trans. K. Baynes, Cambridge, MA: MIT Press.
McCarthy, Thomas (1991) *Ideals and Illusions*, Cambridge, MA: MIT Press.

16
FRENCH POLITICAL THOUGHT IN THE TWENTIETH CENTURY

Jeremy Jennings

1. The Impact of the Dreyfus Affair

A nation's political thought is a reflection of its history and the political turmoil of the age in which it is written. This is particularly true of French political thought in the twentieth century. The century began with a society riven in two by the Dreyfus Affair. The wrongful imprisonment of Captain Alfred Dreyfus for high treason led not only to the subsequent publication in 1898 of novelist Emile Zola's open-letter condemning the miscarriage of justice but also to bitter and sustained polemic among writers on politics who remained sharply divided between those who held and wished to advance the values associated with the Revolution of 1789 and those who sought a return to the "true France" of the Roman Catholic religion, the army and the monarchy. The Third Republic, born out of humiliating defeat by Prussia, the loss of Alsace and Lorraine, the trauma of the Paris Commune and its bloody repression, found itself challenged by those convinced that the republic's sustaining ideology of *solidarité* was little more than a hollow sham hiding bourgeois self-interest. Theorized by the philosophers Célestin Bouglé and Alfred Fouillée and by sociologist Emile Durkheim, this doctrine sought to reconcile the claims of individual liberty and social justice and did so by placing an equal emphasis upon rights and duties. The overriding ambition was to secure social peace (Spitz 2005; Blais 2007). Yet, as disillusionment spread among those who had supported the Dreyfusard cause, the republic and its politicians appeared to embody little else but corruption and mediocrity.

One thinker who reached these conclusions was Georges Sorel. Drawn to Marxism in the 1890s, Sorel had supported Dreyfus and had done so believing that this was in accordance with the ethics of socialism. However, he quickly concluded that the "rush for spoils" associated with the victory of the Dreyfusards had handed France over to the forces of demagogy and bourgeois stupidity and, in response, he aligned himself with the emerging revolutionary syndicalist movement. To that end, in 1906 he published *Reflections on Violence*, a text which directed its fire against parliamentary socialism and what Sorel saw as a timorous and humanitarian capitalist class intent on holding out the hand of friendship to the proletariat. The proletariat, Sorel proclaimed, would be better served by embracing the tactics of class war and by responding with "black ingratitude"

to the blandishments of the bourgeois republic. In its advocacy of the general strike and direct action this was precisely what the syndicalist movement appeared to be doing.

A similar enmity towards the republic was to be found among those who opposed calls for the release of Dreyfus. Throughout the early years of the Third Republic the supporters of an ever-more strident and populist nationalism had sought repeatedly to overturn the political institutions of the republic, but with the Dreyfus Affair this stance hardened into the charge that the republic was in the grip of the forces of "anti-France"—Jews, Protestants, Freemasons and *métèques*—and that only a return to a monarchy that was "traditional, hereditary, antiparliamentary and decentralized" could save France from the decay produced by a weak and unstable regime. This, for example, was the view of Charles Maurras and his many supporters in the *Action Française* movement founded in 1898.

Given their shared conviction that republican democracy was irredeemably corrupt and that bourgeois society was incapable of producing anything of moral seriousness, it was not unduly surprising that, in the years prior to World War I, Sorel and Maurras found themselves in something of an uneasy alliance. Their enthusiastic admirers, clustered around the *Cercle Proudhon*, gave ample testimony to the political disenchantment and iconoclasm of those who came to be known as "the generation of 1914."

The mood of this generation of young intellectuals was best captured in the survey published by Henri Massis and Alfred de Tarde under the pseudonym of Agathon. *Les Jeunes Gens d'aujourd'hui*, they recounted, had turned against a sterile rationalism and had abandoned the admiration for the scientism of Hippolyte Taine and Ernest Renan. Inspired by the philosophy of Henri Bergson, they were driven by a sense of moral and spiritual idealism. Putting doubt to one side, they sought adventure, action and heroism. For example, Ernest Psichari, the grandson of the arch-skeptic Renan, joined the army, wrote successful novels about his experiences as a colonial soldier, and, like many of his fellows, converted to Catholicism. Charles Péguy, having defended the Dreyfusard movement in "Memories of Youth" (2001), found himself increasingly drawn to the *mystique* of the French nation.

2. The Great War and its Aftermath

When war was declared in 1914 few were those who did not respond to the calls for a *union sacrée*. If, for Péguy, this meant an early death in the battle of the Marne, for others it entailed an intellectual mobilization in support of the French war effort. The land of Descartes, liberty and civilization was pitched against the country of Hegel, State power and barbarism. Two themes were much in evidence. From Catholic writers such as Maurice Barrès came that of an "eternal France" whose mission could be identified with the ends of Christ and of humanity. From republicans such as Emile Boutroux and Alphonse Aulard came the refrain of revolutionary France as the *patrie* of rights, justice and universal emancipation.

Little was heard from those who opposed the war or who, like Romain Rolland, remained "above the fray." This was to change with the advent of the Bolshevik Revolution in 1917. An indigenous revolutionary tradition had delayed the implantation of Marxism into France but from the 1880s onwards adherents such as Jules Guesde and Paul Lafargue began the process of developing Marxist theory in a French context. For the most part—Georges Sorel was the notable exception—this amounted to an endorsement of an unsophisticated and crude economic determinism but it did pave the

way for a later receptiveness towards Marxist ideas and practice in the early decades of the century. At the outset of the Bolshevik Revolution little was known of either Lenin or of events in Russia but, viewed from a distant Paris, it was possible to see the Bolsheviks as heirs to Robespierre and the Jacobins and to characterize 1917 as a continuation of the revolutionary events of 1789–94. When combined with a militant pacifism and the desire to break with the political compromises that had led to the mass slaughter of the Great War, this contributed to the creation of the French Communist Party (PCF) in 1920. The consequences of these developments upon the character of French political thought over the next decades were to be profound.

Much might be said about the manner in which the PCF exercised both a fascination and a stranglehold over significant sections of French intellectual opinion in these years but, at a minimum, it produced a stultifying orthodoxy and placed many writers in a condition of subservience to the dictates of Moscow. As André Glucksmann was later to comment, Marxism nourished not only a set of convictions but also a desire not to see things as they really were (Glucksmann 1975). Naivety and self-delusion have been among the more prominent features of French political thought in the twentieth century.

Accordingly, in the interwar years political commitment became the norm as French intellectuals argued over pacifism, colonialism, the threat of war, the rise of fascism and the future of western civilization more generally. In those circumstances one debate that came to the fore was that of the proper role of the intellectual. Prefigured in the exchanges between the writers Henri Barbusse and Romain Rolland in the early 1920s, this amounted to a discussion about whether the intellectual was the voice of an independent moral conscience or whether intellectuals were to be prepared to subordinate themselves to what were taken to be the needs and the political will of the masses. The clearest statement of the first of these positions came from the pen of Julien Benda.

3. Benda, Nizan, Weil, and the Role of the Intellectuals

As set out in *The Great Betrayal* of 1927, Benda's argument was that the intellectual was to be the defender of the abstract over the concrete and of the universal over the particular. Self-consciously modeled upon a vision of the medieval cleric withdrawn from the daily realities of the world, writers were under a duty to speak the truth with as much clarity as possible and, in so doing, were to disregard the pursuit of practical ends and considerations of temporal advantage. As Benda phrased it, the *clercs* of the modern age were to be "the officiants of abstract justice." In his opinion, however, the intellectuals had abandoned this function, serving the prejudices of the multitude and encouraging "the three great passions of our age": class hatred, nation hatred and race hatred.

The most articulate reply to Benda's argument came in the form of Paul Nizan's *The Watchdogs*, published in 1932. A Marxist and a member of the PCF, Nizan's basic premise was that, if some philosophies were "salutary" to humankind, then others were "lethal" to it. On this view, the supreme function of bourgeois philosophy was "to obscure the miseries of contemporary reality." Refusing to take sides was to acquiesce to exploitation, poverty, war and oppression and to place personal spiritual comfort before the needs of real people. Philosophers therefore were called upon "to embrace the cause of those impure creatures called men" and, in this way, "to denounce all the conditions which prevent men from being human." It was in the "abandonment" and "repudiation" of their fellows and not their denial of "eternal absolutes," Nizan contended, that lay the

"real treason of the intellectuals" (Nizan 1971).

For Nizan, then, the choice was simple. One was either for the oppressed or for the oppressors, and the intellectual, now defined as "the technician of revolutionary philosophy," should feel no shame in betraying the bourgeoisie for the sake of humanity. In the difficult economic and international circumstances of the 1930s such a call to arms was not without appeal. As postwar optimism faded, in France, as elsewhere, the idea of a crisis of civilization took hold. Amid talk of the disorder of capitalism and the failure of parliamentary democracy, writers from across the political spectrum called not only for a radical reform of the institutions of the state and the rational organization of the economy—planning was much in vogue—but also for something akin to the spiritual renovation of society as a whole.

Often articulated through the numerous new reviews which sprang up at the time, this longing to be free from the de-humanizing materialism and productivism of bourgeois society took a variety of forms. One of the most representative of these was the development of the doctrine of personalism in the pages of Emmanuel Mounier's journal *Esprit*. Opposed to both liberal individualism and statist collectivism, Mounier's ambition was to foster both a moral and political renaissance. "A personalist civilization," Mounier wrote, "is a civilization whose structures and intelligence are directed towards enabling all those individuals who comprise it to become persons … where they can enjoy the maximum of initiative, of responsibility and spiritual life" (Mounier 1961: I, 523).

Others of a less Catholic persuasion found it difficult to resist the attraction of the new civilization being built, it was believed, so admirably and so gloriously in Stalin's Soviet Russia. Here, according to André Gide, was a land where, if freedom of speech and opinion was limited, "utopia is in the process of becoming reality" (Gide 1937: 15). Still others—for example, novelist Pierre Drieu La Rochelle—succumbed to the (often homo-erotic) attractions of a virile fascism (of either the Nazi or Italian variety). Most were agreed in their loathing of the unabashed materialism and cultural vulgarity of American society, a dystopia frequently portrayed through the image of Chicago's vast slaughterhouses and meat packing factories (Duhamel 1931; Mathy 1993).

As the political and intellectual tensions of the interwar years mounted, one writer who ploughed a solitary and unclassifiable furrow was Simone Weil. Despite her bourgeois upbringing and elite education, Weil entered the world of militant trade unionism and was soon writing for a series of dissident left-wing publications. In 1934 she published *Oppression and Liberty*. Never an orthodox Marxist, Weil saw that the Soviet Union, far from liberating the proletariat, had succeeded in creating a new form of oppression in the form of a state bureaucracy run on behalf of a privileged intellectual caste. Yet her general (and bleaker) point was that oppression was written into modern industrial production itself, the process of rationalization and automation allowing for only the existence of "the specialized unskilled workman, completely enslaved to the machine" (Weil 2001: 20). In such a situation, Weil concluded, the best those who continued to honor human dignity could hope for was "to encourage whatever is capable, in the sphere of politics, economics or technique, of leaving the individual here and there a certain freedom of movement amid the trammels cast around by the social organization" (Weil 2001: 114).

In the years preceding her premature death in England in 1943 Weil's concern to ameliorate the condition of the manual worker did not diminish but her work was increasingly infused with a profound religious mysticism. Her preoccupation now became the needs of the soul (Weil 2002). And these, she affirmed, were best met through "the

growing of roots" and by placing our duties before our rights. What this would have meant in terms of a precise political program is hard to gauge but it is clear that Weil's call for a "refashioning of the soul of the country" would have entailed a fundamental re-thinking of France's republican political culture.

4. World War II: Collaboration, Resistance, and Liberation

That this was required was made vividly clear by France's abject surrender in the summer of 1940. As Marc Bloch's *Strange Defeat* (1949) argued, if the collapse of the French army and the liquidation of the republic had marked both a military and a political defeat, so too they had been the result of an intellectual laziness and complacency characteristic of French society as a whole. What followed only proved the point. If intellectual collaboration with the German occupier flourished in Paris (Riding 2011), around the Vichy regime of Marshal Pétain there was no shortage of writers—Charles Maurras foremost among them—prepared to endorse the causes of "Work, Family and Homeland" and of an authoritarian and anti-Semitic state.

With the liberation of France in 1944 those accused of intellectual collaboration were "purged." Robert Brasillach, the pro-Nazi editor of the journal *Je suis partout*, was executed. Maurras was imprisoned. Many others were refused the right to be published and lost their jobs. An unforeseen by-product of this frequently squalid process was that it cleared the way for a new generation of writers to emerge. One of those who made the most of this opportunity was Jean-Paul Sartre.

Sartre had steered clear of all political engagement during the 1930s, concentrating rather upon the satisfaction of his literary and philosophical ambitions. He similarly avoided contact with the Resistance during the Occupation. In 1933, however, he had gone to Berlin with the express purpose of familiarizing himself with the phenomenology of Husserl and Heidegger. This led in 1939 to the publication of both *A Sketch of a Theory of the Emotions* and the novel *Nausea*, the latter providing a (subsequently famous) description of the world as a contingent and formless mass. It was during the war—in the year that his first play, *The Flies*, was performed on the Paris stage—that Sartre published his philosophical *magnum opus*, *Being and Nothingness*. Subtitled "An Essay in Phenomenological Ontology," it was this text that laid the foundations for what after the war was to become known as existentialism. Starting from an analysis of what he characterized as the three forms of being and displaying considerable literary ingenuity, Sartre drew a picture of humans who, through a process of negation, could not only be said to be totally free but who, through various forms of "bad faith," sought to avoid the anguish brought about by that freedom. The moral imperative was to secure an authentic existence.

Perfectly capturing the wider mood of the time—the Occupation had familiarized everyone in France with the difficulty of making moral choices—existentialism met two immediate objections: for conservatives, it appeared to be an attack upon bourgeois morality; for communists, it amounted to an advocacy of despair and political quietism. The publication of Simone de Beauvoir's *The Second Sex* in 1949 only added to the controversy: for its critics, existentialism was now associated with pornography.

Sartre's first (if unsuccessful) attempt to respond to these criticisms came in a public lecture entitled *Existentialism and Humanism* (1948). Denying that existentialism saw human beings as condemned to a perpetual struggle with each other, Sartre suggested that if each of us willed his or her own freedom, so we were choosing freedom for eve-

ryone else. What was being proposed, Sartre declared, was an ethic of action and of self-commitment in the name of the dignity of man.

Over the next decade Sartre re-worked this argument in a variety of literary and philosophical guises. Commitment in particular—defined by Simone de Beauvoir in her memoirs as "the writer's total presence in what he has written"—came increasingly to perform the function of giving content to existentialism. But it was not long before Sartre realized that he was in a philosophical impasse. He attempted escape through Marxism and the formulation of an existentialist dialectic.

By the time that Sartre published the *Critique of Dialectical Reason* in 1960 he had become convinced that Marxism was "the inescapable philosophy of our time." Existentialism was simply parasitic upon it, its only task—albeit an important one—being to remind Marxism of the decisive role of human subjectivity. Conflict between individuals, Sartre argued, arose not at a phenomenological level (as he had previously believed) but as a result of our practical activity, principally the desire to satisfy our needs in a world characterized by scarcity. We lived in a condition of "seriality" where our relationship with others was one of separation rather than reciprocity. Seriality, Sartre suggested, would be overcome through the spontaneous organization of "the group," its essential characteristic being that it would be composed of individuals with mutual interests in a post-scarcity society. Everyone would see the Other as himself. This transformation would typically arise, Sartre argued, in situations of external threat or, more likely, of revolution.

Yet written into the *Critique of Dialectical Reason* was a tragic vision of our fate. If, after the revolution, our aim would be to preserve the group, Sartre recognized that its existence was at best fragile and short-lived. Under persistent threat from outside forces and by the existence of disunity from within, the group faced inevitable destruction as, in order to endure, it had to become an institution with both formal structure and hierarchical authority. In the real world something very similar appeared to have occurred in the Soviet bloc.

5. The Cold War Dilemma

How to respond to the practices of actually-existing socialism was one of the central questions addressed in French political thought from the late 1940s until the early 1960s. The need for an answer was made all the more pressing by the emerging Cold War and by the belief among many that France should avoid becoming a satellite of either the USSR or the USA. The dilemma—expressed most succinctly by Maurice Merleau-Ponty (Merleau-Ponty 2000: xxi)—was that, if it was not possible to be a communist, it was impossible to be an anti-communist. In brief, if the Soviet Union had renounced the liberty of the proletariat, the Marxist critique of capitalism was still valid. Merleau-Ponty's own answer therefore involved taking a substantial bet on the "objective sense" of history. By this criterion only the verdict of the future could determine whether the actions of the Soviets were just and, as this was as yet unclear, the Moscow show trials of the 1930s—denounced by Arthur Koestler in *Darkness at Noon*—could be excused. For Merleau-Ponty, it was not a matter of whether or not one accepted violence but of whether the use of violence was progressive or not.

It was this very argument that, according to Simone de Beauvoir, was to send Albert Camus into a rage and which, a few years later, was to lead to his acrimonious break with Sartre (Aronson 2004). Like Simone Weil, the Algerian-born Camus gravitated to the

world of libertarian socialism during the 1930s and then to the Resistance, editing its leading journal *Combat*. A key theme in his early writings—apparent in his novel *The Outsider* of 1940 and the philosophical essay *The Myth of Sisyphus* of 1941—was a desire for lucidity and to "gaze steadfastly" at life as it was. He similarly sought to construct a "modest" politics divested of messianism and of any nostalgia for a terrestrial paradise. This won him few friends in post-Liberation Paris, where he increasingly came to equate Marxism with murder. Matters came to a head with the publication of *The Rebel* in 1950.

Three aspects of Camus' argument stand out. The first was his criticism of the location of all values in history and its political implication that the ends justified the means. The second was his analysis of terror as being inevitable to the process of revolution, thereby disqualifying and discrediting the notion of revolution itself. The third was his distinction between revolution and the revolt of the individual in the name of justice and freedom. Crucially, the rebel did not claim the right to universal murder.

The Sartrean response—written by Francis Jeanson—was withering in the extreme, attacking both the man and his ideas. In essence, Camus was accused of being a bourgeois moralist who sought a false universalization of values. But none of this could diminish the power of Camus' fundamental insight: the city of free and fraternal human beings could not be built through mass murder.

A similar line of argument was developed by Raymond Aron. With Marxism cast unambiguously as "the opium of the intellectuals" and as a "secular religion" (Aron 2003), Aron turned his fire against Soviet totalitarianism and, as he did so, sought to develop a liberal definition of freedom based upon the concepts of the rule of law and a sense of individual responsibility. Drawing upon a distinction made by Max Weber, Aron recommended an ethics of responsibility rather than one of conviction. Politics, Aron wrote, "will always remain the art of irrevocable choice by fallible men in unforeseen circumstances and semi-ignorance" (Aron 2003: 199). According to Aron, to think politically entailed prudence, moderation and sound judgment rather than Promethean vision.

6. Structuralism and After

Aron's views were deeply unpopular in intellectual circles during the 1950s and it was not, therefore, to be his version of economic liberalism but the rise of structuralism that dislodged existential Marxism from its position of dominance (Dosse 1997). Sometimes seen as a theoretical response to the impasse of the political culture of the French Fifth Republic (Wolin 2006: 174), the core tenet of structuralism as a philosophical project was that the autonomy of the subject was an illusion. What this anti-humanist position entailed in terms of structuralism's attitude to politics was neither immediately obvious nor easy to decipher but the idea that the presuppositions of modern humanist philosophy served not to defend human dignity but as a cause of oppression was read by many as an attack upon the intellectual foundations and ethno-centrism of western bourgeois liberal society. If this was most striking in the attempt of anthropologist Claude Lévi-Strauss to "legitimize the principles of savage thought," it was much in evidence in Jacques Lacan's ethics of desire, Roland Barthes' analysis of the sign-systems of petit-bourgeois culture and Michel Foucault's dissection of the prison system in *Discipline and Punish*.

One consequence of the structuralist retreat from Marx was to bring three other German-language writers into prominence: Heidegger, Nietzsche and Freud. As a result,

French thought came to amount to a commentary on German thought. From this followed the near-disappearance of political philosophy as a distinct discipline. This was most clearly visible in the work of Jacques Derrida. If the radical structuralism and Heideggerianism denoted by Derridean deconstruction cast suspicion upon the "logocentrism" of the entire western philosophical tradition from Plato onwards, it provided no adequate grounds for making judgments about political questions. At best—and for all Derrida's later efforts to rectify the situation in works such as *Politics of Friendship* (1997)—it hinted at some vague emancipatory project and a celebration of difference. Increasingly marginal in France, in the 1980s Derrida's ideas successfully crossed the Atlantic and, in translated form, became part of the academic cottage industry associated with postmodernism. In domestic political terms they were an irrelevance.

To the obvious embarrassment of the structuralists, their proclamation of the death of man did not prevent people from fighting all the harder to attain this very status. Evident in the struggle against colonialism (where it received its most articulate formulation in Frantz Fanon's *The Wretched of the Earth*), the most decisive practical refutation of the claim that society was composed of agentless structures came in the form of the student demonstrations of May 1968. Interpreting these events and assessing their legacy has proved far from easy (Ferry and Renault 1990: 34–8). Raymond Aron simply dismissed this "elusive revolution" as a "psychodrama" (Aron 1969). Others such as Gilles Lipovetsky (1994) and Alain Renault (1997) have seen fit to portray May '68 as the dawn of a "new era of the individual" and of a society characterized by consumption and hedonism. According to Lipovetsky, it was the first revolution where "frivolity won out over tragedy" (Lipovetsky 1994: 208).

One casualty of the events of May '68 was the orthodox Marxism of the PCF. If it limped on in the form of the structuralist Marxism of Louis Althusser (Elliott 1987), a new generation of Leftists turned to the writings of the young Marx on alienation and species-being to critique not only modern consumer society but also the Stalinist practices of the Soviet Union. The prospect of a utopian future—frequently infused with a good dose of bourgeois self-hatred—was now transferred to Castro's Cuba, Mao Tse-Tung's China and other points east (Wolin 2010). If Jean-Paul Sartre was to be one of the offenders, embracing anti-colonialism and revolution as a matter of principle, the extent of the misguided thinking this involved was revealed by Julia Kristeva. Herself freed from "the totalitarian blindness typical of our western rationalism," in Mao's murderous Cultural Revolution she saw the "unprecedented renaissance of a new humanity" (Kristeva 1977: 201).

Such *gauchiste* fervor could not survive the reality of the killing fields of Cambodia nor the plight of the Vietnamese boat people. Nor was it impervious to the massive impact upon public opinion of Alexander Solzhenitsyn's *The Gulag Archipelago*, published in France in 1974. No one could now deny the existence of labor camps in the Soviet Union. Moreover, the arrival of significant numbers of dissidents from the Soviet bloc made it increasingly difficult to ignore questions relating to democratic freedoms and the rights of individuals. Michel Foucault, for example, abandoned Maoist militancy in favor of an ever-more public advocacy of human rights.

Among the most prominent of those who recanted were the so-called New Philosophers. Taking their name from a set of essays edited by Bernard-Henri Lévy and published in 1976, these penitents claimed great originality for seeing what outside France had long been apparent: the Soviet Union was a repressive and totalitarian regime beyond reform. Lévy's *Barbarism with a Human Face* was typical of the genre. Citing

Solzhenitsyn as the "Dante of our Time," Lévy wrote that the Gulag was "not a blunder or an accident, nor a simple wound or aftereffect of Stalinism, but the necessary corollary of a socialism which can only actualize homogeneity by driving the forces of heterogeneity back to its fringes." There could, he continued, be "no socialism without camps, no classless society without its terrorist truths" (Lévy 1979: 158). Similar arguments bemoaning the horrors of Marxism were to be found in the writings of André Glucksmann, most notably *The Master Thinkers*. On this account, the idea that history was the bearer of hope and of a radiant future was an illusion. Our most important task was to look evil square in the face.

Just as significantly, the French Revolution of 1789—previously held to be sacrosanct and part of the national patrimony—was implicated in the totalitarian nightmare. With historian (and former PCF member) François Furet in the lead, Robespierre's Reign of Terror was now seen as an integral and not contingent part of the Jacobin project. If, as had long been maintained by the Marxist historiography of the Revolution, the Jacobins anticipated the Bolsheviks of 1917, the only conclusion to be drawn was that terror was central to all revolutionary politics. The French, Furet proclaimed, had to realize that the Revolution was "over" (Furet 1981).

In these new circumstances several developments in French political thought were evident. One was a growing interest in issues such as feminism, gay rights and environmentalism. Another was the partial emergence of a non-Jacobin or "second" Left. The writings of Pierre Rosanvallon are the most significant expression of this second strand (Rosanvallon 2006). Still another was a move from revolution to ethics, typified by a renewed interest in the work of philosophers Paul Ricœur and Emmanuel Levinas.

Of equal significance was a relatively short-lived revival of liberalism. Since the fall of the July Monarchy in 1848, France had been barren terrain for liberal political thought, but the post-May 1968 blossoming of civil society and an increasingly skeptical appraisal of the benefits of state intervention in the economy seemed to indicate that this might be at an end. For the only time in his life, Raymond Aron found himself to be fashionable. Political philosophers and commentators turned to their Anglophone counterparts for guidance, reading Rawls and Nozick as well as their communitarian critics. The writings of Alexis de Tocqueville and then those of Benjamin Constant were subject to close scrutiny as part of a desire to recover and learn from the neglected tradition of nineteenth-century French liberalism.

Yet the attempt to formulate and popularize a distinctively French version of liberalism by writers such as Pierre Manent (1994) and Marcel Gauchet (1997) was one of unfulfilled promise. Rather, the most significant feature of the last two decades has been a revival of republicanism. In part this has been an exercise in political nostalgia but it has also been driven by the need to articulate responses to what, in a French context, have been the challenging phenomena of multiculturalism and globalization.

7. Multiculturalism and Globalization

The debate about multiculturalism began in earnest in 1989 following the controversy over the wearing of the veil or *hijab* in French schools. It has since not subsided (Laborde 2008). The response of traditional republicanism—articulated most forcefully by Régis Debray—was that the wearing of the veil was an expression of religious fundamentalism and was thus incompatible with the secular principles of republican citizenship. During the 1990s this was extended into a broader assault upon multiculturalism as a form of

communitarianism threatening the political integrity of France. Multiculturalism was equated with tribalism, balkanization, the tyranny of the minority and political correctness (Jennings 2000). From this it would be wrong to conclude that the cause of multicultural citizenship has not had its supporters but those prepared to contemplate a conceptual refashioning of republicanism to meet new social realities have been in a minority. Of these the most impressive has been Dominique Schnapper, author of *Community of Citizens* (1998). Accepting that the goal of the republic is to transcend particularistic differences and allegiances, be they ethnic, religious, biological, or economic, she nevertheless recognizes that a tension exists between the universalistic aspirations of republican citizenship and the reality of a society characterized by ethnic and social diversities and inequalities. The republic, she argues, can no longer aspire to be built upon the utopia of an abstract humanity.

This reality—one characterized by relative economic decline—has also informed debates about globalization. Over the last three decades the French state has remained relatively impervious to the demands of economic liberalization and has largely preserved its distinctive and generous model of social provision. Efforts to secure reform have been met not only by extensive street protests but by intellectual resistance from those opposed to what has repeatedly been characterized as "the dominant discourse" or "*pensée unique*" of neo-liberalism. The most articulate exponent of this position has been Pierre Bourdieu, for whom the civilization of republican equality was under attack from a "planetary vulgate" characterized by "the imperialism of neo-liberal reason" (Bourdieu 2008). Since Bourdieu's death in 2002, the flame of resistance to a market economy (as well as to parliamentary democracy) has been passed to Alain Badiou, for whom former President Sarkozy was the "Rat Man" and the prevailing ideology is that of the "Pétainist transcendental" (Badiou 2008).

If, then, political thought in twentieth-century France began under the shadow of the Dreyfus Affair, a century later it remains under the shadow of what Badiou has called "the communist hypothesis" (Badiou 2010). We must have the courage, Badiou maintains, to believe that human society can be more than a collection of individuals pursuing their self-interest and ruled by the law of profit. This might well be true, but Badiou's dismissal of political philosophy as "the servant of capitalo-parliamentarism" also tells us something about both the time warp and the insularity in which most recent French thinking about politics has taken place. In this guise, thinking about politics amounts to little more than heroic posturing. Ironically, Badiou, like Foucault and Derrida before him, finds most of his audience outside France. It is there that "French theory" lives on.

Acknowledgments

I would like to thank Robert Willer for his careful and suggestive reading of a first version of this chapter.

Related Topics

Continental Political Philosophy, Liberalism, Republicanism, Marxism and Contemporary Political Thought, Feminism and the History of Political Philosophy, Multiculturalism, Postmodernism and Politics

References

Aron, R. (1969) *The Elusive Revolution: Anatomy of a Student Revolt*, London: Pall Mall Press.

—— (2003) *The Opium of the Intellectuals*, New Brunswick, NJ: Transaction.

Aronson, A. (2004) *Camus & Sartre: The Story of a Friendship and the Quarrel that Ended it*, Chicago, IL: University of Chicago Press.

Badiou, A. (2008) *The Meaning of Sarkozy*, London: Verso.

—— (2010) *The Communist Hypothesis*, London: Verso.

Blais, M.-C. (2007) *La solidarité: Histoire d'une idée*, Paris: Gallimard.

Bloch, M. (1948) *Strange Defeat*, Oxford: Oxford University Press.

Bourdieu, P. (2008) *Political Interventions: Social Science and Political Action*, London: Verso.

Derrida, J. (1997) *Politics of Friendship*, London: Verso.

Dosse, F. (1997) *History of Structuralism*, Minneapolis, MN: University of Minneapolis Press.

Duhamel, G. (1931) *America the Menace*, Boston, MA: Houghton Mifflin.

Elliott, G. (1987) *Althusser: The Detour of Theory*, London: Verso.

Ferry, L. and Renaut, A. (1990) *French Philosophy in the Sixties*, Amherst: University of Massachusetts Press.

Furet, F. (1981) *Interpreting the French Revolution*, Cambridge: Cambridge University Press.

Gauchet, M. (1997) *The Disenchantment of the World*, Princeton, NJ: Princeton University Press.

Gide, A. (1937) *Back from the USSR*, London: Secker & Warburg.

Glucksmann, A. (1975) *La Cuisinière et le Mangeur de l'Homme: essai sur les rapports entre l'Etat, le marxisme et les camps de concentration*, Paris: Seuil.

Jennings, J. (2000) "Citizenship, Republicanism and Multiculturalism in Contemporary France," *British Journal of Political Science* 30 (4): 575–98.

Kristeva, J. (1977) *About Chinese Women*, London: Marion Boyars.

Laborde, C. (2008) *Critical Republicanism: The Hijab Controversy and Political Philosophy*, Oxford: Oxford University Press.

Lévy, B.-H. (1979) *Barbarism with a Human Face*, New York: Harper & Row.

Lipovetsky, G. (1994) *The Empire of Fashion: Dressing Modern Democracy*, Princeton, NJ: Princeton University Press.

Manent, P. (1994) *An Intellectual History of Liberalism*, Princeton, NJ: Princeton University Press.

Mathy, J.-P. (1993) *Extrême-Orient: French Intellectuals and America*, Chicago, IL: Chicago University Press.

Merleau-Ponty, M. (2000) *Humanism and Terror: The Communist Problem*, New Brunswick, NJ: Transaction.

Mounier, E. (1961) *Œuvres de Mounier*, Paris: Seuil.

Nizan, P. (1971) *The Watchdogs*, New York: Monthly Review.

Péguy, C. (2001) "Memories of Youth," in *Temporal and Eternal*, Indianapolis: Liberty Fund, pp. 3–84.

Renault, A. (1997) *The Era of the Individual: A Contribution to the History of Subjectivity*, Princeton, NJ: Princeton University Press.

Riding, A. (2011) *And the Show Went On: Cultural Life in Nazi-occupied Paris*, London: Duckworth.

Rosanvallon, P. (2006) *Democracy, Past and Future*, New York: Columbia University Press.

Sartre, J.-P. (1948) *Existentialism and Humanism*, London: Methuen.

Schnapper, D. (1998) *Community of Citizens: On the Modern Idea of Nationality*, New Brunswick, NJ: Transaction.

Spitz, J.-F. (2005) *Le Moment Républicain en France*, Paris: Gallimard.

Weil, S. (2001) *Oppression and Liberty*, London: Routledge.

—— (2002) *The Need for Roots*, London: Routledge.

Wolin, R. (2006) *The Frankfurt School Revisited*, London: Routledge.

—— (2010) *The Wind from the East: French Intellectuals, the Cultural Revolution and the Legacy of the 1960s*, Princeton, NJ: Princeton University Press.

Further Reading

Bourg, J. (2007) *From Revolution to Ethics: May 68 and Contemporary French Thought*, Montreal: McGill-Queen's University Press.

Christofferson, M. S. (2004) *French Intellectuals against the Left: The Antitotalitarian Moment of the 1970s*, New York: Berghahn.

Datta, V. (1999) *Birth of a National Icon: The Literary Avant-Garde and the Origins of the Intellectual in France*, Albany, NY: State University of New York Press.
Hanna, M. (1996) *The Mobilization of the Intellect: French Scholars and Writers during the Great War*, Cambridge, MA: Harvard University Press.
Jennings, J. (2011) *Revolution and the Republic: A History of Political Thought in France since the Eighteenth Century*, Oxford: Oxford University Press.
Judt, T. (1992) *Past Imperfect: French Intellectuals 1944–1956*, Berkeley and Los Angeles: University of California Press.
Poster, M. (1975) *Existential Marxism in Postwar France: From Sartre to Althusser*, Princeton, NJ: Princeton University Press.
Schalk, D. (1979) *The Spectrum of Political Engagement: Mounier, Benda, Nizan, Brasillach, Sartre*, Princeton, NJ: Princeton University Press.

17
THE POLITICAL PHILOSOPHY OF CHINA

Tongdong Bai

1. Chinese Political Philosophy?

It is both discouraging and encouraging that the *Routledge Companion* contains this one chapter of 6,000 words on Chinese political philosophy. It is discouraging because, unlike Western political philosophy in which each important figure gets an individual chapter, the entirety of Chinese political philosophy gets only one chapter. But it is encouraging because there is at least one chapter on Chinese political philosophy. For some—perhaps many more a decade or more ago—might think that there is no such thing as "Chinese political philosophy," and this chapter would thus be already one chapter too many.

The non-believers in Chinese political philosophy actually have a point. That is, many, if not most, classical Chinese texts are not written in the form of treatises, and they don't contain argumentation as extensive as most works in Western political philosophy. They often have the form of records of dialogues (between a master and his pupils), aphorisms, concrete political advice to the rulers, or policy debates with other ministers or counselors. They are often written not as purely theoretical discourses, and they lack apparent system and a sense of detachment. As the Chinese historian Qian Mu pointed out (1996), one important reason for these differences is that, from the so-called Spring and Autumn period on (i.e. after roughly 770 BCE), the Chinese intellectual elite often became part of the ruling elite, sharply different from the situation in medieval Europe. The Chinese intellectual elite in the past could thus put their political thoughts and theories into practice, and had little need and time to put them into theories that are detached from practice. In fact, in the opening paragraphs of *On the Social Contract* Jean-Jacques Rousseau made a point that supports Qian Mu's:

> I shall be asked if I am a prince or a legislator, to write on politics. I answer that I am neither, and that is why I do so. If I were a prince or a legislator, I should not waste time in saying what wants doing; I should do it, or hold my peace.
>
> (Rousseau 1978: 46)

Of course, this defense only explains why the writings of many Chinese political

thinkers are different from those in the Western tradition, and suggests that, if given opportunity (or if less involved in real-world politics), these Chinese thinkers could have written works that bear more resemblance in style to writings in the Western tradition. But we need to further realize that the apparent lack of system doesn't mean that there is no system in Chinese political writings. Rather, in the Chinese commentary tradition, a reader is required to "connect the dots" and grasp the implicit yet inherent system of reflections within the Chinese classics. This exercise may be very rewarding because, although their works lack leisurely speculations, many political writers of ancient China instead imbued their works with first-hand experiences of politics on the highest level.

Moreover, the claim that Chinese classical texts are not philosophical because they do not contain as much argumentation as Western philosophical works, even if it were true, is based upon an unexamined assumption that argumentation is the only way of philosophizing. If we take philosophy as a reflection on and examination of human experiences, we may accept the possibility that there are different ways for reflections and examinations to be expressed. Furthermore, the aphoristic and dialogical format of many Chinese classics can be taken as an implicit form of argumentation, an argument sketch that leaves the details for readers to fill in, with only important steps and key and difficult points made explicit. As Friedrich Nietzsche put it (1954: 40), "[i]n the mountains the shortest way is from peak to peak: but for that one must have long legs. Aphorisms should be peaks—and those who are addressed, tall and lofty." Therefore, the differences between Chinese classical texts and their Western counterparts may be more apparent than real.

2. The World "Classical" Chinese Thinkers Were Faced With

How can we select materials from a long tradition of Chinese political philosophy? First, we can leave aside works that had little influence or had some influence only for a short period of time, such as those of the Mohist school. Second, we can focus on philosophical thoughts in the so-called Spring and Autumn and Warring States periods (for short, SAWS; roughly from 770 BCE to 222 BCE), for China went through dramatic changes in these periods that set the basic political framework for the following 2,000 years of Chinese history, and the debates among political philosophers during the SAWS contributed to these changes. With these criteria, we can focus on three schools of thought, Confucianism, Daoism, and Legalism, and we can sketch their ideas and their historical and contemporary implications, leaving many finer points aside.

First, let us understand what changes China experienced during the SAWS. Before the SAWS, in the so-called Western Zhou dynasty, the political structure was a feudalistic, pyramid-like, and expanding system. The kings of Zhou (especially of the first generations) enfeoffed their relatives, loyal and competent ministers (many of whom were also the king's relatives), nobles of the past Shang dynasty, etc. These people became princes in their own principalities. Some of these principalities were in remote areas of the empire, and, in a way, they were colonies in otherwise "barbaric" areas (Qian 1996: 57). The establishment and expansion of these de facto colonies thus helped to broaden the imperial reach. When the principalities expanded, their rulers did the same as the kings did, enfeoffing their own relatives and ministers. In the entire empire, the king ruled over princes (of various ranks), princes over lesser lords, and so on. On each level, it was one master ruling over a limited number of subjects, making it possible for the

master to rule through personal influence and codes of conduct. But this hierarchical system was collapsing during the SAWS because familial bonds had been weakening after generations; as the empire expanded to the limit allowed by the technologies of that day, infighting became inevitable; and the contractual bonds also became less and less effective within ever-larger groups whose members had more and more in-group conflicts with each other. In the Spring and Autumn period, the king of Zhou was given nominal homage only, and eventually the boundaries of each principality were no longer respected. Through wars and conquests, seven large states emerged. This was the beginning of the Warring States period. In these periods, the rulers had to rule directly over states that kept becoming larger and more populous and the survival of these states and their rulers depended upon their physical strength alone.

I will not elaborate on the similarities and nuanced differences, but will make a rather bold statement instead: the transitions which these Chinese states experienced are comparable to the European transitions from the Middle Ages to (Western) modernity. In this sense, Chinese thinkers during the SAWS were already dealing with modernity issues in some form. In the feudal systems of China and Europe, on each level of the pyramid-like ruling structure, there was only a small community of a few hundred or a few thousand people. But the states kept expanding due to war and conquest as well as natural population growth. This might appear an insignificant change, but in politics "size matters." The pyramid-like feudal structure didn't work on a large scale (after all, the number of trusted relatives and friends is likely to be limited), and the social glue of the previous era, the noble codes of conduct and virtues that were based upon a comprehensive conception of the Good—which are likely to be maintained when the community is small—didn't work well in a large community. As a result, the old political structure collapsed, and a new one was desperately needed. In particular, the ruling class that was built upon nobility by birth was seriously weakened, and a new way of selecting the ruling class and a new social glue had to be discovered or invented for both the ruling class and the society as a whole. In short, a key distinction between antiquity and modernity lies in the change of the states' sizes and the consequences of that change. This is the background against which we should understand Chinese "classical" (in terms of the time) or "modern" (in terms of the problem) political philosophers.

3. Confucians: the Revolutionary Conservatives

One group of "classical" Chinese philosophers that tried to address the problem of "modernity" were the early Confucians, the most influential of whom are Confucius (551–479 BCE), Mencius (372–289 BCE) and Xun Zi (313–238 BCE). As was mentioned, the old social glue—*li*, that is, rites, rituals, and codes of conduct—became politically ineffective within the ruling class, due to the expansion of the states and the emergence of a society of strangers, and it didn't seem to be applicable to the relations between the rulers and the ruled. Classical Confucians, instead of doing away with the old political structure completely, tried to salvage it by reinterpreting it. They emphasized the significance and argued for the universal applicability of *li*, and offered a new foundation for it, or a new social glue, *ren*, that is, humanity (in the sense of being humane and compassionate to others). According to early Confucians, the reason that we observe certain codes of social conduct is that we have compassion for others. The external practices are expressions of our internal moral sentiment. Mencius offers a very vivid and powerful illustration of this moral sentiment:

> Now, when men suddenly see a child about to fall into a well, they all have a feeling of alarm and distress, not to gain friendship with the child's parents, nor to seek the praise of their neighbors and friends, nor because they dislike the reputation [of lack of humanity if they did not rescue the child]. From such a case, we see that a man without the feeling of commiseration [i.e. compassion] is not a man.
>
> (*Mencius*, 2A6; Chan 1969: 65)

A few things should be noted here. First, Mencius doesn't say that all people will act on an instinctive moral sentiment of compassion. Acting compassionately depends upon a careful and assiduous cultivation of moral sentiments. Second, some might argue that this instinctive feeling of compassion is not as universal as the above story wants us to believe. For example, a child witnessing this scene might not have a feeling of alarm and distress, but Mencius could say that this is not because the child doesn't have compassion, but because, being immature, he or she doesn't understand the consequences of the other child's falling into a well. But what about the sociopaths who understand the consequences of a dangerous situation like this and, precisely because of their understanding, enjoy it greatly? Mencius could argue that they have the heart of compassion, but it is covered up (lost) because they have failed to cultivate this feeling. But if these people never had even the heart of compassion, Mencius would simply say that they are not human. That is, for Mencius, "human" is not merely a biological concept, but a moral concept that has a clear social dimension. Indeed, Mencius even implies (*Mencius*, 3A4) that a "human being" who doesn't have certain social relations is not really a human being. In other words, the noble savage, the autonomous individual, for Mencius, is not really a human being, but a human look-alike. In short, we have to understand Mencius's account of human beings as a normative account which, nevertheless, is descriptive of most human beings.

Second, from both the common meaning of the term "compassion" and from Mencius's story, we can see that compassion is a feeling for strangers, which is to replace the old social glue that was only effective in a close-knit and small society. It might be of some interest to note that compassion is not one of the four Greek cardinal virtues, and it doesn't seem to have been valued by ancient Romans either. In this sense, Nietzsche was right to claim that pity as a virtue was introduced later in Western history (1994 and 2002; in particular, see, for example, 11–37 (the First Essay) of 1994). But Nietzsche was perhaps wrong in his diagnosis of this "symptom." After all, there was no Jewish slave revolt in Chinese history, and the reason for compassion to be valued may have been the need to find a social glue for a large society of strangers (whereas in the Greek *polis* which was comparatively small, friendship served as an effective social glue). Besides, that pity becomes a virtue during Nietzsche's times might further support my claim that China's changes during the SAWS are analogous to Europe's during modernization of the nineteenth century.

Now, how do we cultivate compassion or humanity? Confucius's answer is "to take as analogy what is near at hand" (*Analects*, 6.30). What is commonly near at hand is one's own self and family. In his answer to a king's concern that he (the king) is fond of money and women, Mencius says that this fondness is important in enabling the king to understand that his people must have the same fondness as he does, to take care of his people's needs, and thus to be a good king (*Mencius*, 1B5). To rule requires compassion, and family is a good place to cultivate this expansion of care. For family has dual

roles in one's life. On the one hand, family is what most people feel akin to and thus what is commonly categorized as belonging to the private; on the other, care for family members is for most people the first step to go beyond one's personal self and toward others or the public. Thus, if we focus on the latter aspect of family, it will be possible for people to expand the sense of familial care outwards, eventually embracing everything in the universe. This ever-expanding circle makes universal love possible. But it should be noted that, for Confucians, it is still natural and justified for even a person of universal love to care about the closer ones more, such as trying to save his or her drowning mother before saving anyone (or anything) else. In short, the goal of the Confucian model of moral cultivation is to love universally and unequally (hierarchically).

One target of this Confucian model is the conflict between the private and the public, which is a long-standing theme in political philosophy and a daunting concern for all political thinkers of the SAWS, when, due to the collapse of the old feudal structure, it was a world of "all against all." But instead of denouncing the private (or the part of the private that is in conflict with the public), as Plato's *Republic* apparently and Han Fei Zi actually did, early Confucians realized that there is something in the private sphere that is constructive in the public. So, their solution to the conflict is to cultivate this constructive part in order to use it to suppress the part that is indeed conflicting. For they understand that it is impossible to do away with the private sentiments of preferring our private concerns. Then, instead of attempting this impossible task, we should turn these preferences into something constructive and complementary for our contribution to the public sphere.

Therefore, Confucians try to find a middle ground between the unabashed celebration of self-interest and the total negation of the private. This strategy not only applies to the conflict between the private and public in domestic politics, but also has wider and contemporary implications. For example, this middle way offers a basis for the Confucian understanding of international relations (including just war, international aid, etc.). It is natural and justified that one loves one's own country more than other countries. But this doesn't mean that one should thus practice *realpolitik*, paying no attention to the interest of other peoples or other states. After all, peoples from other states are also people, and humane treatment is required, and we should therefore win them over by our humanity and by being a moral exemplar before resorting to war. Influenced by this Confucian understanding, rulers of traditional Chinese states were rarely praised, and were often criticized, if they were keen on waging war. This understanding offers an alternative to both *realpolitik* and cosmopolitanism. For Confucians, the former is too cruel, and the latter is too naive and will end up as empty claims (loving everyone equally often results in loving no one). The Confucian solution, in contrast, offers a realistic utopia.

Following a similar rationale, when dealing with environmental and animal rights issues, Confucians may offer an attractive alternative. For Confucians, these issues are important not for the sake of the environment and the animals themselves, but for the sake of human beings' being human (physically and spiritually).

Another area in which the Confucian thinking can make a contribution is equality between men and women in politics (which is not to say that early Confucians themselves were in favor of this equality). As implied in Plato's *Republic*, even if men and women have roughly the same potential to participate in politics, they can't be equally competent in actuality if women have to spend years raising children, which is still a fact of life today. The underlying hypothesis is that spending time in private life is use-

less, if not detrimental, to public service. However, Confucians deny a simple break between the private and the public. Indeed, they defended the rite of three-year mourning (i.e., one should excuse oneself from public duty for three years if one's parent dies), presumably on the basis that to spend three years remembering one's parent (and thus reinvigorating one's filial love, etc.) might help a politician to become a better leader than he or she otherwise would be. Similarly, they can offer a defense of female politicians' spending a few years raising their children.

The Confucian "middle-way" strategy is also a salient feature in their treatment of political equality. One crucial issue during the SAWS is to find a new mechanism to select the ruling class. Nobility by birth was largely gone (other than the inherited throne), and early Confucians were rather revolutionary in advocating some form of egalitarianism that opens the door of politics to all. For example, Confucius was said to be the first private teacher, publicly offering to all those lessons of statecraft which had hitherto been reserved to the nobility. Both Mencius and Xun Zi went a step further to claim that every human being has the potential to become a sage-ruler. Early Confucians also made it abundantly clear that the legitimacy of the state lies in the satisfaction of people's interests. For example, according to Mencius, the throne is given to the next ruler by Heaven, and "Heaven sees with the eyes of its people; Heaven hears with the ears of its people" (*Mencius*, 5A5). To be clear, the interests that the government needs to satisfy include not only the people's short-term material interests, as today's states often focus on, but also their long-term and future generations' material interests and their spiritual needs (such as educating the people about the meaning and significance of basic human relationships). Although early Confucians emphasized the role of morality, they understood that, for common people, economic stability was crucial to their morality. Thus, if the government fails to secure basic goods for the common people, they may go astray. It then would be entrapment if the government punished them (*Mencius*, 1A7 and 3A3). It should be noted that, for Mencius, everyone, in spite of the environment, can stay moral. So, for Mencius, both the criminal who goes astray and the government, if they have failed to secure basic goods, should be held guilty for the crime committed. Moreover, if the ruler keeps failing to do his job to satisfy his people's needs, it is justified to remove him, even violently, from office (*Mencius*, 1B6 and 1B8).

So far, early Confucians sound, aside from the fact that they consider government a source of good and not a necessary (or even unnecessary) evil, a lot like Enlightenment or even democratic thinkers who advocate equality. But there is another side of early Confucian teachings. According to Mencius, although the satisfaction of the people's interests is the ultimate ground of legitimacy of a state, the popular will is not the sole deciding factor when political decisions are made. Rather, the involvement of the ruling elite (the "meritocrats") is necessary. The ruling elite, which is to replace the nobility by birth that constituted the ruling class before the SAWS, should be wise and compassionate ("nobility by merits"). However, although Mencius believed in equal potentials of all human beings and advocated upward mobility, he also seemed to believe that those who can actually become members of the ruling class are always few in number, in spite of governmental efforts to satisfy economic needs and to offer education to everyone (which are, for Confucians, duties of the government). If the masses fail to meet the criteria of being wise and compassionate, and have to or choose to work outside politics, they can't understand politics well, including whether a policy is good for them. In this case, their voice should be balanced by the political involvement of the ruling elite

who are selected from the people when they prove themselves to be wise and virtuous. Therefore, the ideal government for early Confucians should be a hybrid of two components: the popular (democratic) and the elite (meritocratic) elements.

The terms "elite" and "meritocracy" might sound offensive to today's democratic ears because, in the West, they often carry a meaning of inherited status, and lead to a sense of disenfranchisement and resentment. But since the SAWS, thanks to Confucians and other thinkers, upward mobility has been deep in the Chinese psyche. Although few people can actually become rulers, the door is nevertheless open to everyone and the government has the duty to help everyone to make it. Moreover, for Confucians in particular, the justification of a person's politically superior status comes from their ability to serve the inferior. The rationale behind their thinking is that, if political inequality is inevitable, instead of hopelessly and sometimes counterproductively getting rid of it, we should turn it to the service of the disadvantaged. Using Rawls's idea of the difference principle, according to which economic inequality is tolerated if it serves the most disadvantaged the best (Rawls 1971: 75–83), we can call the Confucian idea a "political difference principle."

The Confucian hybrid regime could have implications for today's political world. The political ignorance of the common people in today's democracies has been an established fact in political science. This and other facts make the popular vote alone a bad way of selecting legislators. Some democratic thinkers try to correct the wrongs from within, e.g. by improving economic equality and education, informing the voters, getting rid of the influence of money in politics, etc. But some might argue (as I did in my own works—see, e.g., Chapter 3 of Bai 2009) that the problems with the popular vote are rather deeply rooted in the modern world and can't be solved from within the present democratic system. If this argument stands, then we can see the merit of a hybrid regime. It preserves the good elements of popular democracy, but tries to balance democracy's over-emphasis on popular will with meritocratic arrangements. If modern democracy appeared as a way to counter the abuse of the nobles, maybe it is time to counter the abuse of the uninformed and not-so-virtuous masses with the meritocrats. (It is true that there are still meritocratic elements in today's democracies, but they might have been not strong enough.) Not only did early Confucians offer such a hybrid model, but in China's past, later thinkers and politicians tried different ways to select the competent and the compassionate and to use these meritocrats to counter the power of the allegedly absolute monarch (the emperor). (Thus it was a (common) mistake to call China's past regimes authoritarian or even despotic.) If we replace the emperor with the popular will, the above attempts can become models for us to reflect upon and to experiment with in today's politics.

4. Daoists: Return to the Age of Innocence

Early Confucians embraced modernity through their reinterpretations of antiquity, and are confident in the human capacity to solve the problem of modernity. Early Daoists, by contrast, have a fundamentally pessimistic view of human capacity. Instead of intervening in Nature, human beings might be better off through non-action. A few clarifications need to be made here. Chinese philosophy is often accused of advocating quietism, probably because of many people's failure to understand the idea of non-action, which doesn't mean taking no actions, but rather taking no *unnatural* actions. This leads to the meaning of the natural. Here lies another common misunderstanding of

Chinese philosophy. That is, many believe that the ideas of non-action and naturalness are quintessential Daoist ideas. But the idea that Nature should be the inspiration of and guidance for human actions was perhaps first introduced in the *Analects*, where the ideal ruler is described as someone who practices non(unnatural)-action and leads the state by being a moral exemplar. Indeed, even the Legalists model politics after Nature, and when political institutions are well established (i.e., follow natural laws), the ideal ruler should also commit to non(unnatural)-action. So, the issue is really what "Nature" means to these different schools of early Chinese political philosophy.

The founding text of Daoism, the *Laozi*, is a rather mystical text full of aphorisms and hard to decipher. So the following is only one possible way to make sense of it. It seems to argue that what is natural for a human being is to satisfy the bare necessities, and what is unnatural is human beings' attempts to compete for more than these necessities … for example, delicacies instead of mere food, mansions instead of simple shelters, beautiful women instead of any woman, etc. All social and political ills result from people's competitiveness arising out of their unnatural desires. The assumption seems to be that, if we are content with satisfying our bare necessities, we can live peacefully with each other, without deceit, crime, war, and other social ills. This is somehow guaranteed by Nature since our desires are so far natural. When social ills arise from human (vis-à-vis natural) desires, many (Confucians included) try to right the wrongs by making various human efforts. But compared with the great Nature, we human beings are rather insignificant and incompetent. Nature is indifferent to human desires (to think otherwise, such as imagining a God who cares about human beings, is humans' hubris and foolishness), and humans are so incompetent that our attempts to cure these social ills can only backfire. For example, the Confucian celebration of (Confucian) virtues is implicitly encouraging more competition (to become virtuous), which is precisely the root cause of social and political problems. The Confucian cure, then, is like what a Chinese proverb says, "quenching thirst by drinking poisonous wine." Instead of curing social diseases, the Confucian solution actually perpetuates them, on this account.

Then, what is the solution? The *Laozi* seems to be as elitist as most early Confucian sources. For it believes that only a Daoist elite can see the problem of letting desires run wild, and can thus choose to return to the stage of innocence. The masses can at best be made to stay in the stage of innocence. But there is also an anti-elitist side of the *Laozi*. That is, the masses wouldn't be able to pursue unnatural things if not for the lustful elite, the crooks who "invented the wheel." There are a few suggestions in the *Laozi* about how to get rid of the elite while not falling prey to human hubris and taking (unnatural) actions. But this proves to be difficult during the SAWS, when desires (and the crooks) were already running wild. The *Laozi* may have realized that desires' running wild is an inevitable result of society's becoming large and populous, and no cure is available that won't backfire. But in a small and closed community, there is little for people to emulate, and there is little concentration of wealth, which makes luxuries and inventions difficult and even impossible to devise and to maintain. Thus, the ultimate solution is to find a way to go back to a stage of small, closed states with few people, which is described in Chapter 80 of the *Laozi*. The ideas presented here might remind some readers of the Romantic thinkers, such as Rousseau (who makes similar arguments in his *Discourses*). This similarity, again, might indicate the similarity between the transitions experienced by Rousseau and by the authors of the *Laozi*. (Who they were is a controversial issue.)

How, then, can we go back to that stage? The path that the *Laozi* offered, albeit only

through subtle innuendos, is to let the ambitious fight against each other. Eventually, they will kill each other off, and technologies and luxuries will be destroyed with them. Then the Daoist elite who have "lain low" so far can rise up, and maintain society in a stage in which each state is small and closed and no (unnatural) action needs to be taken to contain the bad. This apparently happened in China. The most fierce state during the Warring States period, the state of Qin, conquered all other states, and unified China. But it soon collapsed. The first few rulers of the dynasty after Qin (the Han) allegedly followed Laozian political ideas, including a policy of "allowing the people to rest."

Unfortunately, this stage didn't last long (if it actually existed in the first place), and Chinese society soon went back to the path of development (and, for Daoists, of doom).

The other founding (and likely later) text of Daoism, the *Zhuangzi*, seems to appreciate the hopelessness of the primitivist solution in the *Laozi*. Thus, it takes a radically apolitical attitude toward the world. Instead of saving mankind, enlightened Daoists can and should only try to save themselves by avoiding politics and enjoying the freedom with Nature.

5. The Legalists: Builders of Modern Bureaucracy and Institutions

Finally, we come to the so-called "Legalists." I will focus on the most important thinker of this school, Han Fei Zi. He was said to be a pupil of Xun Zi the Confucian, but became perhaps the most ferocious critic of Confucianism. According to him, even the wisest people can't agree on the meanings of subtle moral teachings, and even Confucius himself could only convince some seventy people to follow him closely. So, a large state can't be built upon a comprehensive understanding of the Good (and a group of virtuous people). This might seem like an early argument for pluralism. But Han Fei Zi only argues for the inevitability of pluralism in a large state, and his conclusion is not to celebrate it, but to declare that we need, in order to maintain stability, to look somewhere else to find a unifying glue for a large state with many diverse people. (The stability issue is still a concern for some liberal thinkers today, and it may be the pluralists who owe us an explanation of what elements in today's world make it possible for us to have pluralism and stability at the same time.) Moreover, although he acknowledges the fact that we could be kind to others, Han Fei Zi thinks that this kindness can't be cultivated to resist any serious challenge in a political setting.

Thus, according to Han Fei Zi, the Confucian moral teachings are completely useless to regulate the people of a large state. It is laws and institutions which are effective. They are built on an understanding of what all human beings must bow down to, that is, rewards and punishments. But unlike Jeremy Bentham (and some other utilitarians) who made a similar claim, Han Fei Zi acknowledges the reality of moral sentiments (instead of accusing them of "the principle of utility misapplied" (Bentham 1948: 13)). He also argued, however, that these sentiments are too weak to be politically effective in a large state, and he therefore understood that morality will be ineffective when states become too large and too populous. These differences make Han Fei Zi's position subtler and stronger than some utilitarians'.

For Han Fei Zi, laws and institutions can regulate people through rewards and punishments. The latter should be based upon verifiable results truly beneficial to the state instead of beautiful words, the ruler's own personal preferences, or "virtuous"

deeds (in a Confucian sense). What a state needs, which is verifiable, is for its people to harvest as many crops and chop off as many enemies' heads as possible. Like the Confucians, Han Fei Zi tried to invent a new way to fill up the ruling class, and his ideal is also a meritocracy that is open to all. Everyone should be treated equally and measured by their merits alone. The key difference is that Han Fei Zi doesn't think that Confucian virtues and wisdom, but rather military and agricultural achievements should serve as the criteria of selection. For Han Fei Zi, virtues are not only useless, but also a dangerous distraction to the state's interests. To establish virtues as an independent criterion of politics weakens political institutions and laws, and challenges the authority of the ruler. Indeed, many believe that Han Fei Zi argues for the absolute power of the sole ruler. But he also clearly and repeatedly warns against the ruler's using his own preferences and understanding of politics to interfere with the functioning of institutions. In this sense, both the ruler and the subjects are faceless parts of a well-functioning automaton. If the ruler tries to use his own political wisdom and preferences rather than what is dictated by the natural laws of politics, his state is doomed to be weakened and eventually destroyed. The ruler should guarantee people's equality before state laws (which are based upon natural laws of politics), and, although, being the guarantor of them, he himself is not subjected to these state laws, he *is* subject to the same laws of nature as the state laws are.

Han Fei Zi's idea that the influence on politics of rulers and of the Confucian virtues should be limited is based upon a premise that is akin to early Daoists' ideas. That is, for Han Fei Zi, we human beings are incompetent in comparison with Nature, or Dao (the Way). There are many passages in the *Han Fei Zi* that suggest his profound appreciation of the *Laozi*. But for the *Laozi*, Dao is mystical, and the final solution is for human beings to go back to a stage of small states with few people. It seems, however, that Han Fei Zi believed that he discovered how to make Dao work to provide explicit guidance of politics in "modern" circumstances. By following the natural Dao of politics, and by curbing our hubristic urge to apply human wisdom and virtues to politics, a large state can be run without the risk of falling into chaos.

6. Final Remarks

The aforementioned three schools have played significant roles in traditional China's politics. After devastations, rulers sometimes adopted a Daoist policy of *laissez-faire*. Most other times, however, against Han Fei Zi's and perhaps some Confucians' original intentions, the imperial policies are a hybrid of Legalist and Confucian teachings, celebrating Confucian moral values and institutionalizing them in a Legalistic manner. The fact that this mixture worked challenged Han Fei Zi's view (and perhaps some democratic and liberal thinkers' ideas) that, in the age of inevitable moral pluralism, morality can at best retreat to the private sphere. Maybe there can be a thinner version of morality that can get around the curse of inevitable pluralism.

Political ideas from China (meritocracy based upon equal access, bureaucratic system of civil servants, secular yet moral teachings as the foundation of a state, etc.), perhaps only in their simplified and distorted forms, inspired early Enlightenment thinkers (which may have been resulted from the similarities between the China's transitions during the SAWS and Western modernization). Maybe it is time to look at traditional Chinese political philosophy closely, in order for us to understand the present and the self by understanding the past and the other.

Acknowledgments

I wish to thank Fred D'Agostino and Jerry Gaus for their great editing job, and Joseph Chan for recommending me to do this job. The research for this chapter is supported by "The Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning," the "New Century Excellent Talents in University" grant, by Shanghai Educational Development Foundation (Shuguang Project), by the "Pujiang Talents" grant from the Shanghai Government, by the funding for the Shanghai Philosophy and Social Sciences Projects, and by the research grants from Fudan University (Guanghua, two 985 projects (2011RWXKZD009 and 2011RWXKZD009), etc.).

Related Topics

Plato, Rousseau, Continental Political Philosophy, Pluralism, The Difference Principle, Global Justice and Politics, War, Democracy

References

Bai, Tongdong (2009) *The New Mission of an Old State: The Contemporary and Comparative Relevance of Classical Confucian Political Philosophy* [in Chinese], Beijing: Peking University Press.

Bentham, Jeremy (1948) *An Introduction to the Principles of Morals and Legislation*, New York: Hafner Press.

Chan, Wing-Tsit (1969) *A Source Book in Chinese Philosophy*, Princeton, NJ: Princeton University Press.

Nietzsche, Friedrich (1954) *Thus Spoke Zarathustra*, trans. Walter Kaufmann, London: Penguin Books.

—— (1994) *On the Genealogy of Morality*, ed. Keith Ansell-Person, trans. Carol Diethe, Cambridge: Cambridge University Press.

—— (2002) *Beyond Good and Evil*, ed. Rolf-Peter Horstmann and Judith Norman, trans. Judith Norman, Cambridge: Cambridge University Press.

Qian, Mu 钱穆 (1996) *Outlines of the History of China* 国史大纲, Beijing: Shangwu Press.

Rawls, John (1971) *A Theory of Justice*, Cambridge, MA: Harvard University Press.

Rousseau, Jean-Jacques (1978) *On the Social Contract with Geneva Manuscript and Political Economy*, ed. Roger D. Masters, trans. Judith R. Masters, New York: St. Martin's Press.

Further Reading

Bai, Tongdong (2011) *A New Mission of an Old State—The Comparative and Contemporary Relevance of Classical Confucian Political Philosophy*. See also "Preliminary Remarks: Han Fei Zi—First Modern Political Philosopher?" *Journal of Chinese Philosophy* 38: 4–13. (This article contains a detailed argument on the similarity between China's transitions during the SAWS and Western modernization, as well as a more detailed discussion of Han Fei Zi's ideas.)

—— (forthcoming) *China: The Middle Way of the Middle Kingdom*, London and New York: Zed Books. (For a far more elaborate version of this chapter.)

Chan, Wing-Tsit (1969) *A Source Book in Chinese Philosophy*, Princeton, NJ: Princeton University Press. (To study Chinese political philosophy requires an exposure to Chinese classics. This anthology has a nice selection of these classics.)

Fung, Yu-lan (1966) *A Short History of Chinese Philosophy*, New York: The Free Press. (Perhaps the best introduction to Chinese philosophy.)

Kung-chuan Hsiao (1979) *A History of Chinese Political Thought, Volume One: From the Beginnings to the Sixth Century* A.D., trans. F. W. Mote, Princeton, NJ: Princeton University Press. (A comprehensive history of Chinese political thought which is not yet translated into English in its entirety.)

Schwartz, Benjamin (1985) *The World of Thought in Ancient China*, Cambridge, MA: The Belknap Press of Harvard University Press. (Another nice introduction to ancient Chinese thought.)

18
INDIAN POLITICAL THEORY

A. Raghuramaraju

1. Rta and Dharma

Searching for the earliest political order in ancient India, one comes across two important concepts, *Rta* and *Dharma*. They are found in the *Rig Veda* and recur regularly in subsequent texts. Rta literally means "the course of things" and it is the order of the universe. It is the preordained divine natural order that is self-existent. Even gods abide by it as they are created by it. The primary duty of everyone including human beings is to follow this order. This order is so predominant and overriding that all other kinds of orders including causal, regulative, hierarchical, serial and sequential, social and political orders flow from this order. So this is the first theoretical account of an order of which the political is a mere part. Though permanent in nature and beyond change, this order nevertheless appears in different manifestations. This is perhaps one of the earliest sources where the same being appearing differently is accounted for.

However, Rta subsequently comes to have a moral dimension. Morality or dharma enables human beings to follow and conform to this transcendental order. There are different accounts of dharma rather than one single definition. Its root is *Dhr*, meaning "to sustain." In Taittiriya Aranyaka, it is said that:

> Dharma is the foundation of the whole universe. In this world people go unto a person who is best versed in dharma for guidance. By means of dharma one drives away evil. Upon dharma everything is founded. Therefore, dharma is called the highest good.
>
> (Dandekar 1963: 220–1)

This gets repeated in many other works including the epic Mahabharata. Further, pointing out the variance in the meaning of dharma from one epoch or *yuga* to another, Manu says that during the *Krita Yuga*, it stands firmly and is followed completely. Subsequently, because individuals behave contrary to dharma, it deteriorates gradually. In the *Treta Yuga*, it loses one leg, two in *Dwapara* and still one more in the *Kali Yuga*. Accordingly, Manu prescribes different dharmas to different yugas. Penance is the highest dharma in Krita, sacred knowledge in Treta; offering sacrifices is the highest dharma in Dwapara and giving away what one has in the Kali Yuga (Manu 2000: 12). Here it is worth noting that after positing an independent and perfect order in the form of Rta, these early sources reveal a gradual tendency to deviate from it, and they seek to

work out ways in which humanity can strive to restore that perfect order. This idea is also found in the Mahabharata where it is said that, whereas there was, originally, neither kingdom nor king, indeed no government, because people governed themselves by means of dharma, this state was corrupted so that, subsequently, *danda* became necessary, meaning punishment for those who defy or don't follow dharma (Badrinath 2007: 422–3).

Crucially, on this account, political and moral orders are subservient to the divine natural order. This both sets limitations and gives identity to moral and political philosophy in classical Indian thought. The given, that which is beyond change, prevails over individual and social choices and actions. This is not so much, as it is sometimes taken to be, a doctrine of determinism. Rather, it is a doctrine about the fittingness of conformity to a permanent and unchanging order. This has far-reaching implications as in this discourse the emphasis is on conforming to that which is given rather than on making individual choices. The influence of this crucial doctrine is, until the advent of modernity, felt on all subsequent theories of the political. Prior to modernity, then, the emphasis in Indian political thought is on equilibrium, coordination, permanence, adherence to the given rather than change, invention and development.

For instance, strict adherence to *sastras*, particularly the Vedas—obeying mother, father, teacher and God—brings privileges over choice-related aspects such as friendship. Manu declares that the "trouble that a mother and father endure in giving birth to human beings cannot be redeemed even in a hundred years." Thus, he declares that one "should constantly do what pleases the two of them, and always what pleases his teacher. He should not assume any other duties without their permission" (Manu 2000: 40). Even Kautilya in his Saptanga theory of the state, while including mitra, an ally, as the seventh element, wants the ally to be a long-term associate of the family, thereby emphasizing continuity, or a form of permanence. All these go to establish that the earlier theories emphasized conformity to an order that pre-dates everything. So the purpose of human existence including political activity is to strive continuously to devise an individual and social path that conforms to this order. So, strictly speaking, there was, during this early period, no idea of politics that is independent of this transcendental order.

While this forms the traditional account, there were of course small-scale societies that were ruled by community chiefs. These societies were governed by customary rules or via physical power and did not make a significant mark outside their territories.

Let us next consider an important approach to organizing individual and social existence, namely, the *purusharthas*. What is important to note here is that rather than provide a scheme of political organization, traditional thinkers, in developing the purusharthas, laid emphasis on individual liberation, and treated the social and political as facilitating this.

2. Purusharthas

There are four purusharthas: *dharma*, *artha*, *kama* and *moksha*. Dharma is righteousness. Artha is material possession. Kama is the enjoyment of this-worldly sensuous pleasures. Moksha is seeking spiritual liberation, thereby attaining salvation. These classifications establish an account of the individual and his objectives that incorporates both the social and the transcendental. The four purusharthas are interrelated and do not operate in isolation. Their relationship is hierarchical though not necessarily unidirectional.

Dharma is of two kinds: *varnashrama dharma* and *samanya dharma*. The former is specifically related to each varna and ashrama and the latter is general and universal. The universal dharma consists of *ahimsa* (non-injury to all living beings) *satyam* (truthfulness), cleanliness, freedom from malice, compassion and tolerance.

The other two, namely, artha and kama, are not autonomous and are placed under the supervision of dharma and are directed towards moksha. For instance, kama, which is seeking sensuous pleasure, is not ignored. Pleasure is not to be obtained by following individual desires and wants but is related to other purusharthas. If it is left unsupervised it can relapse into a lower level. It is in fact regulated by dharma that allows only ethically sanctioned enjoyment. This denial of individual desires and the dominance of discipline came in for severe attack subsequently, especially in literary genres. However, it must be noted that the sensuous enjoyment, though supervised by other dharmas on this account, is not reductionist or monotonous. There are erotic texts such as *Kamasutra* by Vatsyayana, *Ananga Ranga* of Kalyana Malla that reveal a wide variety in modes of sensual enjoyment. This controlled supervision by dharma elevates kama. Yet kama is not dharma; in a limited sense it is enjoyment but this is subordinated, teleologically, to dharma and moksha. It is regulated in the sense that a man can have sexual relations only with his wife and within the institution of marriage but not with any woman or women he desires. The Bhagavadgita opines that kama and dharma are commensurable and that the satisfaction of desire need not be inconsistent with the principle of righteousness.

Whereas Dharma Sastras centralizes dharma, Artha Sastra, while not rejecting the primacy of dharma, tends to restore more importance to artha and kama (Kautilya 1992: 122).

Moksha is liberation. There are different paths to achieve moksha. They consist of: *karma yoga*, *jnana yoga*, *bhakti yoga*, etc. While one can complain about not finding more acknowledgment of change in classical Indian thought, what cannot be ignored is that these texts did take into account a good balance between change and permanence, subsuming the former within the latter. It is this combination that provides a desirable equilibrium between progress and contentedness. The emphasis here is on harmony and equilibrium rather than on mere equality and progress.

Though not directly so, but at least indirectly, the political seems to be closer to dharma. It is the job of the ruler to maintain dharma. In addition to this classification that lays the path for the individual, incorporating individual, social and transcendental, there is another cluster that is predominant, this time in classifying different stages within an individual's life, each station again pointing towards transcendence.

3. Ashrama Dharmas

Yet another social stratification that is available and became prominent is that of the different *ashrama dharma*, or ranks or positions. They are:

- *Brahmacharya* or the student who has to study the Vedas, strictly follow his teacher, and live only on alms;
- the *Grhastha* or the householder who has to perform the duties assigned by scriptures, pursue his own profession, marry a woman from the same caste but from another gotra (deriving from the same primordial paternal ancestor), produce children within the institution of marriage, and, imperatively, attend to the needs of those who are dependent on him rather than seeking his own pleasure;

- the *vanaprastha* or the hermit who withdraws into the forest and spends time in austerity which includes celibacy, wearing only deer skin rather than woven clothes, consuming forest products rather than depending even on the charity of others, and spending time in doing penance; and, finally,
- the *sanyasi* or the ascetic who renounces everything and seeks union with the Supreme God. He must have complete control over his senses, withdrawing from active life, renouncing all worldly attachments (including not staying in any one place), and seeking inner purity.

An interesting thing about these ashrama dharmas is that there is at least no sequential order amongst them with regard to the last ashramas. For instance, it is possible to remain as a Grhastha and achieve the detachment of the hermit or sanyasi. These four stages of life are organized in such a way that they realize four ends in life. This shows how the socio-political is not construed as a self-sufficient purely mundane state of affairs but, rather, as a contributing factor in human beings' achievement of the ultimate purpose of life.

Highlighting the ultimate purpose of human existence, the Bhagavadgita identifies the following as the virtues of an ideal man. He is (Dandekar 1963: 298–300):

- one who renounces all the desires that arise in the mind,
- who is contented within his own Self,
- whose mind is unperturbed in the midst of sorrows and
- who entertains no desires amid pleasures;
- he is one from whom passion, fear, and anger have fled away—
- he is called a sage of steadfast intellect.

The distinguishing objective of these classical teachings is to work out a harmonious relation between individuals' salvation, which is primary, and the facilitating role of social and political organizations. From these moral and transcendental classifications let us turn to an elucidation of the different parts of the state that is available in classical Indian political theory, namely, the *Saptanga* theory of the state.

4. The Saptanga Theory of the State

The Saptanga theory of the state presents a consistent and coherent organization of various parts of the state envisaged by classical Indian philosophers. It consists of: the king, ministers, *janapada*, fort, treasury, army, and allies.

(i) Elucidating the qualifications of a king the Saptanga theory holds that he should have been born in a high family in an auspicious time, endowed with character, should seek guidance from experienced and wise elders, be religious, truthful, consistent in his behavior, strong and energetic, and should not delay or defer his duties. Further, he should have clear perception, retention, analytical knowledge, critical acumen, strong memory, a capacity to handle difficulties, and presence of mind with regard to time and place; he should feel ashamed to do anything evil, be skilled in discriminating between different conditions, know when to show mercy to his enemies and when to resist them depending upon their strength and weakness. He should be free from passion, anger, greed and obstinacy. According to *Artha Sastra*, a king should, in addition to the above,

also keep away from another's wife, nor covet another's property. He should be able to manage his economy well. The happiness of his subjects underpins his own happiness. According to this treatise, the king should, except in cases of urgency, attend to the needs of his people in an orderly way, not postponing action on them (Kautilya 1992: 119–25).

Further, the king should have the qualities of eight deities. Like Indra who showers rain, he should shower benefits on his people; like the sun who sucks up moisture he should be able to extract taxes; he is to penetrate everywhere with his spies like the wind; with the rod of chastisement (*danda*) he is to control all his subjects like Yama, the deity of death, subdues all in the end; he must punish the wicked as Varuna binds sinners with his rope; he is to gladden his subjects by shining upon them as the full moon gladdens men; he is to visit criminals with his anger and destroy wicked subordinates as fire burns all; and he is to support his subjects as the earth supports all creatures (Ghoshal 1959: 132). (Most of these qualifications of the king are also found in the Ramayana, especially the *Aranya Kanda* of this epic where Rama tells his brother Bharata, who goes to meet him in the forest, about the duties of the king (Valmiki 2000: 206–9).)

(ii) A minister should discharge different services including civil administration and revenue. He should be, among other things, native, born of high family, influential, trained in arts, endowed with foresight, bold, eloquent, possessed of enthusiasm, dignity, endurance, etc.

(iii) Janpada refers to both territory and population. The territory should be firm in the midland, should be capable of providing affordable subsistence to its own people and if necessary to outsiders also. It should look beautiful, be rich in arable land, mines, and timber and elephant forests, with pastoral land to graze cattle. However, it should not depend entirely on rain and should have facilities for markets. The terrain should not be muddy, rocky, salty, uneven or have thorny tracts or forests infested with treacherous and wild animals. It should be easy to defend, capable of subduing invaders. It should be able to support an army and provide for surpluses for taxation, be conducive to hard-working farmers, obeying subjects, and a population largely consisting of lower classes, namely, the economically productive classes, namely, vaishyas and sudras (Kautilya 1992: 152–7). (These hierarchical societies found it convenient to have obedient subjects whose job was to produce and who had no political participation or political ambitions.)

(iv) The fort shows itself in defensive fortifications on all the boundaries of the kingdom. This may consist of water-forts, mountain-forts, desert-forts, and forest-forts (Kautilya 1992: 157).

(v) Treasury contains lawfully inherited property from the king's ancestors and those that are earned by the king himself. It mainly consists of gold, silver, and is full of manifold, big precious stones and bars of gold. It should be able to endure a calamitous drought even of a long duration (Dandekar 1963: 250).

(vi) In order to maintain continuity with the past and to ensure stability, the army should also have been inherited from the father and grandfather. It should be constant in its loyalties and obedient. The sons and wives of soldiers should be contented and

well provided for. The army should be equipped for military campaigns in foreign lands; they should be capable of bearing pain, experienced in many battles, and expert in the science of all the weapons of war. They should regard the rise and the downfall of the king as equivalent to their own and consequently not engage in double-dealing with him (Dandekar 1963: 250). Kautilya allocates this job to the Kshatriya caste, whereas Manu considers that in the event of need and emergency other castes such as Brahmins and Vaishyas but not Sudras can also do the job of the army.

(vii) The allies should also be inherited from father and grandfather. They too should be constant in their loyalties, obedient, capable of preparing for war on a large scale and quickly—this is the perfection of the allies (Dandekar 1963: 250–1). As already referred to, Kautilya prefers the ally to be hereditary as this according to him ensures continuity, assures security and does not give rise to any trouble, conflict, or sabotage.

This order of society is said to ensure the safety of the people. Setting the chronological order of stable political orders, the *Santi Parva* of the Mahabharata says that once the state was established on a firm basis, it came to be regarded as the greatest safeguard of one's wife and property. Therefore, it was wisely laid down that one should first select a king in whose dominion to live; then should he select a wife; and then earn wealth.

These seven parts of sovereignty promote its sustenance. While acknowledging the importance of these parts and their role, the emphasis, however, is laid on the capacity of the king. An able king, it is maintained, can perfect other imperfect constituents of the sovereign. An imperfect king, on the other hand, can destroy even efficient and well-developed constituents. Further, an able king with a small dominion but endowed with perfect constituents of sovereignty can, using his statecraft, conquer the entire world (Dandekar 1963: 251).

The moral and transcendental aspects associated with purusharthas, varnasharamad-harmas and the Saptanga theory of the state, get a distinctively political color when we consider the caste classification accepted in these theories.

5. Caste and Varna Theory

A systematic division of society along caste lines dates backs to Purusha Sukta in the Rig Veda, and continues to the present. It is also found in Chandogya Upanishad. It is said:

> Accordingly, those who are of pleasant conduct here—the prospect is, indeed, that they will enter a pleasant womb, either the womb of a Brahma, or the womb of a Kshatriya, or the womb of a Vaishya. But those who are of stinking conduct here—the prospect is, indeed, that they will enter a stinking womb, either the womb of a dog, or the womb of a swine, or the womb of an outcaste (*candala*).
>
> (Hume 2003: 233)

Manu, too, endorses this division when he says that God created from his mouth the priest, from his arms the ruler, from his thighs the commoner, and from his feet the servant, a division which was effected to ensure the proliferation and prosperity of the people (Manu 2000: 6–7). Assigning the duties of each of these he goes on to add that

the Brahmin has to study, teach, perform the prescribed rituals including officiating at other people's rituals, giving and receiving gifts. Kshatriya has to study, perform the rituals prescribed for him, and protect all life. Vaishya has to study, perform prescribed rituals, and undertake agriculture, cattle-rearing and trade. The Sudras have to attend to the service of the twice-born, i.e. the three higher varnas. While a good deal of discussion of this hierarchical classification emphasizes the difference between caste and varna, the fact at the social level remains that caste and untouchability are features of Indian society. This needs to be addressed independently of whether this was grounded legitimately in the texts or resulted, instead, from a misreading of them.

While these ideas, as understood in the orthodox schools, have appealed to the higher castes, particularly by associating with those castes a good mixture of material and transcendental aspects (either in the form of purusharthas or ashrama dharma), it increasingly became evident that these ideas have been instruments of grave injustice to the lower castes. This blot remained an embarrassment and led to constant attack upon these ideas. However grand their political or individual virtues and ideals might be, the treatment of the lower strata of society by members of the higher castes came to be understood as sharing of injustice and violence. Caste discrimination subsequently became the rallying point or most important topic for political discussion. Let us now look at how this happened. The first ever such attacks came from two heterodox schools of Indian philosophy, Jainism and Buddhism, both of which rejected the authority of the Vedas. The clear rejection of caste came from Buddhism.

6. Buddhism

Caste discrimination and the consequent discrepancy between empirical facts and universal values came under severe attack by Buddhism. Buddha clearly and explicitly attacked the caste division of society, and discrimination on the basis of birth. He instead advocated equality. He found the social life of the Vedas to be governed by rituals, dogmas and hierarchy. Emphasizing the importance of rationality, experience, and self-development, he rejected the Vedic dogmas. In *Samyukta Nikhaya* it is said that the Buddhist doctrine is not based on heresy, but means "come and see." Deviating from the Vedas and Upanishads, Buddha emphasized the importance of human over god and insisted that equality is more significant than caste inequality. Turning attention from the stratification of society on the basis of caste he laid emphasis on the universality of individual suffering. Elaborating the nature and source of suffering he proposed four noble truths: that there is suffering, that it has a cause, that it can be removed, and that there is a way to accomplish this. The first noble truth is the tyranny of pain and suffering. This is the noble truth concerning suffering. Birth is painful, decay is painful, disease is painful, death is painful, union with the unpleasant is painful; painful is the separation from the pleasant, and any craving that is unsatisfied, that too is painful. In brief, the five aggregates which spring from attachment are painful. Craving is the cause of suffering.

In contrast to both the extreme asceticism of the Vedas and Upanishads and to an indulgent life-style, Buddha proposed the middle path to reach Nirvana. This is an eightfold path: right beliefs, right aspirations, right speech, right conduct, right mode of livelihood, right effort, right-mindedness and right rapture.

Buddha gave importance to *sangha* which is based on brotherhood. Disciples join this order to realize the perfect life. The "brotherhood" is open to all including women

and *sudras*. It was, in fact, the widowed queen of Suddhodana who was responsible for making Buddha include women in the order. When she decided to renounce worldly things and become a hermit she asked three times for permission to join the order and was denied as Buddha believed that the presence of women disturbed the mind of the monks. Finally, when she approached again, questioning why only men were included in the order which is devoted to liberation or Nirvana, she was admitted. Retrospectively speaking, this intervention saved Buddhism from authorizing another type of exclusion, namely women. It is this that made many, like Ambedkar, look towards Buddhism as an alternative, though Islam like Buddhism emphasized brotherhood and denounced caste. What is important to observe here is that, notwithstanding the exclusiveness, there is a gradual tendency towards inclusion, which falls between extreme radicalism and mere stagnation.

Now let us look at what Buddha says about how to maintain inclusion along the lines of equality and about some democratic features where there is more participation in social activities.

The ideal society for Buddha is the republic, the *Vaggians* and not monarchy. The former is based on equality, the latter on inequality. Towards the end of his life he gave a detailed account of the reason for the decline and prosperity of republic. In *Maha-Parinibbana-Sutta* he emphasizes the need to "hold full and frequent public assemblies" and to carry out together in concord our various "undertakings." Individuals should strictly adhere to the decisions arrived at and not deviate from the ancient institutions of the Vaggians as established in former days. They should honor, esteem, revere and support the Vaggian elders and shrines. In these republics no women or girls belonging to their clans are detained by force or abduction. He also says that they should not fall into craving, but, instead, should take delight in a life of solitude. They should not engage in, or be fond of, or be connected with business, not partake in idle talk, nor indulge in slothfulness, or fall under the influence of sinful desires, or become the friends, companions, or intimates of sinners, nor come to a stop on their way (to Nirvana).

According to Buddha other important conditions for the strength of the republic are that the brethren shall be full of faith, modern at heart, afraid of sin, full of learning, strong in energy, active in mind, and full of wisdom. Yet other conditions include the brethren exercising themselves in the sevenfold higher wisdom, that is to say, in mental activity, search after truth, energy, joy, peace, earnest contemplation, and equanimity of mind. In addition, the brethren should exercise themselves in the sevenfold perception due to earnest thought, that is to say, the perception of impermanency, of non-individuality, of corruption, of the danger of sin, of sanctification, of purity of heart, of Nirvana. Six other conditions of welfare include perseverance in kindness of action, speech, and thought amongst the saints, both in public and in private. The brethren should divide without partiality, and share in common, with the upright and the holy, all such things as they receive in accordance with the just provisions of the order, down even to the mere contents of a begging bowl. This will ensure their prosperity and avoid their decline.

The influence of Buddha continued, and one of the major impacts is found in Emperor Ashoka, who renounced violence after the dreadful Kalinga war. In his Rock Edicts he makes a plea for universal love, following the path of Righteousness, Dharma (13th Rock Edict). He reports how his own effort to follow this path culminated in making the godly mix with the human thus erasing the difference between great and small people. He says that father, mother and teachers should be obeyed, and that pity should be felt

for all creatures, including animals (1st Rock Edict). According to him, the king should not only love his own people but also those from beyond his frontiers; thus he championed the welfare of the whole world. While recognizing cultural variation between the sects he, however, pointed out one common base. One should control one's tongue so as not to disparage another sect while praising or trying to praise one's own sect, for on certain occasions one should show regard for other sects. By doing this one strengthens one's own sect and helps the other, while by doing otherwise one harms one's own sect and does a disservice to the other. Whoever honors his own sect and disparages another man's, whether from blind loyalty or with the intention of showing his own sect in a favorable light, does his own sect the greatest possible harm (12th Rock Edict). Moreover, along the roads he had banyan and mango trees planted to give shade to man and beast, and had ponds dug and shelters erected (7th Pillar Edict).

Rejection of caste, accepting the importance of equality and democratic participation show the internal resources within Indian political thought to correct evil practices. Notwithstanding this, what haunted Indian political theory is the fact that while the classical Hindu theories advocated universalism and gave sermons on the eternal truth, they nevertheless tolerated caste division. This embarrassed subsequent periods. Consequently, the later theories have had to reckon with this. Buddha's clear rejection of caste became a rallying point in the subsequent discussions. There are those like S. Radhakrishnan who, while accepting Buddha's rejection of Vedic rituals and the caste system, seek to incorporate him within Hinduism rather than recognizing his radical posture. He says:

> Buddha did not break away completely from the spiritual ideas of his age and country. … Buddha himself admits that the dharma which he discovered by an effort of self-cultivation is the ancient way, the Aryan path, the eternal dharma. Buddha is not so much creating a new dharma as discovering an old norm.
>
> (Radhakrishnan 2008: 303)

In sharp contrast, B. R. Ambedkar claimed that Buddha rejected the entire Vedic tradition. He says: "The theory of chaturvarna was as repugnant to the Buddha as the theory of sacrifices was repulsive to him" (Ambedkar 2010: 213). He concludes that Buddha "rejected the philosophy of Vedic Rishis as useless" (Ambedkar 2010: 207). These two opposing strategies reveal the contemporary tension on these issues.

Subsequent to the Buddhist revolt against Brahmanism and Vedic authorities, another important intervention into Indian political theory, and this time from the outside, is in the form of Islam. Islam further widened the scope of inclusion by laying emphasis on brotherhood, equality amongst human beings, and rejection of the caste hierarchy though their views on women remained exclusivist. Indian society and politics tried, in a selective manner, to respond, incorporate, or even maintain a distance from Islam. This influence also has permeated deeply some important respects in the *bhakti* tradition that opposed the high tradition of Hindu orthodoxy.

7. Colonialism

This brings us to the colonial period. In recent times the colonial intervention remained influential, yet became important because of the way India reacted to it. Modernity in

the West that rejected its own traditions, subsequently, and more or less for the same reasons, sought to do to other societies what it had done to its own. Westerners found India to be deeply traditional, backward, bereft of social change, and lacking in progress. Deeply rooted in customs, stagnation, and spiritualism, India did not make progress in science and technology that in turn could have contributed to its material progress. At the political level Westerners found an India that was ruled by aristocratic kings. Indian culture did not have ideas of individualism, citizenship, freedom, liberty or nation. So the colonizers ventured to make India a modern state and embarked on changing its society through development, progress and modernization. To acknowledge an existing autocracy in traditional Indian social relations and to provide the conditions for progress were two ideas that were used by colonial rulers to legitimize their presence. They argued that if Indian people are ruled by autocratic kings, and had no idea of modern ideals of freedom, liberty, nationhood and self-rule, they could be ruled by the British in the same manner. The other argument to justify their rule is that only colonialism using modernity could bring change and progress to Indian society. A refined combination of this argument was considered by Edward Said in his book *Orientalism*, where he showed how the identification of the East with spiritualism and of the West with materialism was used to legitimize colonial power. The virtues of the West, embodied in its modernity, were and continue to be accepted by a section of the colonial subjects, a dominant one at that, who also admitted that, looking through the idea and ideals of modernity, the vices and evils pervading Indian society were clear.

However, there have been serious and systematic attempts arising from the Indian national movement and freedom struggle to counter colonial accounts of the Indian past. They are of two kinds: a scholarly one articulated by R. G. Bhandarkar, R. L. Mitra, B. G. Tilak, and K. P. Jayaswal, who rejected the colonial account of the Indian polity as stagnant, and by their research work have shown instead a course of political and social progress in India from ancient times. In this context they have unearthed and discovered many ancient manuscripts, the most significant one being the *Arthasastra* of Kautilya in 1905, published by Shama Sastry in 1909. This discovery had a remarkable and lasting impact on the history of the study of the Indian polity, providing clear evidence that Indian thinkers in the past had deliberated in a systematic manner on various aspects of the organization of society. While this served the purpose of making a case for anti-colonialism, it failed, however, to appeal to the masses, particularly the lower castes. It is against this background that we have to recognize the importance of nationalist thinkers such as Swami Vivekananda, Sri Aurobindo, Mahatma Gandhi, and B. R. Ambedkar. They had an impact at a more popular level. Their intervention provoked an interesting intersection between India and modernity. These thinkers did not completely reject modernity, nor did they merely accept it. They attempted to understand the resources coming from the outside, incorporating some or a lot of them yet not completely surrendering their own heritage. In this process they created some important variations on both colonial and classical Indian approaches, blending three elements: retaining the classical teachings, being sensitive to the resources for criticism that were already available within that tradition, and, finally, being sensitive to Western ideas.

An important feature in post-classical philosophizing in the colonial or post-colonial period is the change in language, namely the shift from Sanskrit to English on the one hand and the vernacular on the other. This provided Indian political thinkers three routes that are at work in today's Indian political thought. They are: accessing past

resources from Sanskrit and Pali, the possibility of reaching local populations through the vernacular, and access to the outside through English. So, while Bankim wrote in Bengali, others such as Vivekananda and Aurobindo broke away from region-specific vernacular media and changed to English. These reveal the sensitivity to the contemporary context in Indian political philosophy and the complexity of reactions to that context.

Partha Chatterjee credits Gandhi as a person who for the first time in Indian politics provided an "ideological basis for including the *whole* people within the political nation" (1986: 110). Gandhi may have provided the ideological basis as well as attempting to bring the masses within the mainstream of the political process, but his predecessors made a claim to bring the masses into the mainstream of political discourse, thereby setting the theoretical background to Gandhi's activities. Thus, the other important element in modern Indian political thought was the attempt, beginning well before Gandhi, to bring the "masses" into the mainstream of philosophical discussion, as for instance in the writings of Swami Vivekananda and Sri Aurobindo. In the process of bringing the masses into the fold, these thinkers simultaneously embarked on an internal criticism which holds the proponents of Hinduism responsible for the poor conditions of the masses. In this context they recognized some positive aspects of Western modernity and at the same time sought to maintain a balance between materialism and spiritualism. This call to bring the masses into the mainstream demanded major structural changes in Hindu society. Vivekananda admits that

> the great national sin is the neglect of the masses, and that is one of the causes of our downfall. No amount of politics would be of any avail until the masses in India are once more well educated, well fed, and well cared for. … If we want to regenerate India, we must work for them.
>
> (1994: Vol. V, 222–3)

He adds, "The one thing that is at the root of all evils in India is the condition of the poor" (Vivekananda 1994: Vol. IV, 362). Sri Aurobindo also admits how the so-called spiritual achievements in India are only "confined to a few" and have not spread to the "whole mass of humanity" (Aurobindo 1972: Vol. 22, 1). All these show a movement towards inclusion.

The rich resources in Indian political thought are evident, then, in this trajectory from an emphasis on the unchangeable divine order to the importance of a humanly accessible dharma, and the relevance of artha and kama, and thence to the inclusion of both (sudras) dalits and women and the incorporation into Indian political thought of modern ideals of individualism, equality, freedom and democracy. The way in which this variety is processed by Indian political theory is to be watched with interest.

Related Topics

Plato's Political Philosophy, Hobbes, Liberalism

References

Ambedkar, B.R. (2010) *The Essential Writings of B.R. Ambedkar*, ed. Valerian Rodrigues, New Delhi: Oxford University Press.

Aurobindo, Sri (1972) *Letters on Yoga: Parts One and Two, Vol. 22*, Pondicherry: Sri Aurobindo Birth Centenary Library.
Badrinath, Chaturvedi (2007) *The Mahabharata: An Inquiry in the Human Condition*, Hyderabad: Orient Longman.
Chatterjee, Partha (1986) *Nationalist Thought and the Colonial World: A Derivative Discourse?*, Delhi: Oxford University Press.
Dandekar, R.N. (1963) "Dharma, the first End of Man," in *Sources of Indian Tradition*, compiled by Theodore de Bary, Stephen Hay, Royal Weiler and Andre Yarrow, Delhi: Motilal Banarsidass.
Ghoshal, U.N. (1959) *A History of Indian Political Ideas: The Ancient Period and Period of Transition to the Middle Ages*, London: Oxford University Press.
Hume, R.E. (2003) *The Thirteen Principal Upanishads*, New Delhi: Oxford University Press.
Kautilya (1992) *The Arthashastra*, ed. L.N. Rangarajan, New Delhi: Penguin.
Manu (2000) *The Laws of Manu*, trans. Wendy Doniger with Brian K. Smith, New Delhi: Penguin.
Radhakrishnan, S. (2008) *Indian Philosophy Vol. I*, New Delhi: Oxford University Press.
Said, Edward (1979) *Orientalism*, New York: Vintage Books.
Valmiki (2000) *The Ramayana*, trans. Arshia Sattar, New Delhi: Penguin.
Vivekananda, Swami (1994) *The Complete Works of Swami Vivekananda, Vols. 1–9*, Calcutta: Advaita Ashrama.

Further Reading

Guha, Ramachandra (2010) *Makers of Modern India*, New Delhi: Penguin.
Jayal, Niraja Gopal and Mehta, Pratap Bhanu (eds.) (2010) *The Oxford Companion to Politics in India*, New Delhi: Oxford University Press.
Mehta, V.R. (1999) *Foundation of Indian Political Thought: An Interpretation*, New Delhi: Manohar, 1999.
Raghuramaraju, A. (2011) *Modernity in Indian Social Theory*, New Delhi: Oxford University Press.
Sharma, Ram Sharan (1968) *Aspects of Political Ideas and Institutions in Ancient India*, Delhi: Motilal Banarsidass.

19
ISLAMIC POLITICAL THOUGHT

Andrew F. March

1. The Political Ethics of the Qur'an

The Qur'an presents itself as a universal ethical code for mankind, in sharp contrast to the tribal particularism of pre-Islamic Arab codes. It presents a conceptual and normative scheme for both socio-political ethics and the duties and virtues of individual believers. The values, norms and commands revealed in the Qur'an transform select pre-Islamic ones while also introducing a normative revolution. Such pre-Islamic values as generosity, courage, loyalty, veracity and forbearance are given Islamic validation by transforming them into virtues which believers are commanded to cultivate in the service of Islam within the limits set down by God. For example, the Qur'an praises generosity in the giving of charity, while condemning both profligacy and spending of one's wealth out of vanity and the desire for praise. Similarly, the pre-Islamic value of absolute in-group loyalty (*walā'*) was transformed by the more complex ethical terrain of the new religion. Loyalty to family and tribal kin was not only expanded to the community of believers (which trumped more particular loyalties) but was also constrained by other Islamic commitments. Believers were commanded to not violate oaths or transgress against Divine commands, even in the service of communal interests. Loyalty within socio-political contexts thus becomes subsumed within the general obligation of loyalty to the covenant with one's Creator, which the Qur'an indicates is required by man's recognition that he is created by a sovereign, autonomous God.

However, the Qur'an views pre-Islamic Arabia predominantly in negative terms. It is characterized as *jahiliyya*, a concept that refers to a state of moral recklessness arising from submission to human-social passions and whims. The normative revolution in Islam consists in its theocentricity. God Himself is referred to in ethical terms (many of the Names of God refer to ethically salient features of God's essential nature) and man's attitude towards God is the primary criterion of moral evaluation. The *summum malum* for a created being is *kufr*, a state of being that alludes to unbelief, ingratitude and the (public) denial of God's existence (*takdhib*), and which results in acts of insolence, arrogance, presumptuousness, mockery of revelation and the violation of Divine limits. That the *kafir* goes astray by committing acts that are substantively unjust, criminal, sinful, forbidden or tyrannical (referred to in the Qur'an variably as *zulm*, *jurm*, *fasad*, *munkar*, *sharr*, *su'*, *fahsha'*, *khabith*, *haram*, etc.) is not itself fully constitutive of the state of *kufr*; rather, it is the consciously held and affirmed beliefs in one's own independence and self-sufficiency in the formulation of judgments, moral and otherwise. The reli-

ance on one's own judgment, which lies at the root of error and *kufr*, is often referred to as *hawa*, or (roughly) the lustful, whimsical and passionate inclination of the human animal. Importantly, the reliance on reason to arrive at sound moral and practical judgments can fall under the scope of *hawa*.

The *summum bonum* for a created being is thus *iman*, a state of being that includes belief in God and His revelation, performance of all mandated rituals, good works and observances, obedience to all commands and prohibitions and, perhaps most centrally, the willingness to put all trust in God and to subordinate one's individual judgment to Him. The *mu'min* (believer) is the one who accepts God's guidance (*huda*), fears God (*muttaqi*), is grateful to God (*shakir*) and is upright (*salih*) according to God's prescriptions. (See Izutsu 2002.)

The Qur'an also has much to say on the ethics of social relations amongst humans. While the specific norms and rules of social relations are elaborated within Islamic positive law, it is possible to generalize about some of the broad principles and major themes of Islamic socio-political ethics. The Qur'an exhorts strong bonds of communal loyalty, extensive social solidarity, charity for the poor, and obedience to those in authority. The socio-political vision is a moderately egalitarian one. Just rulership is only that which is exercised in the interests of the ruled according to Divine guidance. Extreme inequalities are condemned, the poor are said to have a claim on the property of the rich, and wealth is strictly detached from evaluations of virtue, desert and piety. At the same time, the Qur'anic vision is less ascetic and unworldly than is, perhaps, the early Christian one. Wealth is not, per se, a sign of impiety, nor are other good things of this world. The Qur'anic emphasis is on remaining within the limits of enjoyment established by God and on purifying individual motivations. As such, it might be said that the Qur'an attends to considerations of moral psychology, or a realistic attempt to take humans as they are according to an understanding of their motivations, capacities and needs.

There are, however, limits to the Qur'an's egalitarian vision. In addition to some of the inegalitarian distributions of roles and rights within the Muslim community along gender lines, Qur'anic ethics distinguishes sharply between Muslims and non-Muslims. The primary solidarity community is the community of believers, at least from the Medinan period on (622 CE). Muslims are enjoined not to privilege relationships of loyalty with non-Muslims at the expense of the Muslim community. Amongst other reasons, non-Muslim communities are all potential political and military rivals to the Islamic one. Similarly, while Muslims acquire rights by virtue of being Muslim (i.e., they do not need an explicit relationship to a state or ruler), non-Muslims may acquire them only contractually. However, the Qur'an enjoins strict adherence to compacts with non-Muslims and this forms a means of constructing relationships of mutual moral obligation.

2. The Caliphate, or Imamate

After the death of the Prophet Muhammad in 632 CE, leadership of the Islamic community passed to a series of successors, or "Caliphs." While Muslims often imagine the early period of Islam under the "Rightly-Guided Caliphs" as a golden age of unity, piety and worldly success, consensus within the community over its leadership was the exception rather than the rule. The first successor to Muhammad, Abu Bakr (r. 632–634), was immediately faced with a revolt from various Arab tribes who had obeyed and,

most importantly, paid taxes to Muhammad. That these revolts against the political authority of Muhammad's successor come down in Muslim tradition as the Wars of *Apostasy* (*ridda*) says much about the fusion of political, military, economic and religious authority in early Islam. To revolt against the ruler was to revolt against the community, which was to leave the religion. Between 656 and 750, the Muslim community was to experience three major civil wars, or *fitna*s, until the descendants of Muhammad's uncle 'Abbas (thus, the "'Abbasids") were able to assert unitary control and establish an imperial capital in Baghdad.

What is important to note about the early strife of the Muslim community is how conflict over the Caliphate, particularly the First Civil War (656–661) following the assassination of the third Caliph 'Uthman and the assumption of power by 'Ali, was able to generate not only political factions or rival dynasties but religious *sectarianism*. That religious sects were to emerge around competing views of the legitimacy of various claims to the Caliphate and the proper role and status of the office suggests that for early Muslims the unitary leadership of the community was a matter of immense religious, even soteriological, significance. Patricia Crone (2004) has recently offered the metaphor of a caravan to express the earliest Muslim views of politics and religion. If the community was a caravan seeking to make its way through this life, its guarantee of success in reaching salvation in the next world was in following the right leader (or "Imam"). Getting it wrong could mean the difference between salvation and damnation; thus, one's position on whether 'Uthman had been justly assassinated or whether 'Ali was justified in consenting to arbitration between himself and 'Uthman's kin, was able to result in sectarian as well as political party differences.

Doctrinally, the primary disagreement between various camps was whether the Caliphate, or Imamate, was primarily an elective office or a hereditary one. Those who came from the "Sunni" majority held the view that any man of the Prophet's wider tribe (the Quraysh) who fulfilled certain conditions of competence and piety could hold the office. In principle, the office passed from one Imam to the next by "election," meaning that the elite members of the community ("those who loose and bind") all swore allegiance to a new Caliph. In practice, the Caliphate was hereditary along a kingship-model both under the Umayyad dynasty (the kin of 'Uthman who were able to wrest power from 'Ali by 661) and the 'Abbasids (who overthrew the Umayyads in 750). On this view, what was important was how the Caliph acceded to power and what he did there. Nonetheless, the early theory and practice of the Caliphate was a thoroughly charismatic one. They called themselves "God's [not the Prophet's] Caliph," at times flirted with messianic themes, and wholeheartedly adopted pre-Islamic Byzantine and Sassanian Persian enunciations of power and kingship. (See Azmeh 1997.)

The other view was that the office belonged solely to the Prophet's narrow family because of certain special qualities they possessed. What we *today* know simply as "Shi'ism" is the view that the Imamate passed directly from the Prophet to his son-in-law 'Ali and then down through his descendants, until the twelfth such Imam "vanished" into a cave as a child in 874, to return one day as the Messiah. We thus refer to the vast majority of Shi'ite Muslims today (particularly those communities in Iran, Iraq and Lebanon) as "Twelver" Shi'ites. On this view, the Imamate is an inherited *status*, not the successful assumption of a political office. The descendants of 'Ali simply *were* the Imams, even though not a single one ever held sovereign power.

However, during the first two centuries of Islam, the phenomenon of "Shi'ism" was much wider than this single sectarian view. All opposition to the Umayyads which held

out hope for a restoration of someone from the Prophet's wider family (the Hashimite clan) to restore justice to the world and make the office of the Imamate meaningful can be referred to as a form of "Hashimite Shi'ism." Thus, when the descendants of the Prophet's uncle 'Abbas seized power in 750, this revolution was full of charismatic and even messianic expectations on the part of many Muslims. It was the refusal, or at times failure, of these Caliphs to assert the full scope of charismatic authority in the realm of theological doctrine and legislation, as well as their success in politically marginalizing the 'Alids, which eventually guaranteed the emergence of the "Sunni" consensus.

Other early doctrines of the Caliphate with origins in 'Ali's claim to leadership included the Kharijite sect. The Kharijites ("those who go out") were supporters of 'Ali who rebelled against him and eventually assassinated him over his willingness to submit to arbitration with his enemies (the proto-Umayyads) during the Battle of Siffin in 657. For these supporters, 'Ali's claim to leadership was above human interference and they vigorously protested against the setting up of a human tribunal above the Divine word. Loudly protesting that "judgment belongs to God alone" (*la hukma illa li-llahi*) they left 'Ali's army and gradually recruited further defectors from 'Ali's army, assassinating their former master in 661. While the Kharijites began as supporters of 'Ali, their politico-theological doctrines were diametrically antithetical to later Shi'ite ones. They advanced a conception of political authority that was both rigorously authoritarian and fiercely egalitarian and meritocratic, at times almost anarchical. When a legitimate Imam is in power, all in the community must obey him even to the death. However, while any pious Muslim may qualify for the Imamate (in their phrasing: "even an Ethiopian freed slave with a head like a raisin"), his power and authority is strictly contingent on ruling with perfect justice and piety. As soon as the community judges that he has violated religion in some significant way, they are entitled to depose him and even kill him. The extreme zealotry of the Kharijites, particularly their willingness to anathematize Muslims for sinful behavior, has led many contemporary commentators to refer to certain radical Muslim groups today with a similar propensity to anathematize and wage war against fellow Muslims on grounds of sin or any doctrinal difference as "neo-Kharijites."

3. Developed Sunni "Constitutionalism"

As noted above, the 'Abbasid dynasty occasioned the triumph of "Sunni orthodoxy." Sunnism—from the phrase "*Ahl al-Sunna wa'l-jama'a*" ("The People of the Prophetic Example and Communal Unity")—has its origins in the politically quietist approach to the civil wars and the rise of formal religious scholarship. While the Imamate was initially important to all, and gradually only to those various "Shi'ites," as the supreme locus of religious knowledge, authority and meaning, Muslims from an early period also began to identify the source of religious knowledge not only in the Holy Book, but also in the *community's* memory of the Prophet's practice and example. Lay Muslims, and later professional scholars, began to collect reports of the Prophet's behavior and utterances, known individually as *hadith*-reports and collectively as the Prophet's *sunna*. By the eighth, and certainly by the ninth, century a class of pious Muslims devoted to discovering and preserving reports of the Prophet's *sunna* as an expression of revelation alongside the Qur'an had developed.

This group was (largely) politically quietist, but also began to put the office of the Imamate in perspective. It was supremely important to have an Imam, and he must

fulfill a set of criteria to be appointed and remain legitimate, but he was no longer an infallible source of doctrinal and moral knowledge. He now had to share his authority and charisma with the wider community (which held collective knowledge through transmitted memory of the Prophet's practice) and eventually with an elite class of scholars—the *'ulama*.

Developed Sunni political thought, emerging from the ninth to the eleventh centuries, came to have the following features. First, contrary to various Shi'ite views which insisted on the right man for the Imamate, Sunnism was determined to keep the present legitimate, to avoid dissonance between moral ideals and political realities as much as possible. Here "legitimate" means something more than merely acceptable or tolerable but less than ideal or utopian. At times, this involved the creative use of legal fictions or redescriptions of historical reality, especially once the 'Abbasid Caliphs lost actual power on the ground to various regional and tribal "warlords" (from the latter half of the tenth century). Thus, rulers in the provinces who usurped Caliphal power became "appointed governors"; paltry figures at the top became "quasi-caliphs" or legitimate by necessity; kings became "servants of the Caliph." Second, Sunnis were committed to keeping as many Muslims as possible within the communal fold. They avoided anathematization (*takfir*) as much as possible, recognized reasonable disagreement in areas of law and theology, and reconstructed the memory of early Islamic history to preserve the reputations of as many of the participants of the civil wars as possible. Third, Sunnism came to emphasize above all the commitment to Islamic legal order on the ground regardless of who held military and executive power.

We thus might speak about the developed Sunni political vision as "constitutional." The tenth and eleventh centuries produced a number of canonical works of "public law," often under the title *Al-Ahkam al-sultaniyya* ("The (Religious) Rulings on Governance"), the most famous of which was written by the legal scholar Abu'l-Hasan al-Mawardi (d. 1058). Mawardi gives a series of rules related to the status, legitimacy and obligations of the Imam, as well as treatments of the delegation of power, the organization of the military, financial institutions and criminal law. In this work, Mawardi shows by both revelation and reason why Muslims must have an Imam and outlines his "*shar'i*" (those prescribed by the religious law) and "non-*shar'i*" duties. "*Shar'i*" duties include validation of the community as the legal Muslim community; validation and leadership of public worship; execution of the Law and in particular the "*hudud*" punishments mentioned in the Qur'an for certain crimes; waging both defensive and expansionary jihad; preserving right religion; and gathering certain required taxes and alms. "Non-*shar'i*" duties related to internal security; building roads, bridges, inns, walls, mosques and other infrastructure; charity and social welfare; medical services; and education and culture.

In keeping with the tradition of theory emerging in response to crisis, we might note that these intellectual efforts did not reflect a codification of dull custom so much as a panicked response to an acute crisis, namely the loss of Caliphal power at the hands of various tribal dynasties. If the early "Sunni" response to the Imamate was about cutting the charismatic Imam/Caliph down to size, the eleventh- and twelfth-century Sunni preoccupation was about nervously building the status of the Caliph back up. Thus, the great orthodox theologians Abu'l-Ma'ali al-Juwayni (d. 1085) and his student, the towering Abu Hamid al-Ghazali (d. 1111), both wrote treatises outlining ways in which the legality of the religious community could be preserved when the secular sovereign was threatened with the loss of power. Whereas Juwayni stressed the need for an effective ruler, Ghazali wanted a legitimate Imam. For Ghazali, the legitimacy of the legal order

and public sphere (judges, marriages, deputies, etc.) depended on being validated by a legitimate Imam. Without an Imam to legitimate the community, "the carpet of the law would be rolled up in its entirety." In typical Sunni fashion, if preserving the legality of the public order meant that the religious scholars had to keep redefining the role and qualifications of the Imam so that the existing one would be legitimate, then this was vastly preferable to either constant civil war or moral confusion. Ghazali's legitimate Imam seems to bear some comparison to the British monarch: it is a legal fiction that she is actually sovereign; rather, she represents the symbolic unity and legality of the institutions of government and religion.

After the Caliphate was destroyed by the Mongols in 1258, later religious scholars (notably, al-Qarafi, Ibn Taymiyya and Ibn Qayyim al-Jawziyya) looked to secular rulers ("Sultans") to fulfill all of the guardianship duties of political life, especially preserving order and unity, in the hope that the scholars would have the space to preserve a *shari'a*-based social life. The doctrine known as "*al-Siyasa al-shar'iyya*" ("Religiously-Legitimate Political Rulership") emerges from this time. It is a familiar shepherd-and-flock vision of politics, but one set against the backdrop of religious moralism. While it represents an extension of the *al-Ahkam al-sultaniyya* public law tradition, it also displays hints of Machiavellianism in its comfort with dirty hands in the ultimate service of religion and its appreciation of religious morality's need for political order. The ultimate message seems to be that Muslims can forgive a sinful ruler who advances their interests but not a weak one who leaves them vulnerable: "God will strengthen His religion even through the dissolute." This realist vision of politics allows for a significant degree of discretionary authority on the part of secular actors, as long as their discretionary policies are not taken for timeless religious morality.

A few important points stand out here for modern readers. First, religion and politics are deeply intertwined in the Sunni vision, but are not quite the same thing. For the medieval Sunni Muslim, the good life and religious fulfillment were not to be found in politics. Rather, politics was there to secure the legal and material conditions for believers to pursue salvation, above all order and stability. In theological-anthropological terms, orthodox Sunni religious thought leaned mostly towards a Hobbesian pessimism about natural human motives, although it also had heavy doses of Aristotelian conceptions of humans as social (if not political) animals. Humans need one another, but their socialization is also a tenuous achievement. Because most humans are selfish and immoral, society needs a strong absolute ruler to hold things together.

Second, while rulership (state authority) is regarded as a religious matter (it is a collective religious obligation on the community to have a ruler and the ruler has both religiously set obligations and sacral functions), the state per se was not the highest object of loyalty, and the state had no monopoly on religious or moral interpretation. All subjects were commanded to obey the ruler, often dubbed the "Shadow of God on Earth," but in principle "there was no obedience to a creature in defiance of the Creator." However, this tension between the near-absolute obligation to obey and the maxim "no obedience to a creature in defiance of the Creator" is a quintessential Sunni dilemma. It was never resolved at the level of theory or doctrine. The closest one finds to a resolution of this tension is something like a "moral luck" doctrine: rebels who could be killed as traitors and sinners also could be honored and legalized ex post facto if successful.

Third, the Sunni conception of politics and public morality is profoundly legalistic. The just ruler is one who wages war in the cause of God, preserves public safety and

advances common welfare, but otherwise leaves the scholars to the business of regulating social life. The scholars aspired to judicial and legislative authority, expecting the rulers to *execute* what they decided the Law was and to effectively safeguard the community from political chaos, heresy and sedition from within, and infidels from without. If the ruler governs according to the Law and in the interests of the ruled, then his rule is honored with the title of "*imama*"—leadership in accordance with God's instructions for mankind. If he rules for his own private benefit and in defiance of both the Law and the interests of his flock, then his rule is denigrated as mere "*mulk*"—secular kingship.

4. Islamic Political Philosophy

Greek philosophical and other texts were translated into Arabic from the early 'Abbasid period. Great philosophers in the Sunni Muslim sphere included al-Kindi (d. 873), al-Farabi (d. 950), Ibn Sina (Avicenna, d. 1037), Ibn Bajja (Avempace, d. 1139), Ibn Tufayl (d. 1186) and Ibn Rushd (Averroes, d. 1198). Philosophy (*falsafa*) sat in a very uneasy relationship with orthodox theology and law for obvious reasons. However, this mostly pertains to outright claims to philosophy's superiority over religion as a path to knowledge and "happiness" and provocative claims about the status and source of prophecy. In reality, philosophy was not so much suppressed and excluded from medieval Islamic learning as gradually integrated into orthodox theology.

The themes discussed by the philosophers (see Nasr and Leaman 1996) which are of broad political relevance include: the philosophical status of the Prophet and the Imam as knower and legislator; the relationship of the claims expressed in revelation to knowledge derived from reason; the classification of people based on their innate capacities and what they can expect to achieve in this world; the nature, status and acquisition of the virtues; and the meaning of "happiness," its relationship to the concept of salvation and the role of social and political life in achieving it. Much Arabic philosophical writing on these questions was directly borrowed from Plato and Aristotle, including reflections on the virtues, the division of people into classes, the trope of the golden mean, and the themes of the necessity of social life and mutual assistance. The Islamic context required additional speculation on the status of the Qur'anic revelation, the figure of Muhammad and subsequent Imams, and the relationship of the revealed Law to ethical perfection.

The greatest political philosopher was certainly Abu Nasr al-Farabi (873–950). Farabi's political utopia (the Virtuous City) consists in human cooperation, while obeying the divinely inspired Philosopher-King, perfecting the intellectual capacities, and developing the practical virtues by obeying the *shari'a*. The study of politics becomes a guide to man's good actions and behavior and enables man to distinguish the good from the bad, which is Farabi's way of arguing that Platonic and Aristotelian philosophy prove the justness of the *shari'a*. (Following Plato, Farabi argues that those who do not try to apply their theoretical perfection to practical and political pursuits cannot claim to be true philosophers.) Nonetheless, while religion is the cornerstone of politics (because of its necessity for the masses as law and social custom), religion and religious law is just a simulacrum of the real opinions of the Prophet, or Philosopher-King.

Farabi, following Plato's *Republic*, thus discusses various non-ideal "cities." Imperfect cities are *ignorant*, *wicked* or *errant*, each of which has sub-types. Ignorant cities all fail to comprehend the true nature of humanity, its place in the cosmos and its natural end, substituting such false goals as mere subsistence, wealth, sensual gratification, honor,

power and domination, or the whim of individuals ("democracy"). Wicked and errant cities are those that used to pursue the true end but now do not; wicked cities abandoned it on purpose, whereas in errant cities the leader knows the true end but deceives the people about it. Crucially, the constitution of the city has eschatological significance: human happiness ultimately consists in assimilation with the Active Intellect that is achieved when one reaches the stage of acquired intellect. Only the citizens of the Virtuous City will be able to achieve this goal and thereby survive after death when their actualized intellectual souls separate from their bodies. Those who lived in *ignorant* cities will suffer no punishment since their ignorance was not culpable; they will simply be annihilated as a natural consequence of their failure to actualize their intellectual powers, which is the condition for the soul's survival after death. Punishment is reserved for the citizens and leaders of the wicked cities and the rulers of the errant cities, and it consists merely in the continuance of their corrupt desires after death but which can longer be fulfilled in the absence of bodily forms, thus resulting in torment.

After Farabi, Islamic philosophers who discussed themes of the soul's purification through intellectual perfection, the cultivation of the practical virtues through habits of self-discipline and the inequality of various types of humans included the "Brethren of Purity" (a mid-tenth-century pro-Isma'ili group), Miskawayh (932–1030), Ibn Sina (980–1037) and al-Raghib al-Isfahani (?–1108). Mention must also be made of the tendency of some Islamic philosophers in the Western lands of Andalus (Southern Spain) to stress the incompatibility of social and political life with individual ethical and intellectual perfection. Ibn Bajja (d. 1139) and Ibn Tufayl (d. 1186) both stressed how often the virtuous man and philosopher lives in an ignorant or tyrannical city and is "like a stranger in the world." While the city itself benefits from the learned and can be improved by ethical virtues learned from the philosopher, the philosopher himself is often better off in isolation from society, governing himself and contemplating the truth. Ibn Tufayl, in his well-known parable *Hayy, Ibn Yaqzan* ("Alive, Son of Awake"), broke with the dominant motif in Islamic philosophy since Farabi in arguing that philosophy and religion were two entirely separate and independent endeavors which can co-exist but are neither compatible nor mutually supportive. The tremendous philosophical corpus of Ibn Rushd (Averroes) represents a rejection of Ibn Bajja's and Ibn Tufayl's pessimism. According to Ibn Rushd, happiness is immortality of the soul which can be attained in a growing conjunction of man's acquired knowledge with the Active Intellect. Because this acquired knowledge is the collective achievement of mankind, philosophical knowledge and happiness are not the sole aim of any single individual. Nonetheless, Ibn Rushd shared in the judgment of Ibn Bajja and Ibn Tufayl that all actual polities since the early Muslim period have been corrupt and thus the role of the city is reduced to providing the necessities for man's existence whereas the philosopher in fact is like a man "among perilous animals."

5. Modernity: The "Reformist-Nationalist" Response of the Early Salafiyya

For Islamic thought, the problem of modernity is inseparable from the problem of the relative power imbalance between the West and the lands of Islam. The variety of intellectual trends within Islamic thought all have as their primary stimulus (in some form or another) the imperative of providing authentic "Islamic answers" to the problems of Western colonialism and imperialism and the corresponding Muslim political and

economic weakness. All of the main debates which form the contours of modern Islamic political thought—the relative status of reason versus revelation, the immutability versus the reformability of Islamic law, the moral status of national or regional versus pan-Islamic political membership, the status of non-Muslim states and relationships with non-Muslims, the legitimacy of democratic forms of rule, the laws of warfare and political violence, the place of technology—have taken place in reaction to Western ascendancy and hegemony. (See Kerr 1966.)

Islamic Modernism emerged as an elite movement in the later part of the nineteenth century as an attempt to bridge Islamic theological and epistemological commitments with Western modernity. It was an attempt both to rehabilitate Islam as a source of knowledge, identity and inspiration for Muslims, as well as to allow Muslims to incorporate those cultural and intellectual aspects of European modernity seen as necessary for competing with Western political and economic power. The core tenet of Islamic Modernism was that Islam itself was not the cause of nineteenth-century Muslim stagnation, but that certain theological and canonical reforms were necessary to awaken Muslims from their submissiveness and quietism.

The main concerns of Islamic Modernism were to demonstrate the compatibility of Islamic revelation with modern science and rationalism, and to define what is authoritative and binding in the Islamic sources and what is contingent historical accretion. (Derivative questions, such as democracy, rights or the status of women, admitted of a range of treatments and views.) Characteristic of Modernism are the positions that the Qur'an either anticipates or accommodates modern science, and that there thus need be no ultimate conflict between revelation or Divine inspiration (*wahy*) and reason (*'aql*); that Muslims ought to orientate themselves directly according to the Qur'an rather than the accrued tradition of legal, theological and political thought; that God has commanded Muslims to thrive in the world and that they thus may borrow material and mundane achievements from the West; and that in order to achieve this consistently with Islam's immutable spiritual and ethical principles the Islamic community is required to reform its legal codes through the practice of "independent reasoning" (*ijtihad*), which, on the one hand, was to rely strictly on actual Divine revelation yet, on the other, liberally take advantage of certain legal devices, namely argumentation from public interest (*istislah*) and (arbitrary) juristic preference (*istihsan*), which had been accorded a much narrower scope in classical jurisprudence (*usul al-fiqh*). The ideas of Islamic Modernism provide the foundation for the moderate Islamist ideology espoused by the leader of the Tunisian al-Nahda party, and a significant Islamic thinker in his own right, Rashid al-Ghannushi, as well as many on the pro-democratic wing of the Muslim Brotherhood in Egypt and elsewhere.

6. Modernity: Revivalism and Sayyid Qutb's "Realistic Utopia"

Islamic Revivalism refers to the broad ideological trend that insists on the centrality of religion in all aspects of Muslim family, social, economic and political life. It emerged as an explicit rejection of both inter-war secularist trends as well as Islamic Modernism. For Revivalists, the latter represents an apologetic, pro-Western betrayal of core Islamic commitments, although Revivalism in some manifestations shares Modernism's rejection of what it perceives as the conservative, quietist, passive nature of traditional orthodox scholarship and the insistence on direct engagement with the Qur'an. While rejecting many of Modernism's reforms and openness to change, and reverting to many

of the doctrinal positions of the medieval legal schools, Revivalism has as its central *raison d'être* the provision of authentic "Islamic solutions" to modern social problems and the weakness of Muslim lands; this aspiration to popular support and tangible results leaves Revivalism at times at odds with the self-restraint, caution and concern with methodology which characterized the medieval religious scholars. (See Abu-Rabi' 1996.)

The Islamist trends of the twentieth and twenty-first centuries have taken many forms, some of them mutually incompatible, from the reformed Muslim Brotherhood in Egypt and Jordan (which today espouses moderate positions on women, democracy, political violence and intra-Muslim religious disagreement) to the puritanical Wahhabi and Salafi movements. These important distinctions should be at the center of any study of modern political Islam; however, here the broad ideological trend that has at its core the commitment to an Islamic political and social order that rejects the need to self-consciously "modernize" Islam can be referred to as "Revivalism."

For more radical Revivalists, everything in life can be characterized as either authentically Islamic or some form of "*kufr*" (unbelief or rejection of God). The greatest threat to Muslims in the modern world, according to Revivalist ideologues such as Egypt's Hassan al-Banna (1906–1949), who founded the Muslim Brotherhood in 1928, and Sayyid Qutb (1906–1966), Pakistan's Sayyid Abu al-A'la Mawdudi (1903–1979, the intellectual inspiration and sometime leader of the Jama'at-i-Islami), and Iran's Ayatollah Ruhollah Khomeini, was not the material weakness of Muslim societies but the incursion of un-Islamic ideas and practices, all of which Mawdudi and Qutb tagged with the epithet "*jahiliyya*" (a reference to the "pagan ignorance" of pre-Islamic Arabia).

Perhaps the most important political theorist of Revivalism is Qutb, who elaborated a conception of Islamic political ethics throughout many treatises and a full exegesis of the Qur'an. Qutb's work is noteworthy for its inspiration of many radical, revolutionary Islamic groups who followed Qutb in regarding most existing Muslim polities as "apostate" regimes (and thus eligible for treatment according to the laws of war), but his work also contains a very subtle reconstruction of orthodox Islamic law and theology as a response to modernity and secularism. Qutb's political theory stressed two important points: (1) that Islamic law and human nature (*fitra*) are in perfect harmony, and (2) that the demands of Islamic law are easy and painless for ordinary human moral capacities. Qutb thus not only defends Islam as the normatively required doctrine for Muslim societies, but also as a coherent "realistic utopia"—a comprehensive theory of political life which not only posits a true doctrine of the good and the right but also contains a psychological account of that theory's own feasibility and perpetual stability. Qutb's well-known fascination with the earliest generation of Muslims (the *salaf*) are not merely an appeal to the normative authority of the founding generation but serve two functions: first as an account of an actual historical model of the feasibility and realism of an ideal Islamic political order, based on a thoroughly political and psychological interpretation of the experience of that early community, and second as a "Paradise lost" story providing a quasi-Rousseauvian genealogy of the political origins of moral vice in society. Qutb's account of Islamic law and morality as realistic and practical for humans given their actually existing psychology thus bears a certain comparison to modern Rousseauvian and Kantian concerns. But instead of asking modern humans to search for the "law they would give themselves," Qutb's argument is that the Islamic *shari'a* is so calibrated to respond to human needs and limitations that it is in fact the law we would give ourselves. Qutb is thus "modern" in his recognition of the needs of post-traditional

humans, in evidence partially in his revision of the classical negative and pessimistic theological anthropology which dominated pre-modern Islamic thought, but is also "reactionary" in his insistence on an unreformed application of a simplified understanding of Islamic law. While post-Qutbian Islamists are often tempered politically, and open to some reform of Islamic law, the claim that Islamic law represents the manifestation of Islam as the "natural religion" (*din al-fitra*) is a ubiquitous point of doctrine and a central aspect of the Islamic intellectual response to secularism.

Related Topics

Medieval Political Thought, Natural Law and Rights Theory, Religion in Public Life

References

Abu-Rabi', Ibrahim M. (1996) *Intellectual Origins of Islamic Resurgence in the Modern Arab World*, Albany: State University of New York Press.

Al-Azmeh, Aziz (1997) *Muslim Kingship: Power and the Sacred in Muslim, Christian and Pagan Polities*, London: I.B. Tauris.

Al-Farabi, Abu Nasr (1985) *The Principles of the Opinions of the Citizens of the Virtuous City* [*Mabadi' ara ahl al-madina al-fadila*], trans. Richard Walzer in *al-Farabi on The Perfect State*, Oxford: Clarendon Press.

Crone, Patricia (2004) *God's Rule: Government and Islam*, New York: Columbia University Press.

Izutsu, Toshihiko (2002) *Ethico-Religious Concepts in the Qur'an*, Montreal: McGill-Queen's University Press.

Kerr, Malcolm H. (1966) *Islamic Reform. The political and legal theories of Muhammad 'Abduh and Rashid Ridā*, Princeton, NJ: Princeton University Press.

Mawardi, Abu al-Hasan (1996) *The Ordinances of Government* [*al-Ahkam al-Sultaniyya*], trans. W.H. Wahba, Reading, UK: Garnet.

Nasr, Seyyed Hossein Nasr and Oliver Leaman, eds. (1996) *History of Islamic Philosophy*, London: Routledge.

Qutb, Sayyid (1996) *Social Justice in Islam*, trans. William E. Shepard in *Sayyid Qutb and Islamic Activism: A Translation and Critical Analysis of* Social Justice in Islam, Leiden: Brill.

Further Reading

Abou El Fadl, Khaled (2001) *Rebellion and Violence in Islamic Law*, Cambridge: Cambridge University Press.

Black, Deborah L. (1996) "Al-Farabi," in Nasr and Leaman, eds., *History of Islamic Philosophy*.

Daiber, Hans (1996) "Political Philosophy," in Nasr and Leaman, eds., *History of Islamic Philosophy*.

Ibn Taymiyya (1982) *Public Duties in Islam: The Institution of the Hisba* [*al-Hisba fi'l-Islam*], trans. Muhtar Holland, Leicester: The Islamic Foundation.

Johansen, Baber (2008) "A Perfect Law in an Imperfect Society: Ibn Taymiyya's Concept of 'Governance in the Name of the Sacred Law,'" in Peri J. Bearman, Wolfhart Heinrichs, Bernard G. Weiss, eds., *The Law Applied: Contextualizing the Islamic Shari'a: A Volume in Honor of Frank E. Vogel*, I.B. Tauris.

Kamali, Mohammad Hashim (2002) *Freedom, Equality and Justice in Islam*, Cambridge, UK: Islamic Texts Society.

Khadduri, Majid (1984) *The Islamic Conception of Justice*, Baltimore, MD: The Johns Hopkins University Press.

Kuran, Timur (1989) "On the Notion of Economic Justice in Contemporary Islamic Thought," *International Journal of Middle East Studies* 21: 171–191.

March, Andrew F. (2010) "Taking People as They Are: Islam as a 'Realistic Utopia' in the Political Theory of Sayyid Qutb," *American Political Science Review* 104: 189–207.

Shari'ati, Ali (1979) *On the Sociology of Islam*, trans. Hamid Algar, Oneonta, NY: Mizan Press.

Part II

POLITICAL THEORIES AND IDEOLOGIES

20
ANARCHISM

Roderick T. Long

1. Defining Anarchism

Anarchism is a political philosophy, or better, a family of political philosophies, favoring a social order based on voluntary association and rejecting the legitimacy of the state. Although for most of its adherents anarchism means an opposition not solely to the state but also to various broader forms of hierarchy, domination, or aggression (with the details differing from theorist to theorist), anarchists agree at the very least on the desirability of abolishing the state in particular, whatever else they might agree or disagree about.

The Greek word *anarkhia* means the absence of a ruler or *arkhon*; that the term has traditionally been associated with disorder and social chaos is due solely to the presumption, rejected by anarchists, that such chaos is the natural result of the absence of rulers. Indeed, for anarchists it is precisely uncoerced, anarchic association that is responsible for social order, while government produces mainly disorder.

Throughout most of its history, the word "anarchist" has been used as a term of abuse; the first thinker to embrace the term and apply it to his own ideas appears to have been the French writer Pierre-Joseph Proudhon (1809–1865) in 1840, and it took some time for the term to catch on with most other anarchists. The term remains difficult to define precisely, inasmuch as the question of which thinkers and positions count as genuinely anarchist has long been a matter of intense debate among self-described anarchists, with virtually every faction having been read out of the movement by some other faction at some point.

2. Before Modern Anarchism

Visions of a utopian society without government have been put forward by a number of premodern thinkers, including Stoics such as Zeno of Citium (Greece, fourth–third century BCE), Epicureans such as Diogenes of Oenoanda (Anatolia, third century CE), and Taoists such as Bao Jingyan (China, fourth century CE). While such premodern proposals for anarchist societies were often premised on the precondition of radical moral transformation, in the Islamic world anarchist thinkers like al-Asamm and Abbad ibn Sulayman, in ninth-century Basra, appealed to precisely the opposite premise—that people are not sufficiently virtuous either to produce a trustworthy ruler or to agree upon one; to this the Najdiyya sect added the claim that rulership is incompatible with the ideal of human equality (Crone 2000).

In medieval Christian thought, Augustine's (fourth–fifth century CE) view that the state is the consequence of original sin (by contrast with Thomas Aquinas's view that

the state is natural and so would have been appropriate for human beings even in an unfallen condition of grace) doubtless contributed to millenarian visions of a stateless future; in subsequent centuries this vision was translated into action by religious movements such as the Anabaptists, who rejected private property, military service, and participation in human government. A later movement of similar orientation was led by the English reformer Gerrard Winstanley (1609–1676), whose "Diggers" or "True Levellers" not only rejected, on scriptural grounds, both private land ownership and the authority of the state (and indeed all authority of one human being over another), but sought to put these ideas into practice by taking over undeveloped land and cultivating it along communal lines. The following century would see the emergence of secular forms of anarchism as well.

3. Roots of Modern Anarchism

A number of historical developments appear to have contributed to the development of modern anarchist ideas. The growth of anti-monarchical sentiment generally favored republicanism, but arguments against monarchy—particularly those stressing political equality and the need for consent of the governed—were sometimes extended to critiques of government as such. European fascination with the New World, and the impression (whether correct or not) that Indian tribes lived without government, helped to make statelessness seem more viable. The rise of industrial capitalism would also spur anarchist reflections, both positively (with market competition being seen as a source of social order distinct from government) and negatively (with the dominance of wealthy elites under capitalism demonstrating that even democratic institutions were not enough to establish true freedom and equality).

Anarchist ideas were also given an impetus by the attempts of political thinkers to respond to the arguments of Thomas Hobbes (1588–1679). In his 1651 *Leviathan*, Hobbes had painted a dismal picture of the "state of nature," or life prior to the rise of government, arguing that without external enforcement there would be insufficient trust to ground cooperation and social order; the moral he drew was that, given the fragility of cooperation, and the preferability of nearly any government to none, subjects of oppressive rulers should remain content with their governments rather than risk plunging society into anarchy by attempting to overthrow them. In order to respond to Hobbes, thinkers more sympathetic to revolution, even if they held no brief for anarchy, felt the need to argue that the state of nature was less horrendous than the "war of all against all" depicted by Hobbes.

In his *Two Treatises of Government*, John Locke (1632–1704) argued that inasmuch as God has made us all equal, none can exercise legitimate authority over another except either (a) by that other's consent, or (b) as a response to acts of aggression by the other. Hence anarchy is, morally, the default position. Locke further maintained that recognition of at least the rudiments of the moral law is sufficiently widespread that we could expect there to be a fair bit of social order, including private property, without the state—enough order to make anarchy preferable, if not to limited constitutional government, at least to absolute monarchy.

Jean-Jacques Rousseau (1712–1778) also depicted anarchy favorably in his 1754 *Discourse on the Origin of Inequality*. By contrast with Locke, Rousseau envisions the state of nature as a primitivist utopia without private property or selfish motivations, and takes both Hobbes and Locke to task for attributing to pre-political humanity the sorts

of motives, ideas, and institutions that are in fact the product of developed civilization. Despite presenting the state of nature in a largely positive light, Rousseau does not advocate its (re)establishment; he describes it in the *Discourse* as unstable and unrecoverable, and in his 1762 *Social Contract* he makes clear that the task of political theory is not to strike off the chains of government but to legitimate them. Nevertheless, by presenting morally attractive (though deeply different) portraits of stateless society, both Locke and Rousseau helped to inspire positive commitments to anarchism in later thinkers.

One such thinker, possibly, was Edmund Burke (1729–1797), whose 1756 *Vindication of Natural Society* is one of the first explicit secular defenses of stateless society in the modern era. Scholars continue to debate whether Burke's anarchism was meant sincerely or, as he himself claimed, satirically (Kramnick (1977) makes a compelling case that it was some of each) but in any case the *Vindication* shows that anarchist ideas were in the air. And in his 1775 speech on *Conciliation With the Colonies*, Burke did point to the persistence of social order in the American colonies during the Revolution as evidence of the viability of anarchy.

The same example is invoked by Burke's archrival Thomas Paine (1737–1809) in his 1791–92 *Rights of Man*, where he argues that most cooperative social order is the result not of government but of a combination of self-interest and moral sentiment, and so would survive even without the state. While Paine's system is broadly Lockean, he is still more optimistic about the prospects for stateless social order than Locke had been, perhaps owing to the influence of invisible-hand explanations of social order from theorists like Adam Smith (whom Paine praises at Burke's expense). Another Smith devotee, French economist Jean-Baptiste Say (1767–1832), went even further, suggesting that it might one day be possible to dispense with the state altogether (Say 2003a: 101; 2003b: 324–7, 484). Thomas Jefferson (1743–1826), an admirer of both Locke and Smith, suggested in a 1787 letter to James Madison that anarchy might be the best political system if only it could be applied to large populations; while French historian Augustin Thierry (1795–1856), a follower of Say's, developed a theory of class struggle that made the state the linchpin of the ruling class's power, and forecast a time when the principle of government would give way to the principle of association (Raico 1993). And as early as 1793, the young Johann Gottlieb Fichte (1762–1814) was defending the right of secession down to the individual in his *Contribution to the Rectification of the Public's Judgment of the French Revolution*.

The contemporary division of anarchists into (misleadingly labeled) "individualist anarchist" and "social anarchist" camps may be seen (albeit with some oversimplification) as in part a reflection of the influence of these competing visions of stateless society—the individualist anarchists drawing more on the tradition of Locke and Smith, embracing private property, and seeing market mechanisms as an important means of coordinating individuals' actions in the absence of state enforcement (contrary to a popular misconception, most individualist anarchists have no objection to social cooperation and organization); and the social anarchists drawing more on the tradition of Rousseau, rejecting private property, and seeing market relations as a form of exploitation and domination to be abolished along with the state.

4. Godwin and Proudhon

Paine did not take the viability of anarchy as a reason to embrace anarchism himself; but Paine's arguments were explicitly credited by William Godwin (1756–1836)—arguably the first indisputable exemplar of modern secular anarchism—with convincing him of

the dispensability of the state. In his 1793 *Enquiry Concerning Political Justice*, Godwin maintained that, given the vast proportion by which the ruled generally outnumber their rulers, the power of governments necessarily rests on public opinion rather than force; Godwin concluded that public opinion must therefore be sufficient to maintain social order on its own, without the intermediary of government. Godwin does not fit easily into either the individualist anarchist or social anarchist category; he maintained that property should be shared rather than held privately—but also that private property should not be interfered with while it lasts, since the communal alternative must be pursued via moral education rather than by expropriation.

Another thinker who defies easy categorization is the aforementioned Proudhon, whose most important works include *What Is Property?* (1840), *System of Economic Contradictions* (1846), and *General Idea of the Revolution* (1851). While famously defining private property as a form of robbery, Proudhon nevertheless defended an alternative, less absolutist form of individual holding, called *possession*, depending on ongoing occupancy and use, and subject to reallocation with changes in population; such possession he saw as a synthesis of the opposed principles of communism on the one hand and bourgeois property on the other. Proudhon attacked the economic phenomena of rent, profit, and interest as expressions of property rather than possession; he called for the replacement both of capitalist firms and of the state by free associations of workers, under a system he called "mutualism." In response to those who argued that liberty presupposes a framework of order provided by the state, Proudhon maintained that "liberty is the mother, not the daughter, of order." In a later work, *Theory of Property* (1866), Proudhon offered a qualified endorsement of private property in the absolutist sense as a counterweight against excessive social and political power. It is a testament to Proudhon's complexity that individualist anarchists such as Benjamin Tucker and social anarchists such as Mikhail Bakunin have alike claimed him as their standard-bearer.

5. The Rise of Individualist Anarchism

The first half of the nineteenth century saw the emergence of a number of individualist anarchist thinkers, most of them also involved in the abolitionist, feminist, and labor movements, and frequently the free-love movement (i.e., the separation of state and sex, or more broadly of domination and sex) as well (Sears 1977; Blatt 1990; McElroy 1991; Perry 1993; Passet 2003; Martin 2009).

In England, Thomas Hodgskin (1787–1869) radicalized classical political economy and the Lockean theory of property, arguing that governments existed to enforce artificial property claims based on political privilege at the expense of natural property claims based on labor (Stack 1998).

In the United States, Josiah Warren (1798–1874) was circulating *The Peaceful Revolutionist* (1833), which has been claimed as the first anarchist periodical. Warren defended the complete sovereignty of the individual, and also argued for the cost principle—i.e., the moral requirement that producers should sell their products at no higher price than that necessary to recompense them for the cost of producing them; he was also involved in founding intentional communities that put his ideas into practice (Sartwell 2011). The Warrenite cause was taken up by a number of able disciples, most notably Stephen Pearl Andrews (1812–1886).

Back on the other side of the Atlantic, three very different contributions to individualist anarchism were in development. In Germany, Max Stirner (1806–1856), in his

lyrical manifesto *The Ego and Its Own* (1845), rejected as "spooks" all abstract norms—such as God, government, rights, and morality—that seek to limit individual self-assertion; individuals should regard all of reality as their property, and cooperate with one another only on the basis of mutual and temporary self-interest. Stirner's book is often taken to be the principal guiding text of individualist anarchism; one scholar even describes the entire individualist anarchist tradition as merely "a footnote to Max Stirner" (Graham 2005: xiii). In fact, however, Stirner's influence on individualist anarchism has been fairly slight, with most of its thinkers either being unaware of Stirner or rejecting him. (Even the most prominent American Stirnerite, Benjamin Tucker, had already developed his anarchist system in fairly full detail before ever reading Stirner.)

In France, Belgian-born economist Gustave de Molinari (1819–1912), working in the tradition of Say and Thierry, seems to have been the first to describe how security firms competing under conditions of *laissez-faire* might replace the monopoly state in the provision of the traditionally governmental function of protection against crime. Molinari also defended private property and the wage system, but argued for the need for workers to associate into "labor-exchanges" to redress the imbalance of bargaining power between labor and capital.

In England, Herbert Spencer (1820–1903) developed an evolutionary theory of social progress having as its culmination a stateless society governed by a Law of Equal Freedom. By contrast with Stirner, Spencer saw the process of civilization as a development toward greater and greater altruism; by contrast with Molinari, Spencer placed his hopes for an anarchist future in this blossoming of altruism rather than in the operation of economic incentives. While Spencer has acquired the reputation of being a conservative defender of bourgeois privilege, he rejected private ownership of land (while endorsing private ownership of other means of production), and called for the abolition of the wage system in favor of enterprises collectively owned and managed by their workers. The doctrine for which he is perhaps best known—that the poor and sick should be allowed to die off in order to improve the species—is one he never maintained, and in fact vehemently repudiated, defending a duty of positive beneficence, though he did criticize forms of charity that relied on state compulsion or that he interpreted as fostering dependence.

Both Molinari and Spencer retreated from a complete antistatist position in their later years; but Spencer, at least, inspired a cadre of radical Spencerians such as Auberon Herbert, Wordsworth Donisthorpe, and Victor Yarros, who continued to develop the anarchistic side of Spencer's social theory. Molinari may likewise have found at least a partial disciple in his countryman Paul Émile de Puydt, whose doctrine of "panarchy" called for different systems of government to compete for citizens' allegiance within the same geographical territory; and French anarchist Anselme Bellegarrigue put forward ideas bearing various similarities to those of Proudhon, Stirner, and Molinari, though the extent to which he was influenced by any of these thinkers remains unclear.

In the United States, individualist anarchism continued to play an important role in the abolitionist movement, often in conjunction with a struggle to bring greater empowerment to all workers, slave and free (Perry 1993; Martin 2009; Blatt 1990). Abolitionists who based their opposition to slavery on individualist anarchist principles included (in addition to the aforementioned Warren and Andrews) Adin Ballou (1803–1890), William Lloyd Garrison (1805–1879), William Batchelder Greene (1819–1878), Ezra Heywood (1829–1893), Charles Lane (1800–1870), Nathaniel P. Rogers (1794–1846), Lysander Spooner (1808–1887), and of course Henry David Thoreau (1817–1862).

Garrison and Spooner were among the most influential. Garrison was an advocate of nonresistance who favored promoting emancipation through education and persuasion rather than through political methods; he famously denounced the U.S. Constitution as "a covenant with death and an agreement with hell" for its compromises with slavery. Spooner, by contrast, was no pacifist, advocating liberation through federal intervention if possible and violent insurrection if necessary, and lending his support to John Brown's vigilante activities against slaveholders; he also held that the Constitution, if interpreted according to its strict meaning (with contested normative terms being read in light of a somewhat anarchistic version of natural law) rather than the unknowable intentions of its ratifiers, discountenanced slavery. With the coming of the Civil War, however, Garrison disappointed many of his pacifist supporters by embracing the Union cause, while Spooner, convinced of the essential hypocrisy of both Union and Confederacy, denounced the Constitution he had once defended, arguing that human legislation has no claim on our obedience except insofar as it rests on the genuine and explicit consent of the governed as individuals.

American individualist anarchists also campaigned for sexual freedom and gender equality, with prominent "free-love" periodicals being *The Word* (1872–1893), edited by Ezra Heywood and his wife Angela, and *Lucifer the Lightbearer* (1893–1907), edited by Moses Harman and his daughter Lillian Harman. Yet individualist anarchist principles were also enunciated by socially conservative Christians like educator David Lipscomb (1831–1917), founder of the university that bears his name.

The most influential individualist anarchist in the United States during the period after the Civil War was Benjamin Tucker (1854–1939) of Boston, editor of the periodical *Liberty* (1881–1908). Tucker's anarchism was a blend of Warren, Proudhon, Stirner, and Spencer, with individual rights to liberty and (occupancy-and-use-based) property being grounded in contractual agreements among self-interested individuals, rather than in abstract moral law. Tucker's proposal for competing security firms also resembles Molinari's, though whether through influence or coincidence is unclear. The reduction of price to labor cost was for Tucker a prediction of the results to be expected from the abolition of state-supported privilege for the capitalist class, not an independent moral imperative as it had been for Warren and Andrews.

Tucker called his position "consistent Manchesterism," since he shared the Manchesterites' endorsement of competition and rejection of monopoly, but accused them of inconsistently failing to oppose the monopolies of security, currency, and land enjoyed by the state and the capitalist class. Tucker also called his radical free-market proposals "socialist" (his disciple Francis Tandy would prefer "voluntary socialism") since their implementation would lead to the abolition of class privilege and the wage system; though as the term "socialism" became increasingly associated with state control and the suppression of free exchange he began to use it pejoratively instead. Tucker's attitude toward labor unions was sympathetic but cautious, worrying about the dangers of regimentation, bureaucratization, and seduction into a quest for political power; but some of Tucker's associates, such as Dyer Lum (1839–1893) and Joseph Labadie (1850–1933), embraced the union movement and promoted broadly Tuckerite ideas within it.

Like many of her generation, Voltairine de Cleyre (1866–1912) was radicalized by the blatant injustice of the 1886–87 Haymarket trials. Initially an individualist anarchist in Tucker's circle (though never sharing his commitment to egoism), and a close associate of Lum's, de Cleyre was one of the first to describe herself, if in a somewhat

tongue-in-cheek manner, as a "capitalistic anarchist," although she more often followed Tucker in using "capitalist" as a pejorative. In later years she retreated from identification with any one school of anarchism, favoring a pluralistic vision where various different economic arrangements, whether individualistic or communistic, could coexist; she also attacked marriage as an impediment to individual growth.

Across the Atlantic, the cause of individualist anarchism was kept alive by Scottish-born German writer John Henry Mackay (1864–1933) and French writer Émile Armand (1872–1963), both of whom, working broadly in the Stirner-Tucker tradition, extended the free-love principles of the American individualists to a defense of homosexual emancipation (and in Armand's case, polyamory); as well as by German sociologist Franz Oppenheimer (1864–1943), who distinguished between the "economic means" (production and trade) and "political means" (theft and exploitation) of satisfying one's needs, and defined the state as the organization of the political means.

6. The Rise of Social Anarchism

In the meantime, anarchist theories of a more communist or collectivist character had been developing as well. One important pioneer is French anarcho-communist Joseph Déjacque (1821–1864), who argued, against Proudhon and the individualists, that what workers have a right to is not the product of their labor but, rather, the satisfaction of their needs. Déjacque appears to have been the first thinker to adopt the term "libertarian" for his position; hence "libertarianism" initially denoted a communist rather than a free-market ideology.

The most famous communistic ideology, Marxism, has served as both inspiration and adversary for social anarchists. Marx and Engels arguably count as anarchists themselves, since they call for the (eventual) withering away of the state, and establishment of a stateless communist society; but their advocacy of proletarian dictatorship in the interim, their relentless attacks on anarchism both in print and in the First International, and the literal violence against anarchists perpetrated by regimes calling themselves Marxist, have earned them a somewhat ambiguous place in the history of anarchist thought. Although favorable views of Marxism are, unsurprisingly, most common in the social anarchist wing, many aspects of Marxist thought find admirers among individualist anarchists (Carson 2004) and even anarcho-capitalists (Hoppe 1990); and Marx himself had been influenced by anarchist thinkers of a more individualist temper, including Thierry, Hodgskin, Proudhon, and Stirner.

Russia soon became a hotbed of both Marxist and anarchist ideas; three of the most prominent Russian anarchist thinkers—all members of the nobility—were Mikhail Bakunin (1814–1876), Lëv Tolstoj (1828–1910), and Pëtr Kropotkin (1842–1921). Bakunin's "collectivist" version of anarchism, which assigned control over the means of production to federated workers' collectives and contemplated the continued existence of wages and monetary exchange, followed a middle path between two of Bakunin's chief influences, Proudhon's market-based individualism and Marx's marketless communism. Bakunin sharply criticized Marx's plan for a "dictatorship of the proletariat" during the transitional period between the overthrow of capitalism and the withering away of the state, arguing that this would simply mean a new ruling class and the perpetuation of oppression; and the two thinkers clashed within the First International. Kropotkin for his part, while sharing Bakunin's misgivings about Marxism, rejected

Bakunin's collectivism, with its retention of wages and money, in favor of pure communism. In his hostility to money, Tolstoj was perhaps closer to Kropotkin.

Bakunin was an advocate of violent revolution and a fierce opponent of all religion, while Tolstoj defended a mostly pacifist position inspired by Christianity and thought that inner moral transformation must precede political reform. Kropotkin developed an evolutionary and ecological approach to anarchism, stressing the role of mutual aid in human history and in the animal kingdom generally.

Anarcho-communist ideals, largely though not solely through Kropotkin's influence, were widely influential, being promulgated by Élisée Reclus (1830–1905) in France, Gustav Landauer (1870–1919) in Germany, Errico Malatesta (1853–1932) in Italy, Emma Goldman (1869–1940) and Alexander Berkman (1870–1936) in the United States, arguably Oscar Wilde (1854–1900) in Britain, and by still others in Spain, Asia, and Latin America. One movement having a high degree of overlap with both collectivist and communist forms of anarchism is anarcho-syndicalism, in which the means of production are owned and their use directed by federations of labor unions characterized by bottom-up democratic self-management; a major anarcho-syndicalist theorist was Rudolf Rocker (1873–1958).

In the late nineteenth and early twentieth centuries, anarchism, particularly social anarchism, played an important role in popular life, especially the labor movement. Anarchist militias played important roles in both the Russian Revolution and the Spanish Civil War, fighting (unsuccessfully, in the event) for an alternative to Marxism and reactionary conservatism alike. On an individual level, "direct action" and "propaganda of the deed" encompassed acts of terrorism, a controversial and debated tactic within the anarchist community, and one that has continued to shape public perceptions of anarchism to this day; the list of state dignitaries assassinated by anarchists includes a French president, an American president, an Italian king, a Russian tsar, an Austrian empress, and two Spanish prime ministers. In turn, the anarchist movement suffered heavily from state repression, including the mass arrest and deportations of anarchists in the United States (the "Palmer Raids," 1919–20) and, in subsequent decades, the still more sanguinary fate that awaited anarchists in Europe in regions under fascist, Nazi, and Communist rule.

7. The Rise of Anarcho-capitalism

In the 1930s through the 1950s, individualist anarchism reemerged on the American right, as part of the fledgling free-market libertarian movement, with such authors as Albert Jay Nock (1870–1945), Rose Wilder Lane (1886–1968), and Robert LeFevre (1911–1986). In the 1960s and 70s, under the influence (though without the approval) of radical individualist philosopher Ayn Rand and radical individualist economist Ludwig von Mises, thinkers like Murray Rothbard (1926–1995), Karl Hess (1923–1994), Roy A. Childs Jr. (1949–1992), Samuel E. Konkin III (1947–2004), George H. Smith (b. 1949), Randy E. Barnett (b. 1952), and Linda and Morris Tannehill developed these anarchist tendencies into a full-blown theory of *anarcho-capitalism*. (See Rothbard 2006.) These anarcho-capitalists accept private property and economic *laissez-faire* (typically on pragmatic grounds, natural-rights grounds, or both), and take the state's monopoly of the rights-protection industry to be an illegitimate interference with free competition; in place of the state, they revive the Molinari–Tucker model of competing security firms. Another important anarcho-capitalist thinker to emerge during this era is David Friedman (b. 1945), son of economist Milton Friedman, who points to the

success of the medieval Icelandic legal system as evidence of the viability of a non-monopolistic legal order (Friedman 1989). The generic term "market anarchism" is sometimes used to include both anarcho-capitalism and the market-friendly varieties of more traditional individualist anarchism.

While most "minarchist" libertarians (i.e., advocates of the minimal state rather than anarchy) had an adversarial relationship with the New Left, many of the anarcho-capitalists—particularly Rothbard, Hess, Childs, and Konkin—were more sympathetic, maintaining that defenders of the free market had much to learn from leftist analyses (as well as vice versa). In particular, they rejected the right-wing tendency to see big business as the victim of government regulation, arguing instead that big business represents an unjustly privileged elite that is more often the beneficiary of such regulation (Rothbard 1965a, 1965b; Childs 1971). Konkin notably abandoned the term "anarcho-capitalism" in favor of "agorism" (from *agora*, the Greek word for market), situating himself on the anti-capitalist left (Konkin 1983, 2002).

Anarcho-capitalism was brought to academic notice largely via Robert Nozick's criticisms in *Anarchy, State, and Utopia* (Nozick 1974: 3–146). Nozick maintained that a competitive security market lacked both pragmatic superiority—because conflicts among security firms, whether resolved violently by conquest or peacefully by arbitration, would in either case tend to give rise to a territorial monopoly and so reestablish the state—and moral superiority, since a security firm's right to protect its clients by prohibiting risky adjudicative procedures used by other firms could lead to the establishment of a monopoly state via an invisible-hand process that would violate nobody's rights. In response, anarcho-capitalists have argued that security firms would be compelled, by their need to retain customers, to choose the lower-cost option of arbitration over the higher-cost option of violence; that arbitration networks among firms do not constitute a state so long as they allow free entry; and that the right to prohibit risky adjudicative procedures is too weak to license a monopoly (Barnett 1977; Childs 1977; Davidson 1977; Paul 1977; Rothbard 1977; Sanders 1977; Osterfeld 1980; Roark 1986; Young 1986; Skoble 2008b). More recently the debate has turned to the question of whether arbitration networks could profitably exclude free entry (Cowen 1992; Caplan and Stringham 2003; Cowen and Sutter 2005).

Social anarchists, while generally recognizing individualist anarchism as a genuine (albeit mistaken) form of anarchism, largely deny that anarcho-capitalism counts as anarchist (and similarly reject the anarchist credentials of such forerunners as Molinari and Spencer), on the grounds that anarchism, whether social or individualist, has traditionally been anti-capitalist. This dispute is partly terminological (since by "capitalism" anarcho-capitalists mean a completely free market, while most individualist anarchists have reserved the term for a system of privilege which they take to result from government intervention), but also substantive: social anarchists charge that anarcho-capitalists countenance economic relationships that "genuine" individualist anarchists reject, such as rent, profit, interest, hierarchical workplaces, and the wage system, and so conclude that anarcho-capitalism is simply a rationalization for corporate power and existing property holdings. But it is difficult to find a criterion that draws the division neatly: Spooner, for example, is recognized by social anarchists as a genuine individualist anarchist despite his support for rent, while Konkin and Friedman are rejected despite their opposition to hierarchical workplaces and the wage system; and many anarcho-capitalists have attacked corporate power (Childs 1971) and called for thoroughgoing redistribution to rectify past injustice (Rothbard 1969, 2008). Anarcho-capitalists for

their part tend to challenge the libertarian credentials of social anarchists, generally on equally contestable grounds.

8. Recent Anarchism

The political ferment of the mid-twentieth century saw the revival of social anarchism as well (see, e.g., Ward 1992; Goodman 1994; Zinn 1997; Graeber 2004; Chomsky 2005); in recent decades, social anarchists have played a leading role in the anti-globalization and Occupy movements. The growth of the internet has given additional impetus to the propagation of anarchist ideas, both social and individualist.

At a theoretical level, both social anarchists (Taylor 1982) and anarcho-capitalists (Narveson 1988; De Jasay 1997; Stringham 2005; Skoble 2008a) have employed game-theoretic arguments to show that cooperation represents a stable equilibrium under stateless conditions; likewise, both social anarchists (Clastres 1989; Scott 2010) and anarcho-capitalists (Friedman 1979; Anderson and Hill 2004; Benson 2011) have appealed to historical examples of stateless societies to demonstrate anarchism's viability.

Although there are exceptions (Proudhon being an especially egregious one), anarchists have traditionally embraced feminist concerns, seeing patriarchy, statism, and capitalist privilege as mutually reinforcing forms of social control (McElroy 1991; Long and Johnson 2005), and contemporary anarcha-feminists (e.g., Brown 1993; Lilith 2011) have made the same identification.

The anarchist community continues to be the scene of vigorous debate on many topics, including the legitimacy of private property (or, among those accepting private property, the legitimacy of absentee land ownership) and the extent to which prevailing capitalist economic forms are the result of free markets or of government intervention. (Interestingly, the rejection of intellectual property seems to have become the dominant position among all schools of anarchism.) Also in dispute are the appropriate moral foundations for anarchism (e.g., consequentialist vs. deontological), as well as whether anarchism should respect the authority of morality at all. With respect to religion, for some adherents anarchism's anti-authoritarianism entails atheism, and the slogan "no gods, no masters" is popular in the movement; but for others, it is precisely respect for God's authority that discountenances all human authority. And views on technology vary from the "technoanarchist" embrace of high-tech urban life (Kostelanetz 1999), through the preference for decentralized, human-scale, more environmentally responsible forms of technology (Goodman 1994; Bookchin 2004), all the way to the complete rejection of industry, civilization, and even language itself, as found in anarcho-primitivism (Zerzan 1994).

Anarchists are also divided over tactics and methods; most reject or at least de-emphasize electoral politics in favor of a variety of approaches ranging from violent revolution, through anti-corporate hacktivism, to education, counter-economic activity, nonviolent mass civil disobedience, and building alternative institutions. Moreover, a body of overlapping but by no means unitary social anarchist approaches, variously called poststructuralist anarchism, post-left anarchism, and post-anarchism, has challenged the rigidity of traditional anarchist ideology and activism, and emphasized the importance of achieving autonomous modes of experience and practice in everyday life (May 1994; Black 1997; Bey 2003; Jun and Wahl 2010; Nadia C. n.d); other social anarchists such as Murray Bookchin have criticized these approaches as a self-indulgent "lifestyle

anarchism" that distracts from the task of organizing collective action for meaningful social change (Bookchin 1995).

9. Terminological Note

The term "philosophical anarchism" is ambiguous. Initially it referred to the position of anarchists who favored educational rather than violent means to the establishment of an anarchist society (Yarros 1936); but in recent years it has come to mean the thesis that the state's laws do not in themselves give citizens any moral obligation to obey—*without* any further implication either as to its being morally impermissible for the state to enforce such laws, or as to the abolition of states being desirable. This newer sense, which renders philosophical anarchism not strictly a species of anarchism, seems to have arisen out of a misunderstanding of social anarchist Robert Paul Wolff's 1970 *In Defense of Anarchism*, which is widely interpreted as defending merely the narrow thesis, despite the book's closing call (Wolff 1998: 78–82) for a literally stateless society. (At least one self-described anarchist (Barnett 2005) defends the reverse position; that laws may be morally binding on citizens because they are commanded by a state of the right sort, even though it is impermissible for that state to enforce them, or indeed to exist at all.)

The term "left-libertarianism" has at least three meanings. In its oldest sense, it is a synonym either for anarchism in general or social anarchism in particular. Later it became a term for the left or Konkinite wing of the free-market libertarian movement, and has since come to cover a range of pro-market but anti-capitalist positions, mostly individualist anarchist, including agorism and mutualism, often with an implication of sympathies (such as for radical feminism or the labor movement) not usually shared by anarcho-capitalists (Carson 2007, 2008; Johnson 2008; Chartier 2011; Chartier and Johnson 2011). In a third sense it has recently come to be applied to a position combining individual self-ownership with an egalitarian approach to natural resources (Vallentyne and Steiner 2001, 2007); most proponents of this position are not anarchists.

Related Topics

Marxism and Contemporary Political Thought; Natural Law and Rights Theory; Left-Libertarianism; Libertarianism; Equality; Freedom; Autonomy; Power; Authority and Legitimacy

References

Anderson, Terry L. and Hill, Peter J. (2004) *The Not So Wild, Wild West: Property Rights on the Frontier*, Stanford University Press.

Barnett, Randy E. (1977) "Whither Anarchy? Has Robert Nozick Justified the State?," *Journal of Libertarian Studies* 1 (1): 15–21. Available at: http://mises.org/journals/jls/1_1/1_1_3.pdf [accessed on 1 February 2012].

—— (1998) *The Structure of Liberty: Justice and the Rule of Law*, Oxford University Press.

—— (2005) "Libertarianism and Legitimacy: A Reply to Huebert," *Journal of Libertarian Studies* 19 (4): 71–78. Available at: http://mises.org/journals/jls/19_4/19_4_5.pdf [accessed on 1 February 2012].

Benson, Bruce L. (2011) *The Enterprise of Law: Justice Without the State*, 2nd edn, Independent Institute.

Bey, Hakin (2003) *T.A.Z. The Temporary Autonomous Zone, Ontological Anarchy, Poetic Terrorism*, 2nd edn, Autonomedia.

Black, Bob (1997) *Anarchy After Leftism*, CAL Press.

Blatt, Martin (1990) *Free Love and Anarchism: The Biography of Ezra Heywood*, University of Illinois Press.
Bookchin, Murray (1995) *Social Anarchism or Lifestyle Anarchism*, AK Press. Available at: http://www.studiesinanti-capitalism.net/StudiesinAnti-Capitalism/Documents_TWO_files/SocialAnarchismOrLifestyleAnarchism.pdf [accessed on 1 February 2012].
—— (2004) *Post-Scarcity Anarchism*, AK Press.
Brooks, Frank H., ed. (1994) *The Individualist Anarchists: An Anthology of* Liberty *(1881–1908)*, Transaction Publishers.
Brown, Susan L. (1993) *The Politics of Individualism: Liberalism, Liberal Feminism and Anarchism*, Black Rose.
C., Nadia (n.d) "Your Politics Are Boring As Fuck." Available at: http://www.crimethinc.com/texts/selected/asfuck.php [accessed on 1 February 2012].
Caplan, Bryan and Stringham, Edward (2003) "Networks, Law, and the Paradox of Cooperation," *Review of Austrian Economics* 16 (4): 309–326.
Carson, Kevin A. (2004) *Austrian and Marxist Theories of Monopoly-Capital: A Mutualist Synthesis*, Libertarian Alliance. Available at: http://www.libertarian.co.uk/lapubs/econn/econn102.pdf [accessed on 1 February 2012].
—— (2007) *Studies in Mutualist Political Economy*, BookSurge. Available at: http://www.mutualist.org/sitebuildercontent/sitebuilderfiles/MPE.pdf [accessed on 1 February 2012].
—— (2008) *Organization Theory: A Libertarian Perspective*, BookSurge. Available at: http://www.mutualist.org/sitebuildercontent/sitebuilderfiles/otkc11.pdf [accessed on 1 February 2012].
Chartier, Gary (2011) *The Conscience of an Anarchist: Why It's Time to Say Good-Bye to the State and Build a Free Society*, Cobden Press.
—— and Johnson, Charles W., eds. (2011) *Markets Not Capitalism: Individualist Anarchism Against Bosses, Inequality, Corporate Power, and Structural Poverty*, Minor Compositions. Available at: http://radgeek.com/gt/2011/10/Markets-Not-Capitalism-2011-Chartier-and-Johnson.pdf [accessed on 1 February 2012].
Childs, Roy A., Jr. (1971) "Big Business and the Rise of American Statism; part 1," *Reason* (February): 12–18, and part 2 (March): 9–12. Available at: http://praxeology.net/RC-BRS.htm [accessed on 1 February 2012].
—— (1977) "The Invisible Hand Strikes Back," *Journal of Libertarian* Studies 1 (1): 23–33. Available at: http://mises.org/journals/jls/1_1/1_1_4.pdf [accessed on 1 February 2012].
Chomsky, Noam (2005) *Chomsky on Anarchism*, AK Press.
Clastres, Pierre (1989) *Society Against the State: Essays in Political Anthropology*, trans. Robert Hurley and Abe Stein, Zone.
Cowen, Tyler (1992) "Law as a Public Good: The Economics of Anarchy," *Economics and Philosophy* 8: 249–267.
—— and Sutter, Daniel (2005) "Conflict, Cooperation and Competition in Anarchy," *Review of Austrian Economics* 18 (1): 109–115.
Crone, Patricia (2000) "Ninth-Century Muslim Anarchists," *Past & Present* 167 (May): 3–28.
Davidson, James Dale (1977) "Note on *Anarchy, State, and Utopia*," *Journal of Libertarian Studies* 1 (4): 341–348. Available at: http://mises.org/journals/jls/1_4/1_4_8.pdf [accessed on 1 February 2012].
De Jasay, Anthony (1997) *Against Politics: On Government, Anarchy and Order*, Routledge.
Friedman, David (1979) "Private Creation and Enforcement of Law: A Historical Case," *Journal of Legal Studies* 8 (2): 399–415. Available at: http://www.daviddfriedman.com/Academic/Iceland/Iceland.html [accessed on 1 February 2012].
—— (1989) *The Machinery of Freedom: Guide to a Radical Capitalism*, 2nd edn, Open Court. Available at: http://www.daviddfriedman.com/The_Machinery_of_Freedom_.pdf [accessed on 1 February 2012].
Goodman, Paul (1994) *Decentralizing Power: Paul Goodman's Social Criticism*, ed. Taylor Stoehr, Black Rose Books.
Graeber, David (2004) *Fragments of an Anarchist Anthropology*, Prickly Paradigm Press.
Graham, Robert, ed. (2005) *Anarchism: A Documentary History of Libertarian Ideas, Volume One: From Anarchy to Anarchism (300 CE to 1939)*, Black Rose Books.
—— (2009) *Anarchism: A Documentary History of Libertarian Ideas, Volume Two: The Emergence of the New Anarchism (1939–1977)*, Black Rose Books.
—— (2012) *Anarchism: A Documentary History of Libertarian Ideas, Volume Three: The New Anarchism (1978–2010)*, Black Rose Books.
Hoppe, Hans-Hermann (1990) "Marxist and Austrian Class Analysis," *Journal of Libertarian Studies* 9 (2): 79–93. Available at: http://mises.org/journals/jls/9_2/9_2_5.pdf [accessed on 1 February 2012].

Johnson, Charles (2008) "Liberty, Equality, Solidarity: Toward a Dialectical Anarchism," in Long, Roderick T. and Machan, Tibor R. eds., *Anarchism/Minarchism: Is a Government Part of a Free Country?*, Ashgate: 155–182. Available at: http://radgeek.com/gt/2010/03/02/liberty-equality-solidarity-toward-a-dialectical-anarchism [accessed on 1 February 2012].

Jun, Nathan J. and Wahl, Shane (2010) *New Perspectives on Anarchism*, Rowman & Littlefield.

Konkin, Samuel Edward, III (1983) *New Libertarian Manifesto*, Koman Publishing. Available at: http://agorism.info/NewLibertarianManifesto.pdf [accessed on 1 February 2012].

—— (2002) "Smashing the State for Fun and Profit Since 1969." Available at: http://www.spaz.org/~dan/individualist-anarchist/software/konkin-interview.html [accessed on 1 February 2012].

Kostelanetz, Richard (1999) *Political Essays*, Autonomedia.

Kramnick, Isaac (1977) *The Rage of Edmund Burke: Portrait of an Ambivalent Conservative*, Basic Books.

Lilith (2011) *Gender Disobedience: Antifeminism and Insurrectionist Non-dialogue*, Anarchist Library. Available at: http://theanarchistlibrary.org/pdfs/letter/Lilith__Gender_Disobedience__Antifeminism_and_Insurrectionist_Non-dialogue_letter.pdf [accessed on 1 February 2012].

Long, Roderick T. and Johnson, Charles (2005) "Libertarian Feminism: Can This Marriage Be Saved?" Available at: http://charleswjohnson.name/essays/libertarian-feminism [accessed on 1 February 2012].

Long, Roderick T. and Machan, Tibor R., eds. (2008) *Anarchism/Minarchism: Is a Government Part of a Free Country?*, Ashgate.

Martin, James J. (2009) *Men Against the State: The Expositors of Individualist Anarchism in America, 1827–1908*, Ludwig von Mises Institute.

May, Todd (1994) *The Political Philosophy of Poststructuralist Anarchism*, Pennsylvania State University Press.

McElroy, Wendy, ed. (1991) *Freedom, Feminism and the State: An Overview of Individualist Feminism*, 2nd edn, Independent Institute.

McKay, Iain, Elkin, Gary, Neal, Dave, and Boraas, Ed (2008) *An Anarchist FAQ*. Available at http://www.infoshop.org/page/AnAnarchistFAQ [accessed on 1 February 2012].

Narveson, Jan (1988) *The Libertarian Idea*, Temple University Press.

Nozick, Robert (1974) *Anarchy, State, and Utopia*, Basic Books.

Osterfeld, David (1980) "Internal Inconsistencies in Arguments for Government: Nozick, Rand, and Hospers," *Journal of Libertarian Studies* 4 (3): 331–340. Available at: http://mises.org/journals/jls/4_3/4_3_7.pdf [accessed on 1 February 2012].

Passet, Joanne E. (2003) *Sex Radicals and the Quest for Women's Equality*, University of Illinois Press.

Paul, Jeffrey (1977) "Nozick, Anarchism and Procedural Rights," *Journal of Libertarian Studies* 1 (4): 337–340. Available at: http://mises.org/journals/jls/1_4/1_4_7.pdf [accessed on 1 February 2012].

Perry, Lewis (1993) *Radical Abolitionism: Anarchy and the Government of God in Anti-Slavery Thought*, University of Tennessee Press.

Raico, Ralph (1993) "Classical Liberal Roots of the Marxist Doctrine of Classes," in Yuri N. Maltsev, ed., *Requiem for Marx*, Ludwig von Mises Institute: 189–220.

Roark, Eric (1986) "Nozick's Failed Defense of the Just State," *Journal of Libertarian Studies* 21 (1): 5–39. Available at: http://mises.org/journals/jls/21_1/21_1_2.pdf [accessed on 1 February 2012].

Rothbard, Murray N. (1965a) "Left and Right: The Prospects for Liberty," *Left & Right* 1 (1): 4–22. Available at: http://mises.org/journals/lar/pdfs/1_1/1_1_2.pdf [accessed on 1 February 2012].

—— (1965b) "Liberty and the New Left," *Left & Right* 1 (2): 35–67. Available at: http://mises.org/journals/lar/pdfs/1_2/1_2_4.pdf [accessed on 1 February 2012].

—— (1969) "Confiscation and the Homestead Principle," *Libertarian Forum* 1 (6): 3–4. Available at: http://mises.org/journals/lf/1969/1969_06_15.pdf [accessed on 1 February 2012].

—— (1977) "Robert Nozick and the Immaculate Conception of the State," *Journal of Libertarian Studies* 1 (1): 45–57. Available at: http://mises.org/journals/jls/1_1/1_1_6.pdf [accessed on 1 February 2012].

—— (2006) *For a New Liberty: The Libertarian Manifesto*, 2nd edn, Ludwig von Mises Institute. Available at: http://mises.org/rothbard/foranewlb.pdf [accessed on 1 February 2012].

—— (2008) *The Ethics of Liberty*, New York University Press. Available at: http://mises.org/rothbard/ethics.pdf [accessed on 1 February 2012].

Sanders, John T. (1977) "The Free Market Model versus Government: A Reply to Nozick," *Journal of Libertarian Studies* 1 (1): 35–44. Available at: http://mises.org/journals/jls/1_1/1_1_5.pdf [accessed on 1 February 2012].

Sartwell, Crispin, ed. (2011) *The Practical Anarchist: Writings of Josiah Warren*, Fordham University Press.

Say, Jean-Baptiste (2003a) *Oeuvres Complètes, IV: Leçons d'Économie Politique*, Economica.

—— (2003b) *Oeuvres Complètes, V: Oeuvres morales et politiques*, Economica.
Scott, James C. (2010) *The Art of Not Being Governed: An Anarchist History of Upland Southeast Asia*, Yale University Press.
Sears, Hal D. (1977) *The Sex Radicals: Free Love In High Victorian America*, Regents Press of Kansas.
Skoble, Aeon J. (2008a) *Deleting the State: An Argument About Government*, Open Court.
—— (2008b) "Radical Freedom and Social Living," in Long, Roderick T. and Machan, Tibor R. eds., *Anarchism/Minarchism: Is a Government Part of a Free Country?*, Ashgate: 87–102.
Smith, George H. (1991) *Atheism, Ayn Rand, and Other Heresies*, Prometheus Books.
Stack, David (1998) *Nature and Artifice: The Life and Thought of Thomas Hodgskin*, Boydell & Brewer.
Stringham, Edward P., ed. (2005) *Anarchy and the Law: The Political Economy of Choice*, Transaction Publishers.
——, ed. (2007) *Anarchy, State and Public Choice*, Edward Elgar.
Taylor, Michael (1982) *Community, Anarchy, and Liberty*, Cambridge University Press.
Vallentyne, Peter and Steiner, Hillel, eds. (2001) *Left-Libertarianism and Its Critics: The Contemporary Debate*, Palgrave Macmillan.
—— (2007) *The Origins of Left-Libertarianism: An Anthology of Historical Writings*, Palgrave Macmillan.
Ward, Colin (1992) *Anarchy In Action*, Freedom Press.
Wolff, Robert Paul (1998) *In Defense of Anarchism; With a New Preface*, University of California Press.
Yarros, Victor S. (1936) "Philosophical Anarchism: Its Rise, Decline, and Eclipse," *American Journal of Sociology* 41 (4): 470–483.
Young, Frederic C. (1986) "Nozick and the Individualist Anarchist," *Journal of Libertarian Studies* 8 (1): 43–49. Available at: http://mises.org/journals/jls/8_1/8_1_4.pdf [accessed on 1 February 2012].
Zerzan, John (1994) *Future Primitive and Other Essays*, Autonomedia.
Zinn, Howard (1997) *The Zinn Reader: Writings on Disobedience and Democracy*, Seven Stories Press.

Further Reading

Brooks, Frank H., ed. (1994) *The Individualist Anarchists: An Anthology of* Liberty *(1881–1908)*, Transaction Publishers. An anthology of nineteenth-century American individualist anarchists associated with Benjamin Tucker's circle.

Chartier, Gary and Johnson, Charles W., eds. (2011) *Markets Not Capitalism: Individualist Anarchism Against Bosses, Inequality, Corporate Power, and Structural Poverty*, Minor Compositions. Available at: http://radgeek.com/gt/2011/10/Markets-Not-Capitalism-2011-Chartier-and-Johnson.pdf [accessed on 1 February 2012]. An anthology of individualist anarchist writings, mostly twentieth- or twenty-first-century and mostly anti-capitalist.

Graham, Robert, ed. (2005, 2009, 2012) *Anarchism: A Documentary History of Libertarian Ideas*, three volumes, Black Rose Books. A three-volume anthology of mostly social anarchist writings, from the early modern period through the present.

McKay, Iain, Elkin, Gary, Neal, Dave, and Boraas, Ed (2008) *An Anarchist FAQ*. Available at: http://www.infoshop.org/page/AnAnarchistFAQ [accessed on 1 February 2012]. Rather hostile to, and unreliable on, individualist anarchism generally and anarcho-capitalism specifically, but a wonderfully comprehensive compendium of information about social anarchism.

Stringham, Edward P., ed. (2007) *Anarchy and the Law: The Political Economy of Choice*, Transaction Publishers. An anthology of mostly anarcho-capitalist, mostly twentieth-century writings, though excerpts from classic works by Burke, Molinari, Spooner, Tucker, Lipscomb, and de Cleyre are also included. Some critical commentary is included as well.

21
LIBERALISM

Michael Freeden

Liberalism is a well-established theory and tradition, yet an inexact concept. While ideologists and philosophers of liberalism may each subscribe to what they regard to be clear and uncontentious accounts of their creed, the analysis of liberalism is often split between unpacking the diversity of its forms and advocating an ideal-type single version, whether "comprehensive" or minimal. It is thus instructive to explore liberalism as a label for a set of political and philosophical views and arguments that does not denote a single ideational phenomenon, but may be interpreted as an assembly of family resemblances, of Venn diagrams, or even as an umbrella term for different epistemologies and ideologies. Liberalism embraces ideological and philosophical positions possessing the heuristic and classificatory features that discharge the delivery of necessarily simplified understandings of the world, but it also conceals within it a considerably divergent set of beliefs and political understandings, attesting to the pluralist and polyvalent nature of political thought and to the potential problems of imposing precise categories on the messiness of ideological entities.

Throughout its history, liberalism has been diffused by means of two particular models, each with different globalizing tendencies: the first is a contextual-historical model that has presented liberalism as a narrative about the growth of civilization that spreads not through the power of timeless logic and stipulative morality, but through concrete spatial and temporal quasi-teleological contagion or through cultural contestation and triumph. The other is an analytical-philosophical model that has anchored liberalism in an abstract ethical construction of desirable human conduct with immediate universal rational appeal—from social contract theory through psychological approaches to human flourishing to original positions under a veil of ignorance—that ostensibly detaches it from context and contingency. More recently, a third semantic model approaches liberalism as an intricate field of meaning possessing an identifiable and patterned conceptual morphology. The three models form the core of the following analysis.

With regard to its content, liberalism may, among others, be seen as a humanist and occasionally radical doctrine, a political theory commending individualism and autonomy, or a capitalist rationalization of the interests of the bourgeoisie. Yet liberalism has not fragmented to the extent that has been the recent lot of the radical ideologies of the left and the right. Empirical evidence of liberal discourse demonstrates that its core concepts have been fairly constant over time, but its adjacent and peripheral components differ markedly at different times and spaces, and those differences react back on the internal variations exhibited even by the core concepts. If liberty is a clear etymological contender for core liberal status, its association with economic entrepreneurship,

with individual development, or with self-ownership and private property will propel it onto very different trajectories of meaning, and the evolution of liberalism has seen liberty joined by other values of similar importance. On that understanding, liberalism is a concatenation of core concepts of which liberty is only one, and each of which may appear in different relative weightings. Notably, J.S. Mill referred to a mutually sustaining cluster consisting of three concepts: the "free development of individuality" (Mill 1977: 261).

1. The Evolution of Liberalism in Temporal Contexts

Historically, the origins of liberal thinking precede its naming, the source of which was in early nineteenth-century Spain. The centrality of liberty in proto-liberal argument identifies its emancipatory purpose: the release of individuals from oppression or at least political invisibility. For those pre-liberal trends that emancipation was initially founded on a doctrine of natural rights—exemplified in John Locke's writings (Locke 1963)—that were the inalienable property of the individual agent but were inadequately protected without governmental safeguarding. The anthropocentrism contained in that viewpoint attributed reason to individuals and implied an interpersonal equality. For earlier liberals, equality was limited to that shared rationality alone, though over the past century, as we shall see, liberals have expanded their understandings of equality. Those aspects coincided with a desire to articulate legitimate constraints on the mounting power of governments and states: hence the early-modern call to curb absolute rule first manifested in the right to resist tyranny while institutionalizing respect both for persons and for the groups (religious to begin with, then secular) of which they were part.

Temporally, a series of liberalisms advanced on four paths as its more complex forms evolved: constitutionalism, markets, welfare, and cultural decentralization. Analytically, liberalism contains broad divisions into universal versus plural; absolute versus relative; horizontal versus vertical; and static versus dynamic. (See section 2 below.)

The development of liberalism entailed not only a move from one stage to the next, but an accumulative process resulting in parallel and layered nuances. As an ideology, liberalism aspired to institutional reification in its parliamentary struggle for a political voice through Liberal parties—deriving from seventeenth-century theories of consent as the basis of legitimate government—but it also sought expression in the economic structures of civil society and, later on, in the redistributive mechanisms of human well-being. The franchise and the legal limits on power were slowly put into practice during liberalism's political dominance in the nineteenth and early twentieth centuries, while a number of liberal ideas occasionally developed, as in France, *sans le nom*, late nineteenth-century solidarism being such an instance (Logue 1983). Legally, a philosophical anthropology provided a narrative in which an initial and inevitable, yet voluntary, social contract became the driving force that transformed potential social chaos into the collective harmony so central to the early liberal imagination; and the concept of contract served to inculcate the notion of mutual promising as a binding constraint on the conditional conduct of individuals and governments alike. Those ideas were reflected among others in the U.S. Constitution, in which the rule of law and a Bill of Rights offered a measure of representation, equal treatment and impartiality towards the citizens of the state and undertook to do away with the arbitrariness of government, adding to the predictability and respect that liberals sought to establish in social relations. The universalist rhetoric that sustained liberal thinking for much of its exist-

ence was thus reinforced: individual agents were the unit of analysis but the quest for their identical formal and legal regard could also efface their individuality, particularly through the rise of the bureaucracies on which the rule of law relied, as evinced in German theories of the *Rechtsstaat* (Beetham 1989).

Liberalism's association with markets was partly grounded on its pursuit of free trade as another manifestation of the unfettering of individual activity. In alliance with imperialism, liberals curiously intermingled the colonizing of foreign markets with a sense of mission concerning the spread of its values across the globe. Within societies, liberals extolled the spirit of entrepreneurship, the role of voluntary associations, and the sacrosanct status of individuals. These were buttressed by the independence and security that the ownership of private property was thought to confer, and by the rationality of a market that converted selfishness into benefits for others—through what Adam Smith popularized as the invisible hand and Hegel understood as the necessary but unintended interdependence of the members of civil society (Smith 2008; Hegel 1991). In Britain, the market element of liberalism was also significantly bolstered by utilitarian theories that attempted to offer a science of human conduct and motivation linked to the pursuit by individuals of their greatest happiness (Bentham 1960). That provided a new cutting edge to the radical urge within liberalism, abetted by an open-ended horizon of social improvement, but flawed—in its threat to the individuality and well-being of some—by an aggregative overriding of the personal preferences of the few by the greatest number. The conflation of many of those streams proved a powerful basis for the ideological hold of liberalism in its terrain of origin, Europe (de Ruggiero 1959).

The thinker who most successfully articulated and popularized an advanced theory of liberalism—a term he rarely used—was John Stuart Mill. What was all along a bourgeois view of social life both socially and economically, particularly in countries such as Germany, was transformed by Mill into a powerful ethical and political theory. For one, Mill exchanged the structural notion of individualism for the cultural notion of individuality, moving away from an atomistic view of the spatial separation of the fundamental social unit to a developmental view of human progress and individual flourishing, in part in conjunction with others, and for which liberty was both a means and an end. The hallmarks of mid-nineteenth-century liberalism that have been retained in salient strands of that tradition thus include a clear divide between the public and private spheres and a dynamic of change. Alongside those appeared a growing recognition of human difference and the value of eccentricity as engines of social progress. A liberal universalism was emerging that "paradoxically" emphasized the uniqueness of each individual: a tension that was to increase in liberalism's future trajectory with the recognition of groups as significant political entities.

Mill's famous, but highly controversial, distinction between self- and other-regarding acts was meant to protect individuals from deliberate harm by others through safeguarding a personal space around them. The imposition of physical and legal constraints on individual action outside a self-regarding domain was deemed sufficient for a just society (Mill 1977: 225). On more recent libertarian interpretations, that has been interpreted as a negative liberty position endorsing a large area of unlimited personal freedom beyond social control, irrespective of many of the incurred social costs (Nozick 1974; Hayek 1960). By encouraging liberty to take up almost all of liberalism's available ideational space, libertarians arguably distort and diminish the greater complexity of liberalism's conceptual morphology. Yet Mill did not entertain a neutral attitude—as do libertarians—towards what individuals could do within their own confines. He

voiced a cultural elitism that was to typify much subsequent liberal argument in favoring individuals who, while expected to pursue their own good in their own way, would recognize that good as incorporating originality, judgment, and the exercise of mental activity and moral preference in choosing their plan of life (Mill 1977: 262). Not least, Mill fortified the perception of liberalism as the epitome of modern civilization, toward which decent societies would eventually evolve.

Mill also expanded the reach of utilitarianism—while undermining much of its rationale—by reformulating it to promote "the permanent interests of a man as a progressive being" (Mill 1977: 224). That set the scene for the third temporal liberal component, welfare—with Benthamite pleasure metamorphosing into happiness as well-being in Mill's hands, before undergoing further mutation into a strong notion of welfare as human flourishing. By the late nineteenth century liberalism had moved decisively into that third stage, assimilating those rising social theories that reassessed human nature as significantly reliant on the co-operation of other human beings; and, in the case of the British new liberalism of L.T. Hobhouse and J.A. Hobson, crafted an organic view of society in which the whole was dependent on the health of each and every part and in which reciprocal assistance was necessary for both individuals and society to thrive. The incipient sociability of liberalism, already discernible in Locke and Mill, now became the basis for a far more powerful notion of liberal community (Freeden 1978).

The new liberal view of human interdependence coincided with a reinterpretation of the barriers to human liberty. A novel identification of constraints marked poverty, ill-health, ignorance, and unemployment as significant impediments to human flourishing and over which individuals had insufficient control. Mutual aid, alongside mutual forbearance, was therefore essential in order to unlock human potential and to tap into the vitality that Mill had sought to exercise and draw out; and here liberals departed from a traditional hostility to the state as oppressor (Hobhouse 1911). Instead, a benevolent state, under democratic regulation, would assist in minimizing those constraints, fighting dehumanizing deprivation with compulsion if necessary and ensuring a proper redistribution of the material and organizational prerequisites for a decent life (Hobson 1909). The habitual embarrassment of liberals concerning the wielding of power was suspended in the area of social reform. The rationale of that incipient welfare state—with the state now legitimately and necessarily moving into new areas of responsibility towards its members, a development that arguably constituted the most significant domestic creation of twentieth-century liberalism—remained true to liberal basics as an enabling instrument, not an enforcing monolith, even if it contained a degree of paternalism. Significantly, the state was an agent of direct social needs as well as a servant of individuals—the new liberals allowed for a layered and integrated conception of human good as both personal and social, reducing—though not eliminating—the older and sharper distinction between public and private. All that was underpinned by a changing conception of human nature. Whereas both the market version of liberalism and the elitist proclivities of Mill saw individuals as purposive, active and energetic self-exercisers (a view that has, to a large extent, been retained by current liberal philosophers), the welfare version regarded them also as vulnerable, fragile and prey to unavoidable risks that were not of their making but reflected serious defects in social organization (Freeden 2003).

The relative success of liberalism's staying power—its ability to adapt to changing circumstances without abandoning its core precepts—and its universalizing tendencies, have also created a series of difficulties in turn. To universalize an ideology may be a

thought-experiment with little political effect, or a practical desideratum in ultimately successful competition with other ideologies. Neither has happened, and the fourth phase of liberalism is thus witness to a dual de-universalization. On one dimension the global aspirations of liberals have failed to take account of the reception and translation of ideological imports into other cultures, and even of the diverse consumption of an ideology within one society. In many senses, then, the family of liberalisms has remained a European and North American set of beliefs. The kind of political language that it evokes does not dovetail into the types of political languages that exist in other societies, or even into different languages in societies where liberalism does exist. It calls for a high educational standard and has therefore flourished only in societies that possess such standards. Whereas varieties of conservatism have no such requirement, and whereas varieties of socialism have a strong mobilizing and radicalizing edge with mass appeal, liberalism is not a populist discourse and cannot be delivered in easy slogans or sound bites.

The second, and theoretically challenging, de-universalization of liberalism has been its growing recognition of pluralism. Initially, that included only the acceptance of varieties of group conduct and thinking that could be contained within a liberal framework (Stears 2002). Increasingly, however, liberals have come to accept co-existence with non-liberal creeds as part and parcel of a liberal outlook. After World War II, they were accused of unwarranted flexibility in permitting the rise of non-liberal ideologies, whose path was smoothed by the reluctance of liberals to apply illiberal methods to secure their eradication. Isaiah Berlin opposed a deplorable and totalitarian monism to a desirable liberal pluralism, but for him that reflected the incommensurability of values that individuals faced in making choices—an alternative version of the essential contestability of concepts (Gallie 1955–6). Hence Berlin intimated that, unlike the structural harmony posited by many earlier liberal versions such as the invisible hand doctrine, or the substantive harmony that progressive liberals such as Hobhouse desired, value-conflict was not only unavoidable but recognized the richness of human thought and the options facing agentic choice (Berlin 1969, 1990). Yet Berlin's pluralism still accorded pride of place to the rationality of choosing, and the ranking of choices, as against radical relativism. More generally, pluralism differs from toleration in that the latter is a value that is external to whatever values are tolerated, which are, in a fundamental sense, a matter of indifference to the tolerator. That position is different from the one that celebrates the internal diversity of moral and political principles as a hallmark of liberalism itself.

But there is also a role for emotion in advocating and supporting liberal ideas and it is mistaken to regard liberalism as an emotion-free zone. Three kinds of emotion relate to liberalism:

1 Expressivism, spontaneity, and human fellow-feeling are identified as crucial features of human conduct.
2 Liberals are passionate when addressing liberalism and are emotionally attached to their principles, as indeed are the proponents of many ideological positions.
3 Liberals have come to recognize the rise of social theories in which emotion plays a central analytical role.

That further underscores the fact that liberalism is not a neutral set of beliefs but a series of deep ideological commitments—both rational and emotional—to issues such as individual liberty and human rights (Freeden 2005: 19–37).

There is, nonetheless, a completely different significance to pluralism, relying on the multifaceted structure of society as composed of various groups. Here is the kernel of another important shift within the family of liberalisms, for when Hobhouse and Hobson acknowledged the political and ethical importance of the group-cum-community, they conceived of it as possessing the singularity of an organic whole. Twentieth-century pluralism challenged that holism and offered instead a far more fragmented view of social structure. Groups and communities replaced the idea of community as comprising a single social body. In the late twentieth century, liberal theorists of a more philosophical bent reformulated that multiplicity as the "politics of difference" (Young 2000), and associated that move with a broadening of the sphere of human rights to include the protection of the practices, needs and demands of social sub-groups. Liberalism's shift to the particular invokes not only the structural pluralist appreciation that groups have a dynamic, even a personality, irreducible to that of their members, but the additional duty of the state to assist in promoting the diverse cultural features of a society. The variety and eccentricity of individuals extolled by Mill were transferred to larger social units as a consequence of the obvious recognition that cultures are not the products of individuals. Not least, though, the problem of how liberals can grapple with groups in their midst that—however externally compliant they are with liberal constitutional requirements at state level—entertain illiberal views and practices internally, attests to what happens to liberalism once the harmony principle is dismissed, and once the harm principle is joined by the possibility of the cultural overriding of liberal principles inside a liberal society (Kymlicka 1995; Baumeister 2000).

On the theoretical level, then, two consecutive developments took place. The excessive pluralist relativism that allowed the rise of totalitarian ideologies, particularly Nazism, encouraged counter-attempts to tease out the solidity of liberalism in some of its post-World War II philosophical variants, such as the highly influential work of John Rawls with its revived philosophical universalist certainty. That new mainstream philosophical liberalism, of which more below, was nonetheless accompanied in its later years by an alternative philosophical standpoint that endorsed the ethical necessity of cultural difference.

2. The Defining Axes of Liberalism

In recent years, liberals have once again been caught in the vice of their own logic. If liberalisms produce an enticing cocktail of voluntarism, choice, agency, self-development and open-mindedness, that is not to say that there is no robustness in those interlocking creeds. There is even something we might cautiously call—at least structurally—a liberal dogmatism (Hartz 1955). Like any political theory or ideology, liberalism possesses red lines it will not cross and non-negotiable principles over which it will not bargain: an attachment to basic human rights (though the list of such rights may differ within the liberal camp); the primacy of liberty; the sanctity of constitutional arrangements; and, for some, the insistence on respect for redistributive social justice. But concurrently, the toleration principle within liberalism sends out mixed messages about the non-negotiability of some of its principles—different principles concerning market choices or the death penalty may display surprising flexibility within dissimilar local liberalisms.

There are other ways of reworking the liberal story that formulate analytically some of its historical manifestations. One distinction revolves around what may be termed the

horizontal-spatial and vertical-temporal axes of liberalism. For many liberals, the main achievement of their doctrine is a horizontal constitutional one that ascribes rights and duties to individual and social entities, that detaches flashpoints of conflict from each other, and that preserves areas of ostensibly extra-political activity. Central to that liberalism of space and separation are constructs such as the public–private divide, the role of checks and balances in a parliamentary system, or the protection of private property (Starr 2007: 53–82). The state is cast mainly in a series of presences in the role of guardian, harm preventer (internal and external), and boundary policer; and in a series of absences in its silences and putative neutrality beyond such agreed boundaries. On the whole, stasis dominates that perception. For example, the balance between public and private and the safeguarding of different lifestyles is prioritized in the Netherlands over harnessing the state to achieve liberal ends (te Velde 2008).

The vertical axis—with its Millian and Darwinian faith in human progress—has dominated a developmental strand of liberalism that measures its success through promoting a secular gospel of humanity aimed at individual progress and fulfillment, as well as emancipating peoples still at earlier stages of advancement. It rejects the illusion of state neutrality and regards the state as the remover of hindrances not just to action but to growth, the unlocker of individual potential and the facilitator of personal visions. Liberalism here is a temporally oriented creed that elicits the march of civilization on the macro-level, and individual flourishing on the micro-level. On the whole, it is an evolving, dynamic and non-teleological set of principles. In that version, the spatial axis may still serve as a minor theme. That is not always reciprocated in the first version, as growth and development merely remain options that individuals might choose to exercise.

We have far from exhausted our understanding of liberalisms. There is an entirely different angle we might explore: not what the features of liberalism are in some kind of objective, external sense, but what will the various methods of investigating liberalism produce? When we enquire into liberalism, how does our chosen approach determine *what* it is that we then identify as liberalism? One way of relating to that question is whether we are prepared to approach liberalism in a comparative spirit, or whether we wish to deny its comparative nature. That question constitutes another access path to the universality or particularity of liberalism. Even then, though, there is lack of clarity. The assumption that liberalism *expresses* itself as (rather than indisputably *is*) a set of universal claims about the human capacity for rationality, individuality, self-development and above all liberty does not have to succumb to locking each of these concepts into a single meaning, but may yet display a comparative flavor if we accept the essential contestability of concepts and deem each of those values or concepts to range across a spectrum of meanings (Collier *et al.* 2006). Comparison is itself an aspect of a liberal mentality, reducing the dogmatism not of the principles we detect in a theory or ideology, but the dogmatism of the scholar in entertaining an openness about what is being studied. Such comparison can be across cultures or within a culture: empirically, we might find that liberalism is a regional ideology—predominantly European and North American—competing with powerful contenders and adopting local versions of considerable diversity, say Indian liberalism with its prioritization of the protection of minorities; and we might find that within a nation or culture there is considerable disagreement on what constitutes liberalism: social liberalism versus neo-liberalism is one such instance.

3. The Morphology of Liberalism

Another method of exploring liberalism looks at its ideational morphology in order to identify the kind of conceptual configurations that characterize actually existing liberalisms; namely, ideological positions that are named liberal by their adherents and opponents. This then creates a critical mass of overlapping configurations that assists in determining whether additional variants fall within the general configurative pattern. That approach relies on empirical evidence through which core—that is to say, durable and ubiquitous—concepts can be detected. Within the family of liberalisms seven core concepts can claim that recurring status: liberty, individuality, progress, rationality, the general interest, sociability (for which, as noted above, there is surprisingly ample evidence in the writings of liberals), and constrained and accountable power. Each of those, obviously, encompasses a number of different conceptions, so that permutations within liberalism are not only possible but inevitable (Freeden 1996). Thus the absence of constraint implied by liberty will depend on the significance of the perceived barriers to action and development: formal and legal, socio-economic, cultural, ethnic, gender-based and so on. Rationality may entail the capacity for reflective choice, or a calculating economic efficiency, or a harmony among individual wills, or compliance with ethical standards. The generality of the core thus holds liberalisms together in a recognizable family, while its specific decontestations supply the local flavor and the internal diversity. Thus progress may be linked to material gain, to self-realization, or to attaining educational and cultural goals. All these are sited within the semantic field of liberalism. Yet the permutations are not unlimited: the structural flexibility of liberalism's conceptual architecture mirrors its pursuit of tolerance, but its core beliefs limit that flexibility to revolving around a limited number of constraining conceptual nodes and their adjacent concepts. The conjunction of liberty and property carries different ideological meanings from the conjunction of liberty and individual development; while the link between liberty and anarchy may be exiled from the liberal family altogether.

The crucial issues in the morphological analysis of liberalism are therefore triple: First, what are the semantic boundaries of its core concepts that liberals can put up with, before their creed mutates into something else? What happens when there is a shift from a thick liberalism to its thinning to the point where it has insufficient family resemblances to remain a member? Second, it is not just the presence or absence of a core concept that determines the profile of liberalism. The question is, rather: what is the relative weight of each core concept within the liberal conceptual cluster—that is, the distribution of significance a particular liberal variant deploys? When a single core liberal concept expands to take up all available space, the delicate and complex internal balance of liberalism is demolished. Thus, if we run with markets alone—with the privatization of exchange relationships or the establishing of a vested sphere of civil society—at the expense of individual development or liberty as reflective choice, the claim for membership of the liberal family begins to falter. Third, how do we cope with the claim by non-liberal ideologies to possess strong liberal features? What happens when liberalism is misappropriated, as has been the case in recent years when radical-right parties in the Netherlands or in Austria have promoted what they claimed was a liberating protection of certain rights of minorities, or a populism that professed to represent the actual free will of the people? Here we have to exercise the kind of scholarly judgment that assesses "liberal" elements in relation to the fuller conceptual vocabulary

of the ideology in question. The presence of liberty or of some human rights could be an intentional deflector for the other messages of that ideology, or it might be an orphaned fragment that has local significance. No less telling is the presence of potentially illiberal components within various liberalisms (King 1999). It has, for instance, exhibited paternalist tendencies, appealing in real politics to elite expertise and education, and appealing in theoretical terms to reason and truth outside the expression of democratic will, as in Ronald Dworkin's rejection of majoritarianism with regard to a Bill of Rights (Dworkin 1985: 197), or Rawls's insistence that it should be possible to establish liberal principles that meet the urgent political requirement to fix, once and for all, the content of certain basic rights and liberties and thus to take them off the political agenda (Rawls 1996: 161).

It is sometimes claimed that liberalism, in an inchoate form, is the winning ideology. Ideologies do not win unless there is a finishing post—which is inconceivable if an open-ended notion of time is adopted, nor would it be possible to ascertain which liberalism has won, given that there are so many liberal runners. But there is a deeper side to that argument, namely, that liberalism has permeated many other neighboring ideologies: the respect for constitutionalism and rights that conservatives may display or the emphasis on human development and choice that some socialisms exhibit. Much European social-democracy saw itself as the heir to and extender of liberal principles, as Eduard Bernstein's writings reveal (Bernstein 1993). On the whole, though, relatively thin and fragmented components of liberalism have been assimilated and appropriated into other ideologies and political discourses.

4. Recent Philosophical Liberalism

Liberalism as a political philosophy has seen a strong revival in recent decades, spurred on by the work of Rawls. For liberal political philosophers, liberalism is a method—often with irresistible rational and moral appeal—of thinking about the principles of social life and the attainment of just social outcomes. Many analytical and ethical liberal philosophers tend to emphasize the mental and moral capacities of human beings, focusing as they do on what is *uniquely* human, at the expense of physical and emotional attributes and in contrast to the discourses surrounding welfare-state liberalism. Positioned at the center of their liberal ethical theories is an autonomous, rational, purposive, individual agent. One by-product of that is the identification of free choice with free reasonable choice, so that Rawls could assert of the compelling nature of political philosophy that "if we feel coerced, it may be because, when we reflect on the matter at hand, values, principles, and standards are so formulated and arranged that they are freely recognized as ones we do, or should, accept" (Rawls 1996: 45). That differs from the non-negotiability that liberalism displays on key principles, for the latter insistence on the irrevocability of certain values for the good life is based on criteria that are not the result of abstract individual reflective reasoning but emanate from the common values that civilized societies have developed over time. By contrast to Rawls, Mill expressed a *preference* for choosing self-development and the exercising of individual capabilities but permitted individuals to fall short of such requirements. That more permissive liberalism was aptly summed up by the one-time British Prime Minister, Herbert Asquith, reacting to T.H. Green's perfectionist view of human development:

> I believe in the right of every man face to face with the State to make the best of himself and subject to the limitation that he does not become a nuisance or a danger to the community to make less than the best of himself.
>
> (Asquith 1907: 1189)

Liberalism can bear a modicum of individual imperfections and systemic tentativeness.

In many of its philosophical versions liberalism can be seen as a modeling device in which universal ground rules for a society are drawn up, one that permits in particular a fair and equal pursuit of the chosen life plan of every person: liberty as reflective choice; inclusiveness; impartiality; and respecting the individual pursuit of basic individual goods or needs (Raz 1986), team up as sustaining values. In recent decades ethics has become largely coterminous with that philosophical liberalism. Liberalism—in the singular—is an investigation into the agreed normative requirements of a free and well-ordered community.

In Rawlsian versions, justice is the first virtue of such a society, even if liberty has lexical priority within liberal justice. The idea of harmony returns with some force, inasmuch as his political liberalism aims at an overlapping consensus over a core of fundamental principles concerning equal liberty and a redistributive system that compensates the worse off for social inequalities, "serious contention about which must undermine the bases of social co-operation" (Rawls 1996: 157). That, however, involves the shrinkage of the historical domain of liberalism to what Rawls terms its political features—on a narrow interpretation of the political as a set of constitutional arrangements, presented as a theory of the right rather than the good, that can command public agreement whether or not the comprehensive doctrines they support are liberal. On closer inspection that theory, too, includes distinct values of well-being incorporated into a public culture, whereas its claim to neutrality is both not part of the historic liberal tradition and itself a powerful value (Galston 1991; Rosenblum 1989). Rawls's methodological optimism, termed by him a realistic utopia (Rawls 1999: 11), is not one uncovered by the actual study of political thinking that acknowledges its indeterminacy and vagueness. Rawls's regulative ideals are too inconclusive to offer concrete guidelines for policies and activities that collectivities can undertake. They eschew the constraints of the political world, constraints that cannot be removed through ratiocination or searching for optimal virtue.

Rawls's political liberalism was presented as a Western paradigm, but one whose *rightness* would gain it universal recognition, moving from a historicity anchored in a democratic culture to ahistoricity. Its universal appeal would not be forged out of competition over legitimacy with other ideological arrangements, as most liberals have measured their successes, earning their spurs on the battlefield of ideas, but as a non-pluralist, free-standing procedural and moral necessity grounded in human reasonableness. Put differently, this liberalism employed an argument not over substantive goods, but over justifiable rules. In earlier versions of liberalism, acceptance had to be won through the substantive benefits that would accrue from the liberal way of life in struggle with other political options. Rawls assumed coherence rather than tentativeness. "We may," he writes, "reaffirm our more particular judgments and decide instead to modify the proposed conception of justice until judgments at all levels of generality are at last in line on due reflection" (Rawls 1996: 45). Contrast this with another mid-century appraisal of liberalism: "as a form of social thought Liberalism has been empirical, scientific, mundane, and more or less skeptical-minded" (Neill 1953: 18). The sense of paradigmatic shift is palpable.

One pre-eminent consequence of this change has been the drastic contraction of the natural habitat of liberalism, the political arena. The bulk of the philosophical liberal conceptual configuration now spills over into a politics-free area, if politics is understood as including contention as well as co-operation. Individuals are assigned heavy personal responsibility in the names of autonomy and self-determination outside the social networks that many liberals identify as vital to self-realization. The capacity of liberalism for gathering the fruits of the political process, hitherto a central aim of the liberal agenda, has been weakened (Geuss 2008). It has challenged the delicate balance between essence and contingency so typical of liberal argument. And it has pared down the liberal tradition by relegating its comprehensive values to the status of equal contenders with other socio-political doctrines, thus undercutting the specific ideological appeal of liberalism, which its other proponents have been so eager to advance. Paradoxically, though, Rawls's political liberalism is so imbued with liberal assumptions and principles that it is difficult to see it as the basis of anything other than a broader liberal creed. All that is not to underplay the great contributions made to liberalism by recent political philosophers in an attempt to test an optimal structure of liberal principles, to envision a world in which radical liberal principles govern the management of social justice, and in which transparency, reasonableness, justification and accountability lend dignity to collective life (Gaus 2003). The role of such normative experiments is to stretch the liberal imagination, explore its consistency, and adumbrate its ethical content, so that any of liberalism's concrete instances can be judged by its optimal promise, and it is in that intellectual exercise that recent philosophical liberals have excelled.

It would be inaccurate to suggest that there is a liberalism "out there" that simply has to be discovered, tracked, developed or criticized. Our comprehension of liberalism will significantly depend on the method employed to access and explore it: analytical-philosophical; contextual-historical; or semantic-ideological. Though these methods reveal important overlaps in substantive liberal beliefs, they also cast light on very different potentials of liberalism and, not infrequently, pull it in diverse directions. That diversity itself attests to liberalism's remarkable, pervasive and durable power.

Related Topics

Mill, Late Nineteenth- and Early Twentieth-century British Thought, Contractualism and Political Liberalism, Pluralism, Libertarianism, Freedom, Rights

References

Asquith, H.H. *Hansard*, 4th Series (April 18, 1907), London.

Baumeister, A.T. (2000) *Liberalism and the "Politics of Difference,"* Edinburgh: Edinburgh University Press.

Beetham, D. (1989) "Max Weber and the Liberal Political Tradition," *European Journal of Sociology* 30: 311–323.

Bentham, J. (1960) *A Fragment on Government* and *An Introduction to the Principles of Morals and Legislation*, Oxford: Basil Blackwell.

Berlin, I. (1969) *Four Essays on Liberty*, Oxford: Oxford University Press.

—— (1990) *The Crooked Timber of Humanity*, London: John Murray.

Bernstein, E. (1993) *The Preconditions of Socialism*, Cambridge: Cambridge University Press.

Collier, D., Hidalgo, F.D. and Maciuceanu, A.O. (2006) "Essentially Contested Concepts: Debates and Applications," *Journal of Political Ideologies* 11: 211–246.

de Ruggiero, G. (1959) *The History of European Liberalism*, Boston, MA: Beacon Press.

Dworkin, R. (1985) *A Matter of Principle*, Oxford: Clarendon Press.

Freeden, M. (1978) *The New Liberalism: An Ideology of Social Reform*, Oxford: Clarendon Press.
—— (1996) *Ideologies and Political Theory: A Conceptual Approach*, Oxford: Clarendon Press.
—— (2003) "The Coming of the Welfare State," in T. Ball and R. Bellamy (eds.), *The Cambridge History of Twentieth Century Political Thought*, Cambridge: Cambridge University Press.
—— (2005) *Liberal Languages: Ideological Imaginations and Twentieth-Century Progressive Thought*, Princeton, NJ: Princeton University Press.
Gallie, W.B. (1955–6) "Essentially Contested Concepts," *Proceedings of the Aristotelian Society* 56: 167–198.
Galston, W. (1991) *Liberal Purposes*, Cambridge: Cambridge University Press.
Gaus, G.F. (2003) *Contemporary Theories of Liberalism*, London: Sage.
Geuss, R. (2008) *Philosophy and Real Politics*, Princeton, NJ: Princeton University Press.
Hartz, L. (1955) *The Liberal Tradition in America*, New York: Harcourt, Brace & World, Inc.
Hayek, F.A. (1960) *The Constitution of Liberty*, London: Routledge.
Hegel, G.W.F. (1991) *Elements of the Philosophy of Right*, Cambridge: Cambridge University Press.
Hobhouse, L.T (1911) *Liberalism*, London: Williams and Norgate.
Hobson, J.A. (1909) *The Crisis of Liberalism*, London: P.S. King & Son.
King, D. (1999) *In the Name of Liberalism: Illiberal Social Policy in the United States and Britain*, New York: Oxford University Press.
Kymlicka, W. (1995) *Multicultural Citizenship: A Liberal Theory of Minority Rights*, Oxford: Clarendon Press.
Locke, J. (1963) *Two Treatises of Government*, Cambridge: Cambridge University Press.
Logue, W. (1983) *From Philosophy to Sociology: The Evolution of French Liberalism 1870–1914*, DeKalb, IL: Northern Illinois University Press.
Mill, J.S. (1977) "On Liberty," in J.M. Robson (ed.), *John Stuart Mill: Essays on Politics and Society* 18, Toronto and Buffalo: University of Toronto Press/Routledge & Kegan Paul.
Neill, T.P. (1953) *The Rise and Decline of Liberalism*, Milwaukee: Bruce Publishing Company.
Nozick, R. (1974) *Anarchy, State, and Utopia*, Oxford: Basil Blackwell.
Rawls, J. (1996) *Political Liberalism*, New York: Columbia University Press.
—— (1999) *The Law of Peoples*, Cambridge, MA: Harvard University Press.
Raz, J. (1986) *The Morality of Freedom*, Oxford: Clarendon Press.
Rosenblum, N.L., ed. (1989) *Liberalism and the Moral Life*, Cambridge, MA: Harvard University Press.
Smith, A. (2008) *The Wealth of Nations*, Radford, VA: Wilder Publications.
Starr, P. (2007) *Freedom's Power: The True Force of Liberalism*, New York: Basic Books.
Stears, M. (2002) *Progressives, Pluralists, and the Problems of the State*, Oxford: Oxford University Press.
te Velde, H. (2008) "The Organization of Liberty: Dutch Liberalism as a Case of the History of European Liberalism," *European Journal of Political Theory* 7: 65–79.
Young, I.M. (2000) *Inclusion and Democracy*, Oxford: Oxford University Press.

22
CONSERVATISM

John Kekes

1. Overview

It has been said, and said, that conservatism came into being as a result of Burke's reaction to the French Revolution (Burke 1999/1790). This is one of the many misunderstandings from which conservatism has long suffered. Conservatives aim to preserve political arrangements that are worth preserving. And surely it did not require the French Revolution to make reasonable people want to protect whatever they have found good against foolish attempts to abandon it and experiment with something new and untried. Conservatism is the political expression of this elementary requirement of prudence, not a mindless attachment to the past. It has the reasonable aim to preserve political arrangements that have stood the test of time, provided they are supported by those who live under them and enable them to make of their lives what they can. I will proceed by identifying several features of what I take to be the strongest version of conservatism (Kekes 1998). These features jointly define it and distinguish it from weaker versions, as well as from other political outlooks.

A common ground among conservatives is that history is the best guide to what political arrangements ought to be preserved. Conservatives do not think that the reasons for or against political arrangements should be derived from a hypothetical contract, an ideal theory, the supposed common good, or a sacred or secular book. In preference to these alternatives, conservatives look to the history of their society in order to understand what deserves their allegiance and what is inimical to well-being. The preservation of political arrangements, of course, depends on continually adapting them to changing circumstances. Conservatism, therefore, does not involve rigid adherence to any political arrangement, but a flexible judgment about what political arrangements should be preserved and how they should be changed to cope with changing circumstances.

Preserving political arrangements is like preserving one's house. It requires constant repair, refurbishment, additions if circumstances warrant them, anticipating problems and coping with them if they occur unexpectedly, being on good terms with neighbors, having trustworthy people to do the upkeep, and generally making and keeping it a comfortable place conducive to living as one wishes. But throughout all the necessary changes it remains the house that, for better or worse, one lives in. The reason for taking pains with it is to make living in it better.

All this is common ground among conservatives, and even among some of their opponents. More needs to be said, therefore, to explain the disagreements among conservatives and between conservatives and others. The explanation that follows focuses on

the different kinds of reasons that may be given to decide what political arrangements ought to be preserved. The nature of these reasons emerges from four distinctions.

2. Tradition vs. Authority

The first is between authoritarian and traditionalist conservatives. Authoritarians think that if the considered judgments of a long-established social authority conflict with the judgments of individuals, then the first should override the second (Maistre 1965/1821). The problem with this view is that, although individuals are fallible about everything, including their own well-being, they are still likely to know better than any social authority, who is also fallible, what does and what does not foster their well-being. It is not enough if a social authority declares that something is a condition of individual well-being. Individuals must actually find that the condition fosters their well-being. The judgments of individuals, of course, can be influenced by the judgments of a social authority, but no matter how strong that influence is, the judgment must ultimately be made by the individuals whose well-being is at stake. If their judgments are mistaken, as they may be, they must bear the consequences. As the lamentable historical record shows, however, this has not been accepted by countless social authorities who persecuted those who rejected their judgments as heretics, infidels, class enemies, maladjusted, living with false consciousness, in bad faith, or in a state of sin. The result is a repressive society whose dogmatism is reinforced by moralizing that unjustifiably assumes that human well-being can take only one form and the authorities know what that form is.

Traditionalist conservatives, such as Montaigne and Oakeshott, argue against authoritarians that there is no need to insist that either individual or authoritarian judgment should systematically prevail over the other when they conflict. Both are important for human well-being. Instead of engaging in futile arguments about their comparative importance, it is far more illuminating to recognize that they are parts of two interdependent and indispensable aspects of the same underlying activity: the pursuit of well-being by individuals. Its two aspects are the individual and the social, and the connecting link between them is tradition (Casey 1978; Oakeshott 1962 and 1975; Shils 1981).

A tradition is a set of customary beliefs, practices, and actions that has endured from the past to the present, attracted the allegiance of people, and made them wish to continue to participate in it voluntarily. A tradition can be reflective and rule-governed, like chess tournaments, or unreflective and spontaneous, like sports fans rooting for their teams; it might have a formal institutional framework, like the Catholic Church, or it could be unstructured, like mountain-climbing; it could be competitive, like the Olympics; largely passive, like going to the opera; humanitarian, like the Red Cross; self-directed, like jogging; honorific, like the Nobel Prize; or punitive, like criminal proceedings. Traditions can be religious, horticultural, scientific, athletic, political, stylistic, moral, financial, aesthetic, commercial, medical, legal, military, educational, architectural, and so on and on. They permeate human lives.

When individuals gradually and experimentally form their idea of well-being, what they are mainly doing is deciding on what traditions they want to participate in. Their decisions might be conscious, deliberate, clear-cut yes-or-no choices, or unconscious, unreflective ways falling in with familiar patterns, or they could be at various points in between. But however they decide, the bulk of their activities consist in participating in the various traditions of their society.

Individuals certainly make decisions, but they do so in the frameworks of various traditions, which provide the alternatives and the standards that determine what decisions are good or bad, reasonable or unreasonable. Their decisions are the individual aspects of their conformity to their tradition's authority. They make decisions *by* following the standards of the traditions to which they feel allegiance. When a Catholic goes to confession, a violinist gives a concert, a football player scores a touchdown, a student graduates, a judge sentences a criminal, then the individual and the social are inextricably mixed. To understand what is going on in terms of individual decisions is as one-sided as it is to do so in terms of social authority. Both play an essential role, and understanding what is going on requires understanding both the roles they play and what makes them essential. Traditionalism rests on this understanding, and it is a political response to it. The response is to maintain political arrangements that protect those traditions that foster the well-being of individuals.

Traditions are not independent of one another. They overlap and problems or questions occurring in one are often resolved in terms of another. Most traditions have legal, moral, political, aesthetic, stylistic, commercial, and a multitude of other aspects. Furthermore, people participating in one tradition necessarily bring with them the beliefs, values, and customs of the other traditions in which they also participate. Changes in one tradition, therefore, are most likely to produce changes in others, because they are like waves that reverberate throughout the other traditions of a society.

Some of these changes are for the better, others for the worse. Most of them are complex, have consequences that grow more unpredictable the more distant they are, and thus tend to escape human control. Since these changes involve the traditions on which the well-being of the participants depends, traditionalist conservatives will be extremely cautious about changing political arrangements that protect such time-tested traditions. They will want to control the changes in so far as it is possible, to make them no greater than what is necessary for overcoming some specific obstacle to well-being. They will be reluctant to experiment or to support large changes.

Changes, of course, are often necessary. Traditions can be vicious, destructive, stultifying, nay-saying, and thus not conducive to well-being. Human well-being requires drawing distinctions among traditions that are unacceptable, tolerable, and not just tolerated but cherished. Traditions that violate the minimum requirements of well-being, such as vendettas, should be prohibited; those that have shown themselves to make questionable contributions to well-being, such as pornography, could be tolerated but not encouraged, and those whose record testifies to their importance for well-being, such as philanthropy, should be cherished.

The obvious question is *who* should decide which tradition is which and *how* that decision should be made. According to traditionalist conservatives, the decision should be made by those who are empowered to do so by the political arrangements of their society and they should make the decisions by reflecting on the historical record of the traditions in question. Competent decision-makers, therefore, ought to know and understand the historical record. Political arrangements work well if they empower decision-makers who are not ill-educated, passionate about a single issue, inexperienced, or have qualifications that lie in some other field of endeavor. Traditionalist conservatives, therefore, will not favor populist politics.

A society that proceeds in this manner is tolerant, because it is committed to having as many traditions as possible. Traditionalist conservatives, therefore, will place the burden of proof on those who wish to proscribe a tradition. If a tradition has endured, if

it has attracted the voluntary allegiance of enough people to perpetuate it, then there is a prima facie case for it. The case might be defeated, but the initial presumption is in its favor.

Traditionalist conservatives favor limited government, because, according to them, the purpose of political arrangements is not to create heaven on earth by imposing on people some conception of well-being. The political arrangements of limited government should interfere as little as possible with the indigenous traditions that flourish among people who participate in them. This is not to say that there should be no interference with traditions that are contrary to human well-being, only that the interference should be no more than what is necessary for the preservation of traditions that aid human well-being and to which widespread voluntary allegiance is felt. The political arrangements traditionalist conservatives favor are those that enable people to live as they please, rather than to force them to live in a particular way. One of the most important ways of accomplishing this is to have a wide plurality of traditions as a bulwark between individuals and the government that has power over them.

3. Skepticism vs. Metaphysics

The second distinction is between conservatives who look beyond history for the reasons that determine what political arrangements ought to be preserved and those who do not. To be sure, all conservatives agree that history is the appropriate starting point, but some of them believe that it is not an accident that some political arrangements have historically fostered well-being, while others have hindered it. Prominent among conservatives who believe this are Thomists. They think that there is a metaphysical explanation for the historical success or failure of various political arrangements. They believe that there is a moral order in reality. Political arrangements that conform to this order foster well-being, those that conflict with it are bound to cause misery. Metaphysically inclined conservatives are willing to learn from history, but only because history points beyond itself toward more fundamental considerations (Finnis 1980 and 1983; Voegelin 1954–87). They all agree that these considerations center on a moral order, but they disagree whether the order is a Platonic chain of being at whose pinnacle is the Form of the Good; providential, as it is held, among others, by natural law theorists; a Hegelian unfolding of the dialectic of clashing forces culminating in the final unity of reason and action; or some other possibility.

Such disagreements notwithstanding, conservatives of this persuasion are convinced that the ultimate reasons for or against political arrangements are to be found by understanding the nature and implications of the moral order that permeates reality. Disagreements, whether metaphysical, moral, or political, are due to human fallibility. They believe that there is an absolute and eternal truth about these matters. The problem is finding out what it is, or, if it has already been revealed, finding out how the canonical text ought to be interpreted.

The belief that metaphysics is the clue to what political arrangements foster good lives is held not only by some conservatives, but also by some left-wing radicals, such as Robespierre and Marx, who otherwise disagree with conservatives. They believe that the laws that govern human affairs have been discovered: they are the laws of history, sociology, psychology, sociobiology, or ethology. Their shared view is that human well-being is possible only if the political arrangements of a society reflect the relevant laws. Human misery is a consequence of ignorance or wickedness that leads people to favor

arrangements contrary to the laws. History, as they see it, is the painful story of societies banging their collective heads against the wall. They have found the key, the door is now open, history has reached its final phase, and from here on all manner of things would be well, if only their prescriptions were followed.

The historical record of left or right-wing societies whose political arrangements were inspired by metaphysical schemes is one of unmitigated disaster. They all broke countless eggs and none of them produced an omelet. They imposed their certainties on unwilling or duped people, made their lives miserable, all the while promising great improvements just after the present crisis, which usually turned out to be permanent. If the recent past has a moral achievement, it is the realization that it is morally and politically unacceptable to proceed in this way.

Opposed to these metaphysically inclined conservatives and nonconservatives are skeptical conservatives, such as Montaigne (1943/1588) and Hume (1985/1777) in the past and closer to our time (Oakeshott 1962 and 1975; Hartle 2003; Livingston 1984). They need not deny that there is a moral order in reality, only that reliable knowledge of it can be had. Skeptical conservatives are far more impressed by human fallibility than by the success of efforts to overcome it. They think that the claims that some truths are revealed, some texts are canonical, or that someone knows the eternal verities stand in need of persuasive evidence. Such claims are only as credible as the evidence that supports them. But the available evidence is no less dubious than the claims that it is supposed to support. According to skeptical conservatives, it is far more reasonable to look to the historical record of various political arrangements than to endeavor to justify or criticize them by appealing to metaphysical considerations that are bound to be less reliable than the historical record.

Skepticism, however, does not lead conservatives to deny that it is possible to evaluate political arrangements by adducing reasons for or against them. What they deny is that good reasons must be metaphysical. The skepticism of these conservatives, therefore, is not a global doubt about the possibility and desirability of being reasonable, of basing beliefs on the evidence in support of them, and of making the strength of beliefs commensurate with the strength of the evidence. Their skepticism is about drawing political conclusions from metaphysical premises. They want political arrangements to be rooted in the experiences of the people who live with them. Since these experiences are unavoidably historical, it is to history that skeptical conservatives look for supporting evidence. They will not try to deduce from metaphysical premises which orifices of the body are suitable for sexual pleasure or solve problems by appealing to an a priori scheme formulated in ignorance of the conditions that cause the problems and of the misery the a priori solution may cause. Skeptical conservatives thus avoid the absurdity of basing political arrangements on speculation about what lies beyond experience or of being suspicious of all political arrangements, because of a global distrust of reason.

4. Pluralism vs. Absolutism

The third distinction is between conservatives who are absolutists, such as Finnis (1980, 1983), de Maistre (1965/1821), and Stephen (1993/1873) and conservatives who are pluralists, such as Santayana (1951), Oakeshott (1962, 1975), and Kekes (1993). Absolutists accept that there are many values, but they think that they can be ranked on the basis of the standard set by their relative importance for human well-being. Some think that this standard yields a highest value, such as happiness, duty, virtue, or God's will.

Others think that it provides a canonical principle, like the categorical imperative, the satisfaction of preferences, or the Golden Rule. If a choice needs to be made between different values, then the highest good or the canonical principle will determine which value ought to take precedence. Absolutists, then, give as their reason for preferring some political arrangements over others that the preferred ones are judged to be more important by the standard than the alternatives.

Absolutism often has a metaphysical basis, because the reason often given in favor of the absolutist standard is that it reflects the moral order of reality. This is the inspiration behind the attempts to establish ecclesiastical polities, on the right, and egalitarian, Utopian, or millennial ones, on the left. Nevertheless, the connection between absolutism and metaphysics is not a necessary one. Standards need not be metaphysically sanctioned. If, however, their advocates eschew metaphysics, then they must provide some other reason for regarding the standard they favor as authoritative.

It is a considerable embarrassment to absolutists that the standards they offer are as diverse as the values whose diversity is supposed to be diminished by them. Absolutists acknowledge this, and explain it in terms of human shortcomings that prevent people from recognizing the right standard. The history of religious wars, revolutions, left- and right-wing tyrannies, and the persecution of countless unbelievers, all aiming to rectify human shortcomings, testifies to the dangers inherent in this explanation.

Opposed to absolutism is pluralism. According to pluralists, there is an authoritative standard, but it applies only to *some* values that must be recognized by anyone committed to human well-being. The standard is human nature, and the view can be characterized as a minimal, non-metaphysical form of Aristotelianism. It holds that the realization of some values is necessary for human well-being, but it allows for a wide plurality of values beyond the necessary minimum. To understand human nature sufficiently for the purposes of this standard does not require plumbing the depths of the soul, unraveling the obscure springs of human motivation, or conducting scientific research. It does not call for any metaphysical commitment and it can be held without subscribing to the existence of a natural law. It concentrates on normal people in a commonsensical way. It rests on the obvious claim that human well-being depends on the satisfaction of basic physiological, psychological, and social needs: for nutrition, shelter, and rest; for companionship, self-respect, and the hope for a good or better life; for the division of labor, peace, and predictability in human affairs; and the like. The satisfaction of these needs is a minimum requirement of all human lives, whatever the social context may be in which they are lived. If the political arrangements of a society foster their satisfaction, that is a reason for preserving them; if they hinder their satisfaction, that is a reason against them.

Absolutists go beyond the minimum and claim that their standard applies to most of the values on which human well-being depends. Pluralists think that beyond the minimum there is a plurality of values, a plurality of ways of ranking them, and a plurality of conceptions of well-being embodying these values and rankings. They claim that human nature underdetermines the values required for human well-being. According to pluralists, then, the political arrangements of a society ought to protect the minimum requirements of well-being and ought to foster a plurality of values beyond the minimum.

If pluralism takes a conservative form, it provides a universal and objective reason in favor of those political arrangements of a society that protect the minimum requirements and against those political arrangements that violate them (Kekes 1998; Oakeshott 1962 and 1975). It motivates, gives direction to, and sets the goal of intended

reforms. It makes it possible to draw reasonable comparisons among different societies on the basis of how well or badly they protect the conditions of human well-being. Pluralistic conservatism thus avoids the objection that it sanctions any political arrangement so long as a wide enough consensus supports it. It holds that the best guide to what political arrangements are worth preserving beyond the minimum level is the history of the society, rather than any metaphysical consideration. Thus pluralistic conservatism avoids the dangers of dogmatism and repression that beset absolutism.

5. Pessimism vs. Optimism

One of the safest generalizations is that conservatives tend to be pessimists. In the conservative writings of Montaigne (1943/1588), Hume (1985/1777), and Oakeshott (1962 and 1975), for instance, cheerfulness keeps breaking through, but even then, it does so in spite of their doubts about the possibility of a significant improvement in the human condition. Conservatives take a dim view of progress. They are not so foolish as to deny that great progress has been made in science, technology, medicine, communication, management, education, and so forth, and that they have changed human lives for the better. But they have also changed them for the worse. Progress has both beneficial and harmful consequences. It has enlarged the stock of human possibilities, but new possibilities seldom fail to bring with them new dangers. Progress in physics has resulted in nuclear weapons; in technology, it has led to ever-higher numbers of casualties in armed conflicts; in sexual attitudes, it has brought epidemics and the disintegration of families; and in improved living standards and life expectancy, it is causing the exhaustion of unrenewable resources. Conservatives tend to be pessimistic, because they doubt that more possibilities will make lives on the whole better. Their doubt is based on what they believe are permanent obstacles to the significant overall improvement of human well-being.

Conservatism has been called the politics of imperfection (O'Sullivan 1976; Quinton 1978). The label rightly suggests that conservatives reject the idea of human perfectibility. But it is worse than a bad joke to call world wars, the genocide of numerous peoples, tyrannies, systematic torture, ethnic, tribal, and religious mass murders, and other horrors imperfections. Conservatives are much more impressed by the prevalence of evil than this label implies. They think that its prevalence is a permanent obstacle to human well-being that cannot be significantly altered. Imperfection is a misleading label because it suggests that the imperfection is in human beings. Conservatives certainly think that human beings are responsible for much evil, but to think only that is shallow. The prevalence of evil reflects not just a human propensity, but also contingencies that influence what propensities human beings develop, and thus influence human affairs beyond the possibility of human control. The propensity for evil is itself a manifestation of such deeper and more pervasive contingencies. It operates through genetic inheritance, environmental factors, the confluence of events that places people at certain places at certain times, the crimes, accidents, pieces of fortune and misfortune that happen or do not happen to them, the historical period, society, and family into which they are born, and so forth. And the same contingencies affect all individuals, because those with whom their lives are intertwined in countless ways are as subject to it as they are themselves.

The fourth distinction, then, is between conservatives, such as de Maistre (1965/1821), who are led by their pessimism to the hopeless conviction that human nature is predominantly evil rather than good, and conservatives, such as Montaigne (1943/1588), Hobbes (1962/1651), James Fitzjames Stephen (1993/1873), Bradley (1967/1876), Freud

(1959/1920 and 1961/1930) who are led by pessimism to the realistic view that whether the balance of good and evil propensities and their realization tilts one way or another is a contingent matter over which we have some, but not enough, control.

This needs to be stressed: realistic pessimism does not lead conservatives to think that the human condition is devoid of hope. They are, however, realistic about the limited control a society has over the well-being of its members. Their view is *not* that human beings are evil and their evil propensities are uncontrollable but, rather, that human beings have both good and evil propensities and neither they nor their societies can exercise sufficient control to make the realization of good ones reliably prevail over the realization of evil ones. The right political arrangements will help of course; just as the wrong ones will make matters worse. But even under the best conceivable conditions great many contingencies will remain, and they will place beyond human control much good and evil.

The chief reason for this is that the human efforts to control contingencies are themselves subject to the very contingencies they aim to control. And that is the fundamental reason why conservatives are pessimistic and skeptical about the possibility of making significant progress toward the improvement of human well-being. It is thus that the skepticism and realistic pessimism of conservatives reinforce one another. It does not follow, and realistically pessimist conservatives do not believe, that it is a matter of indifference what the prevailing political arrangements are. Although not even the best political arrangements can guarantee the victory of good over evil, they can still influence how much good and evil there is. Whether that will prove sufficient at a certain time and place is itself a contingent matter insufficiently within human control. The resulting attitude will have a negative and a positive component. The negative one is acceptance of the fact that not even political arrangements that best reflect the requirements of reason and morality can guarantee human well-being. The positive one is to strive nevertheless to make political arrangements as reasonable and moral as possible.

If the criticism or the justification of political arrangements is governed by this conservative attitude, it will result in arrangements that look toward fostering what is taken to be good and hindering what is regarded as evil. The political arrangements conservatives favor include those that protect such familiar political values as democracy, equality, justice, liberty, order, peace, private property, prosperity, rights, the rule of law, security, and so forth. There need be no significant difference between the values of conservatives and the ones liberals or socialists favor. But significant differences still remain between conservative politics and the politics of liberals and socialists.

One of these differences is that conservative politics is genuinely pluralistic, while the main alternative contemporary approaches are not. Liberals and socialists are committed to an approach to politics that deems some few values overriding, such as liberty, or equality, or justice, or rights, or some combination of them. It is their essential claim, the claim that makes them liberals, socialists, or whatever, that when the one or few values they favor over the others conflict with the less favored ones, then the ones they favor should prevail. Conservatives reject this approach. Their commitment is to all the values I have listed above, and their essential claim is that what is important is the preservation of the whole system of values. Its preservation in some circumstances requires favoring a particular value over another, sometimes the reverse. And they hold this to be true for each of the values. Conservatives thus differ from liberals and socialists in refusing to make a commitment to the overridingness of any particular value or small number of values.

Another significant difference between conservative politics and most current alternatives to it is the insistence of conservatives on the importance of political arrangements that hinder evil. This difference is a direct result of the pessimism of conservatives and the largely Enlightenment-inspired optimism of others. Their optimism is revealed by the assumption that the prevalence of evil is the result of bad political arrangements. If people were not poor, oppressed, exploited, discriminated against, and so forth, they optimistically suppose, then they would live reasonably and morally. They assume that the prevalence of evil is due to the political corruption of human beings. Optimists suppose that if political arrangements were good, there would be no corruption. What is needed, therefore, is to reform political arrangements so that they would foster the good. The arrangements that hinder evil are unfortunate and temporary measures needed only until the effects of the good arrangements are generally felt.

Conservatives reject this optimism. They do not think that evil is prevalent merely as a result of bad political arrangements. They think that one reason why so many political arrangements are bad is that those who make them act on their evil propensities. Political arrangements, after all, are made by people and they are bound to reflect the propensities of their makers. Since the propensities and the actions they prompt are formed by contingencies over which human control is insufficient, there is no guarantee whatsoever that political arrangements will be made good. Nor that, if they were made good, they would be sufficient to hinder evil.

Conservatives, therefore, insist on the necessity and importance of political arrangements that hinder evil. They stress moral education, the enforcement of morality, the treatment of people according to their moral merit or demerit, the importance of swift and severe punishment for serious crimes, and so on. They will oppose the prevailing attitudes that lead to agonizing over the criminal and forgetting the victim, to perpetuating the absurd fiction of a fundamental moral equality between habitual evildoers and those who suffer from their actions, and to guaranteeing the same freedom and welfare-rights to good and evil people, and so forth.

This leads to yet another major difference between conservatism, on the one hand, and various forms of liberalism and socialism, on the other. Conservatives are not egalitarians, whereas many of their opponents are. Conservatives are not opposed to political and legal equality, but to the idea that on some basic level all human beings have equal worth. How could good and evil, morally better and worse people have equal worth? What could be the point of the hypocritical pretence that permeates liberal and socialist rhetoric that decent people living decent lives deserve the same concern, respect, and resources as murderers, torturers, terrorists, and other criminals who endanger the conditions of well-being?

Conservatives, therefore, will favor political arrangements that hinder evil. They know that such arrangements have often been abused. Conservatives know and care about the historical record that testifies to the dreadful things that have been done on the many occasions when such arrangements have gone wrong. The remedy, however, cannot be to refuse to make the arrangements, but to learn from history and try hard to make reasonable and moral arrangements and avoid their abuse. Conservatives know that in this respect, as in all others, contingencies will make complete success very unlikely. But that is precisely the reason why political arrangements are necessary for hindering evil. Their realistic pessimism will lead conservatives to face the worst and try to deny scope to it, rather than endeavor to erect the City of Man on a far from quiescent volcano.

6. Summary

Perhaps it will be apparent that the strongest version of conservatism is not committed to any religious view. Theists, agnostics, and atheists could all accept it, just so long as they do not hold that political arrangements should have a metaphysical basis. Nor is this conservatism committed to any economic system. It certainly regards prosperity as one of the important political values and it recognizes that it is a reason in favor of an economic system that it is more likely to lead to prosperity than competing systems. But it is only a reason, not an overriding one. For there might be circumstances in which the protection of other political values, such as peace, security, or justice, is more important than prosperity. This version of conservatism, therefore, has no a priori commitment to any economic system. Moreover, no dogmatic answer follows from it to controversial questions about sexual morality. It is opposed to anything that might violate the minimum requirements of human well-being, but it holds that the justification or the criticism of sexual practices beyond that basic level depends on the justification or the criticism of traditions that people who participate in them have found historically worth preserving. And, I repeat, people may be mistaken in supposing that a tradition of theirs is worth preserving. Traditions can be reasonably criticized on the grounds that they are detrimental to the well-being of those who participate in them, even if they themselves think otherwise.

The version of conservatism I have been defending is *traditionalist*, not absolutist; *skeptical*, not metaphysical; *pluralist*, not absolutist; and *realistic* rather than hopelessly *pessimistic*. Two features distinguish this version of conservatism from liberalism and socialism. One is that conservatives deny that there is a highest value or canonical principle that could guide the evaluation of political arrangements, whereas liberals and socialists believe that there is such a value or principle. The other is that conservatives deny that better political arrangements will eliminate the ambivalence toward reason and morality that is inherent in human nature, whereas liberals and socialists are committed to political arrangements that presuppose the opposite.

Related Topics

Hegel, Pluralism

References

Bradley, F.H. (1967/1876) *Ethical Studies*, Oxford: Clarendon Press.
Burke, E. (1999/1790) *Reflections of the Revolution in France*, Indianapolis: Liberty Fund.
Casey, J. (1978) "Tradition and Authority," in M. Cowling, (ed.) *Conservative Essays*, London: Cassell.
Finnis, J. (1980) *Natural Law and Natural Rights*, Oxford: Clarendon Press.
—— (1983) *Fundamentals of Ethics*, Oxford: Clarendon Press.
Freud, S. (1959/1930) *Beyond the Pleasure Principle*, New York: Bantam.
—— (1961/1930) *Civilization and Its Discontents*, New York: Norton.
Hartle, A. (2003) *Michel de Montaigne*, Cambridge: Cambridge University Press.
Hobbes, T. (1962/1651) *Leviathan*, London: Dent.
Hume, David (1985/1777) *Essays, Moral, Political and Literary*, ed. Eugene F. Miller, Indianapolis: Liberty Fund.
Kekes, J. (1993) *The Morality of Pluralism*, Princeton, NJ: Princeton University Press.
—— (1998) *A Case for Conservatism*, Ithaca, NY: Cornell University Press.
Livingston, D.W. (1984) *Hume's Philosophy of Common Life*, Chicago, IL: University of Chicago Press.
Maistre, J. de. (1965/1821) *Works*, London: Allen & Unwin.

Montaigne, M. de. (1943) *Essays*, Stanford, CA: Stanford University Press.
Oakeshott, M. (1962) *Rationalism in Politics*, London: Methuen.
—— (1975) *Human Conduct*, Oxford: Clarendon Press.
O'Sullivan, N. (1976) *Conservatism*, New York: St. Martin's Press.
Quinton, A. (1978) *The Politics of Imperfection*, London: Faber & Faber.
Santayana, G. (1951) *Dominations and Powers*, New York: Scribner's.
Shils, E. (1981) *Tradition*, Chicago, IL: University of Chicago Press.
Stephen, J.F. (1993/1873) *Liberty, Equality, Fraternity*, Indianapolis: Liberty Fund.
Voegelin, E. (1954–87) *Order in History*, 5 vols. Baton Rouge: Louisiana State University Press.

Further Reading

Allitt, Patrick. (2009) *The Conservatives: Ideas & Personalities Throughout American History*, New Haven, CT: Yale University Press.
Hegel, G.W.F. (1991/1821) *The Philosophy of Right*, Cambridge: Cambridge University Press.
Kekes, J. (1992) *The Examined Life*, University Park: Penn State Press.
—— (2005) *The Roots of Evil*, Ithaca, NY: Cornell University Press.
Machiavelli, N. (1985) *The Prince*, Chicago, IL: University of Chicago Press.

23
REPUBLICANISM

Christian Nadeau

Republicanism is, first and foremost, a theory of political liberty. According to the republican theory, an agent is free if and only if he has a status equal to that of any other citizen; furthermore, this status must effectively protect him from domination by other citizens or by the State, i.e. from arbitrary interference with his choices. The greatest obstacle to liberty, according to the republican view, is not interference but, rather, domination, for the latter implies the individual's being in an unacceptable state of vulnerability vis-à-vis another individual, a group of individuals, or a political organization (Pettit and Braithwaite 1990; Pettit 1997a, b, c). The fundamental difference between domination and interference stems from their logical independence: interference can occur without involving the domination of an individual by another agent; conversely, domination can occur even in the absence of interference, as an individual is considered to be dominated just to the extent that he is compelled, for structural reasons, to obey someone else (Lazzeri 2001).

According to the republican theory, having a political status recognized by everyone effectively protects or "immunizes" an agent against all attempts to dominate him, as this presupposes a legitimate association among all citizens (Pettit 1997a: 108). Thus, far from advocating the isolation of individuals from one another, republicanism values their cooperation, and views social institutions as tools necessary for safeguarding the liberties of all; indeed, since it sees political and legal institutions as necessary conditions for liberty, the republican theory posits an analytic relation between the role of the State and the concept of political liberty, with the latter consisting in non-domination. Liberty is not, as it were, a collateral effect of a society's laws, but is, rather, constituted by the institutions which aim to promote all individuals' rights and obligations in order to protect them from illegitimate interference—just as, in the immune system, antibodies collectively react to the introduction of a foreign body (Pettit 1996, 1997a). Now, these antibodies do not properly *cause* the agent's immunity; rather, they *constitute* this very immunity (Honohan 2009). Using the above analogy, we might say that society's institutions are the instrumental means guaranteeing the person's liberty: the antibodies are the means whereby someone becomes immune. In sum, a free individual is one who has institutional guarantees that shield him from any and all arbitrary interferences that could thwart his choices within the legitimate space of his liberty. Those institutional guarantees must be deeply democratic in order to avoid the domination of the State over the individuals that compose it.

1. The Historiographical Debate

Republicanism is an intellectual tradition concerned with the concept of political liberty. This tradition goes back to Antiquity, and has undergone many variations over the centuries, making it difficult to establish clear logical connections between its contemporary proponents and the models that they draw from this tradition. On the other hand, one cannot separate recent theoretical developments within republicanism from the work of numerous historians of political ideas, first and foremost Quentin Skinner. Indeed, it is largely on the basis of Skinner's work that Philip Pettit came up with the concept of liberty as non-domination, which today lies at the heart of republican political theory.

Historians of political ideas have been studying the multiple traditions linked to the history of republicanism for many years and especially the contrast between neo-Roman republicanism as presented by Quentin Skinner (1998) and what Hans Baron called "civic humanism" (Spitz 1995; Geuna 1998). Since Hans Baron's (1955) seminal work on civic humanism, two major families of political ideas have gradually emerged in connection with one another; while they diverge significantly on several points, both nevertheless emphasize the importance of political liberty as guaranteed by legitimate institutions. Civic humanism can be traced back to a classical conception, originally inspired by Aristotle, of "the good life" as a perfectionist ideal. It is based on the idea that the agent fulfills himself by actively participating in the political life of his community. The citizen's active presence makes it possible to establish adequate relations between the State and the citizens, as well as between the citizens themselves (Arendt 1958; Baron 1955; Moulakis 2008; Pocock 1975). It should be added, however, that this category is currently the subject of numerous disputes among historians (Hankins 2000). According to other authors, especially Quentin Skinner, another political tradition, neo-Roman republicanism, which is related to civic humanism but nevertheless different in several key respects, emerged in the Renaissance—in the work of Machiavelli, in particular (Geuna 2009). While civic humanism looked to the political world of Ancient Greece, this alternative tradition turned instead to the thinkers of Rome, especially Cicero. This rereading of Machiavelli has engendered several studies tracing the main historiographical and theoretical stages of republicanism (Viroli 1988, 1990). The main difference between these two ways of conceiving the ideal relations between citizens stems from the different values that each perspective grants to participation in political life and civic virtues: for civic humanism, they are intrinsically good; for neo-Roman republicanism, they are merely means for attaining liberty.

Civic humanism, thus, gives an essential role to the *vita activa* and to the *vivere civile*, i.e. to a way of life dedicated to civic activities. The distinction between civic humanism and republicanism, however, should be understood neither in the same sense as Benjamin Constant's distinction between the liberty of the Ancients versus that of the Moderns, nor as equivalent to Isaiah Berlin's (1969) distinction between positive and negative conceptions of liberty: that is, while the "liberty of the Ancients" and "positive liberty" fairly aptly characterize the principles defended by the authors associated with civic humanism (such as Leonardo Bruni or Collucio Salutati), neither the "liberty of the Moderns," nor "negative liberty"—both of which underlie standard versions of liberalism—adequately captures the ideal of liberty as non-domination.

In fact, the concept of liberty as non-domination emerges in response to dissatisfaction with the seemingly false dichotomy between two traditional conceptions of political liberty, namely "negative" and "positive" liberty.

For the apologists of positive liberty, on the one hand, the real problem is not so much "how far" power can extend, but rather "to whom" it belongs. According to Berlin, the advocates of positive liberty are not concerned with the question of the limits of power, because it is encompassed by the larger, more pressing question, of determining "who" holds power (Berlin 1969; Carter 2008; Spitz 1995).

The negative conception of liberty, on the other hand, is meant to answer the question of the limits of civil institutions' authority. According to the negative conception of liberty, *any* interference by the State or by another individual in one's choices of action constitutes a constraint on liberty. Liberty, therefore, depends on the possibility of making choices without being subject to real and effective constraints. Many liberal authors, following Isaiah Berlin, see negative liberty as the only possible conception of political liberty compatible with value-pluralism. For Berlin, the concept of negative liberty is to be preferred to the positive, for only the former is compatible with the recognition of the plurality of values and their mutually incommensurable character, without requiring individuals to forsake their own ends in the name of emancipation and autonomy, which everyone would ostensibly obtain by way of their political commitment to the common good.

Republicanism is sometimes characterized, according to the vocabulary of liberalism (e.g. Rawls's terminology), as a "comprehensive political theory," based on an ideal of the good life, which, however, is *prima facie* inconsistent with the axiological pluralism defended by liberalism—but, as even Rawls himself remarks, such a characterization improperly conflates civic humanism with republicanism (Rawls 1993).

Most republican authors refuse to assimilate their conception of political liberty to positive liberty. On the one hand, as a deeply *democratic* theory, republicanism upholds the value-pluralism proper to contemporary societies (Guérard de Latour 2009); on the other hand, it must not, in so doing, simply join the ranks of the defenders of the tradition of negative liberty—or at least not the version championed by Berlin and his followers.

Conceived of as a normative theory, negative liberty calls for the reduction, or even the total elimination, of "interference" by others, whereas, according to the republican interpretation, the concept of political liberty involves the absence of *arbitrary* restraints, but does not imply that citizens ought to be protected from *all* kinds of interference whatsoever. The latter requirement would be misguided, in effect, as the State—supposing it to be legitimate, which is by no means assumed necessarily or for all cases—can legitimately exercise interference as a, or even as the only, form of effective *protection* vis-à-vis the arbitrary power relations among its citizens. This is still a negative form of liberty—in the sense that liberty is conceived as the absence of something—yet on the condition of not reducing all constraint of liberty to interference. Properly republican political liberty is not the result of the elimination of the relations between agents; on the contrary, it takes account of the fact that political union actually fosters the protection of all. Thus, if the State, *qua* democratic authority dedicated to promoting and defending its citizens' interests against the hegemony of others, intervenes to hinder or block a particular action, one ought not to regard this "interference" as an obstacle to liberty per se.

The main criticism that republicanism makes of the concept of negative liberty is that the latter seems to lead to the idea that the best means for putting an end to agents' mutual interference is to definitively separate them from one another. The republican model of virtue, by contrast, emphasizes citizens' mutual obligation to defend each oth-

er's liberty as if it were their own: what we have here is a version of an idea going back to the eighteenth century, namely the "love of equality" (*l'amour de l'égalité*) promulgated by Rousseau and his followers (Spitz 2000)—although this does not entail, for all that, a return to positive liberty.

2. Questions and Debates

The concept of liberty as non-domination has given rise to numerous debates since Philip Pettit's first writings on the subject. Quite a few liberal authors, on the one hand, view this concept as being perfectly compatible with the liberal tradition, yet as not effectively adding anything especially interesting or substantive to the traditional liberal conceptions of political liberty developed since John Stuart Mill. On this view, republicanism amounts at most to an amendment to standard theories of political liberty or to a kind of warning serving to draw theorists' attention to the threats to liberty (instead of focusing exclusively on what directly contravenes it); as such, republicanism could have an impact on how we think about the institutionalization of liberalism, say, but does not change the latter's internal logic in any way. Republican authors, on the other hand, present their project as a bona fide option, distinct from the different variants of liberalism, and also directly confront their model with those of other traditions, such as Marxism or libertarianism.

We must first consider the core concepts of the republican theory, the most important of which is without doubt the concept of domination. If we regard it as the possible yet uncertain effect of social relations, then domination can hardly be seen as a genuine problem; more importantly, there seems to be no way to thereby distinguish it from other forms of coercion. Now, the distinguishing mark of domination is that it implies a structural relation between agents, such that some have the capacity to arbitrarily interfere in the choices of others (Lovett 2010). And these structural relations—rather than contingent social relations that actually determine agents against their will—are the true forms of domination.

The classic example is the relation between the good master and the slave: if domination is assessed according to a given social structure, then it might not be possible to identify the effects of domination in certain cases, for example, if the master's goodness extends so far as to allow his slave to be free in all practical respects—yet a slave nonetheless. But the essential thing is that the structure of their relation matters more than the contingent fact of the master's good or bad disposition, for, regardless of the master's goodness, the relation between him and his slave implies a hierarchical order, such that it is always only the former who decides whether or not the latter shall live decently. Indeed, even if the slave were to outwit his master and to succeed in living well by this means, he would nevertheless remain subject to the former's will, because, given the hierarchical order that binds them, any victory by the slave can never be more than provisional. In the *Marriage of Figaro*, for example, the valet desires Susanna, the Countess's chambermaid, and he must use cunning to conceal her from the eye of the Count, his master; yet only when he is no longer in the latter's service can his victory become definitive.

On the other hand, while domination stems from perverted social structures, these structures themselves are not agents of domination per se; rather, it is *agents* who dominate each other, for without them, those social structures would have no effect whatsoever. In a society characterized by domination, one master may die and be replaced by another,

but if all of the masters were to disappear and only the slaves to remain, the very notion of domination would lose all significance; that is to say, domination implies both dominating and dominated agents. But, as we have just seen, no agents can be so qualified without reference to the structure of the relation that regulates their mutual interactions.

The republican theory has provoked much criticism, particularly of its concept of liberty as non-domination. Most objections consist in denying that this concept adds anything significant to the traditional liberal conceptions of political liberty. Others recognize its distinctiveness, yet deem it either incoherent or unacceptable on the ground that it is, ultimately, too closely tied to a conception of "the good life" and hence incompatible with the pluralism intrinsic to contemporary democratic societies. These two types of criticism are mainly voiced by liberal theorists (Goodin 2003; Larmore 2004; Patten 1996; Waldron 2007).

Another group of critics, while favorable towards the republican project, nevertheless doubt whether it can become truly relevant unless it distinguishes itself more clearly from liberalism. These critics argue in various ways for the need to appeal to a substantive moral conception of the relations between individuals—a strategy that, in extreme cases, can lead to giving up the concept of liberty as non-domination altogether, in favor of positive liberty. Other "internal" critics of republicanism attempt to differentiate it from liberalism by advocating more radical mechanisms for contestation and participation by citizens, without, for all that, regarding this conception as a means for guiding citizens towards a putative ideal of the good life (Maynor 2003).

According to the first group of critics, then, republicanism does not, in fact, offer anything especially new or distinctive compared to the standard liberal positions: at most, it serves to highlight certain kinds of arbitrary, hierarchical relations that the liberal theory may have neglected—yet these could simply be added to the latter's standard list of obstacles to liberty (Goodin 2003; Larmore 2004). Many other authors (e.g. Brennan and Lomasky 2006) are skeptical regarding the pertinence of a strong distinction between liberalism and republicanism (social goods versus communal goods; political participation versus political liberty; etc.). Finally, if other liberal critics endorse the republican theory as an original political project, they nevertheless do not regard it as opposed to liberalism at all, but rather as the best version of the latter (Dagger 2006).

For a second group of critics, republicanism can constitute a new political language, provided that it remains true to its radical heritage, close to both Machiavelli and Marx (Bellamy 2002).

A third group of critics rejects the contemporary formulation of the republican theory, deeming it incompatible with the political virtues endorsed by civic humanism. For without the moral perfectionism proper to civic humanism, one loses any relevant separation between republicanism and liberalism. Consequently, if republicanism is to distinguish itself from the liberal tradition on this score, it must embrace the strong moral values that have emerged over the course of its own history.

Although this line of criticism ends up neglecting liberalism's own perfectionist elements, invoked by numerous authors (Raz 1986; Galston 1991), it nevertheless has the merit of bringing to the fore the question of the role of moral character in the republican theory: Ought one to regard republicanism in the same way that Rawls sees liberalism in this respect, namely as a political theory independent of moral debates, or rather as tied, by definition, to a particular ethical conception?

Pettit and Skinner have always emphasized that republicanism must be clearly distinguished from any and all forms of communitarianism or neo-Aristotelian ideal of civic

virtue that could lead to paternalism, for the latter would not only be anachronistic but could even impinge on individual liberties (Bellamy 2002), whereas the advantage of the republican theory is supposed to be just to protect against all forms of arbitrary interference, be it by individuals (*dominium*) or by the State (*imperium*)—even if the latter intervenes for the sake of an ideal of "the good life." At issue here is the moral perfectionism bound up with the model of civic virtue upheld by communitarianism (Etzioni 1993). By renouncing the primary role ascribed to civic virtues within a political conception of the good life, republicanism, it would seem, would not be compatible with the idea that the public good ought always to be considered superior to the private good. Now, if neo-republicans close to Pettit and Skinner reject such an idea, it is precisely because they regard the enjoyment of individual liberties, protected from arbitrary interference from other individuals and from the State, as a primary good (in the Rawlsian sense of the term), rather than as a homogeneous conception of the good hostile to axiological pluralism.

Paul Weithman (2004) suggests that there are two models for thinking about the civic virtues proper to republicanism. According to the first model, civic virtues do not depend on the agents' character traits, but are instead the derivative effects of institutional rules and practices. The "excellence" proper to civic virtue does not, therefore, consist in the *overall* "excellence" of the agent; instead, what matters for civic virtue is *only* the quality of the *relations* between the agent and his community. According to the second model, by contrast, the only way to give a proper account of the civic virtues invoked in the republican project is to recognize that it is, indeed, based on a conception of the good life, which, in turn, involves a tight bond between the agent's virtues and their realization within society: in order to flourish, the agent must actively engage in politics, and the effects of his actions must further be recognized by his fellow citizens as civic virtues (e.g. patriotism). Thus, if the social conditions are favorable, the agent's excellence will completely realize itself within his society, in whose eyes he will in turn appear as a good citizen. Weithman links this strong version of civic virtues with what he calls "perfectionist republicanism." This interpretation of civic virtue seems to conflict with the pluralism inherent to contemporary societies, however; for this reason, many authors would rather link republicanism to a weaker—and hence less controversial—moral conception. Others call for associating the republican theory with a moral conception that would make no claim to the axiological neutrality typical of liberal theories, while still recognizing the "quasi-perfectionist" character of liberty as non-domination (Maynor 2003). According to Maynor, a strictly instrumental view of civic virtues would make it impossible to account for the kind of exchanges between individuals that republicanism ought to aim for: if one grants that the latter seeks to protect citizens in such a way as to make them less vulnerable to the arbitrary power relations that they maintain amongst themselves, then one must concede that the protections tied to the status of non-domination do indeed have an intrinsic value; one must not, therefore, underestimate the important link between a person's development and flourishing, on the one hand, and his status as a free citizen, on the other.

On the other hand, this whole debate has perhaps been skewed by the stress laid on the idea of virtue, whereas the real problem might lie in the normative issues surrounding the concept of citizenship—in which case, we should refocus the debate on the question of whether or not the basic structural conditions tied to citizenship are sufficient to exhaust the meaning of what a properly republican—as opposed to, say, a liberal—citizen is and ought to be.

On the other hand, the principle of non-domination might be taken to mean that all citizens are called upon to participate in public life and to determine the rules in accordance with which they shall achieve their common objectives *themselves*. But what non-domination implies, above all, is not strong civic participation so much as a number of institutional protections designed to shield individuals from attempts to dominate them. Judicial institutions are therefore at the heart of the republican project. It is still necessary to guard against any form of *imperium* by the State against individuals, of course, but well thought-out institutions ought to make it possible to avoid such problems. The role of judicial institutions is to guarantee citizens' liberty, so that everyone can enjoy them to an equal extent (Honohan 2009). Whether it be a matter of political liberty or of the rights of individuals, both cases are inextricably bound to political interactions between citizens; hence the crucial importance of individuals' right to have a say in and oversee political issues that concern them (Sunstein 1988). The question that arises is then to determine the precise role of the judicial review process regarding measures adopted by the legislative or executive powers. In principle, judicial review is meant to provide citizens with a number of protections from arbitrary or authoritarian decisions on the part of the legislative and executive powers. Conversely, even if citizens are able to both protect themselves against excesses by these powers as well as to exert a certain influence on them, do they not still remain at the mercy of the constitution itself (Waldron 2004, 2006)? Is there not an unjustifiable and undemocratic inequality between the constitutional and other kinds of power? That is, either the constitutional power operates above and beyond the will of the citizens, ostensibly in order to guarantee that their fundamental rights be respected—in which case it is deeply incompatible with the ideal of non-domination, as it itself is beyond the reach of the citizens' powers—or, on the contrary, it is viewed as a public forum where debates can be had on behalf of all citizens (Bellamy 2007, 2009; Honohan 2009).

3. Republicanism and Democracy

Thus, if non-domination consists in not being subjected to arbitrary interferences by the State, then citizens must be regarded as participants on an equal footing in all decisions concerning the public sphere. The question is then to determine whether this implies a strong conception of democracy, whereby the citizens' real contribution to collective choices is the true measure of their power. For certain authors, including Pettit himself, while this conception does contain an ideal of self-government, republicanism nevertheless does not favor the participative model of democracy over the representative one. Political liberty does require the existence of a constitutional and judicial system, yet it also requires the possibility of contestation: given that legislators, along with judges and deputies of the public administration (i.e. the executive power), have a certain discretionary power over the content of the laws, it is crucial to prevent this kind of power from progressively becoming abusive.

Since it claims to be a political theory of the normative foundations of political relations as well as a theory of government, moreover, republicanism must give an account not only of the legitimacy of political decisions, but also of the proper functioning of democratic mechanisms. On the one hand, the ultimate goal of the republican government is the common good. On the other hand, everyone will not necessarily agree on one particular conception of what that is. So, if pluralism itself is considered to be intrinsic to the common good, then the latter can only be pursued through some form of

dialogue between the various parties, which in turn presupposes divisions that must be accounted for without thereby losing sight of the republican State's proper objectives.

For this reason, republican theorists, such as Philip Pettit, distinguish between two models of democracy: the "electoral model" and the "contestatory model." The primary objective of the government is to seek non-domination; consequently, the government cannot claim the banner of republicanism if it gives priority to the will of specific groups over the common interest—even if the groups can most stimulate the economic development of the society. By means of electoral mechanisms, the government acts in accordance with the group as a whole—assuming, of course, that that electoral model itself does not give rise to structural discrimination from the outset. Overall, the government cannot claim to be above the people, however, as the latter effectively control it by means of elections. Yet every election involves a certain division amongst the people: the individuals who voted against a certain government may afterwards find themselves incorporated, arbitrarily and against their will, into "the people," conceived as a single and indivisible entity. One does not cease to be a citizen simply because one is opposed to the majority; on the other hand, one does not cease to be a minority as soon as the election has passed. Without monitoring by the citizens, the government can act on the people without the people's being able to act on it. This public system of control as such nevertheless does not imply that citizens must exclusively devote themselves to the business of self-government. For one thing, such an arrangement would not be realistic. Furthermore, it would not be necessary, as long as the government's actions can be submitted to critical scrutiny by the citizens—this is what Pettit calls the "contestatory model" of democracy. While the framing of laws and public policies remains the government's responsibility, the citizens may intervene and oversee the latter in two ways, just as an editorial committee oversees a newspaper: either *ex post*, i.e. by demanding that certain measures be modified after the fact according to various standards, or *ex ante*, i.e. before these measures are even developed, by means of public consultations. To Pettit's eyes, electoral democracy and contestatory democracy can and must be complementary.

The possibility of contestation or "contestability" makes it possible to give a better account of the legitimacy of public institutions than the notion of consent vis-à-vis a supposedly legitimate authority, e.g. an authority that claims to embody the will of the majority (Pettit 1999, 2000). All individuals, as well as, especially, the most vulnerable groups in society (for ethnic, sexual or cultural reasons), must be able to defend their interests (Pettit, forthcoming).

4. New Issues

In spite of the numerous criticisms that it has provoked, republicanism can no longer be seen as just a trend, but must now be reckoned among the major political theories, along with liberalism and libertarianism. So far, however, many sectors of political activity have not yet been studied from the republican perspective, or only to a limited extent.

The neo-republicans' answer to the problems of domination is essentially based on a conception of the State and of government whereby the institutions' primary mandate is to protect, guarantee or "immunize" individuals against any and all forms of arbitrary interference. To repeat, what constitute domination-relations are not just *actual* interferences, but the mere *possibilities* of arbitrary interference. Hence, what needs to be combated, by means of institutional mechanisms, is the individuals' state of vulnerability, although it remains to be seen whether institutions suffice to adequately protect the

citizens or whether the latter must not *also* share a certain common culture in order for all to be equally respected. Clearly, it is difficult to get away from the debate concerning the moral implications of normative principle of domination for the agent himself. Even if the agent is shielded from certain arbitrary practices, what about the esteem in which he is held by others or even by himself, especially if his value is relentlessly denigrated by the majority or by a group in power?

As Pettit himself acknowledges, while domination can take many forms and calls for great vigilance, the precise areas where it needs to be fought can, nevertheless, remain difficult to identify. The problem is to determine (a) where and when the struggle against domination is relevant, since institutions cannot be expected to intervene in situations where certain domination-like relations arise merely by chance and could just as easily have turned out differently, and (b) which criteria, accordingly, would enable us to distinguish a situation in which domination is really present from one in which it is not.

This goes both ways. How to avoid discrediting, from the outset, certain social demands that are made despite the attribution of an identical social status for all but where the dominated individuals may still tend to behave in a servile manner? In other words, should the social struggles of excluded or rejected groups be favored just in order to prevent the emergence of contemptuous attitudes—not only on the part of the majority vis-à-vis the vulnerable groups, but also on the part of the latter themselves? Encouraging such struggles does not imply a return to the paternalistic model rejected by Pettit, Skinner, and the neo-republicans, however; the aim is, rather, to work towards public forums' guaranteeing a fair tribunal for claims that might otherwise go unheard.

It is for this reason that contestatory democracy, precisely in Pettit's sense, is so important, and also why it must be conceived not as a mere procedural mechanism, like an editorial board, that puts the finishing touches on the work of the institutions below it but, rather, as a nexus of social struggles whose normative framework must be thought out by the republican theory.

One example of such social struggles is the fight against cultural discrimination. While this kind of discrimination is most often linked to a claim for identity-recognition (e.g. minority linguistic rights), the resulting social struggles are often motivated by the collateral effects of the discriminations involved. This becomes obvious when one considers that, in many cases, the minority groups that suffer from cultural discrimination *also* suffer from social iniquities, as is the case, for example, with the Aboriginal peoples of Canada.

On the other hand, neo-republicanism cannot deny that adding a widened conception of democracy to its program might either bring it back to the utopian and retrograde model of civic humanism, or else hinder the attainment of its core objectives by the exaggerated importance of social struggles and factions, whereby the strongest and most organized groups would gain the upper hand over weaker ones in terms of social gains and visibility.

After having gradually proven its theoretical validity, and now improved by new theoretical approaches, such as critical theory (Laborde 2008; Le Goff 2007) republicanism has now emerged as a promising tool for thinking about topics as varied as economic theory, social struggles against poverty (Pettit 2007), immigration and multiculturalism, feminism, education (Peterson 2011), or the environment. Also raised anew in light of these issues is the problem of determining the most appropriate institutional design for realizing the republican program and for promoting non-domination. While it is not possible to present all of these topics here, we can nonetheless point to a few of the

fields of study that will most likely move to the center of republican theorists' attention in the coming years.

The republican theory has met with a number of interesting criticisms from a *feminist* point of view, yet none of these seem to radically challenge the idea that the theory of liberty as non-domination could make a valuable contribution to theorizing women's claims (Friedman 2008; Philips 2000). But, for the moment, no comprehensive study of properly feminist issues from the republican perspective has been produced. This field promises to give rise to many new studies, however, concerning not only the fundamental issues surrounding women's vulnerable state in society, but also the institutional means employed to combat the forms of domination to which women in particular are subjected—for these means must themselves be studied according to the republican theory's own parameters of justice and democracy.

As for *ecology*, there are not yet any important studies of the relevance of republicanism to what is now called "environmental justice." Moreover, environmental struggles often give rise to forms of contestation and activism (Young 2001) whose specific mode of action seems to set them apart from the model of contestatory democracy envisaged by Pettit. Although the framework of the republican theory of justice appears to be favorable to ecologists' claims, the political form that these claims actually take—i.e. reaching beyond the level of State sovereignty in order to address international economic institutions instead—suggests that the model of democracy adopted by many republican theorists should perhaps be re-examined.

As we have seen, any interpretation of republicanism and of its program depends to great extent on the moral conception that one associates with the objectives of non-domination as well as on the role that one ascribes to civic virtues therein. The *philosophy of education* thus plays a crucial role here, for it either shows that the moral values associated with republicanism are substantive by definition, or else that republicanism, together with the type of citizenship that it puts forward, points towards a conception of civic education, based above all on the critical capacities of individuals within the public realm—in which case one would have to then formulate the method and goals of a civic pedagogy (Peterson 2009, 2011).

Last but not least, one must also mention the important studies of the contribution of republican theories to the issues surrounding international relations and international justice. As far back as Machiavelli, republican theorists have maintained that no institutional reforms or policies can be effective in the long term if international power relations subsequently undermine them. The problem is then to design regulative mechanisms between States as well as alliances between similarly oriented countries (Anctil 2009; Bohman 2007; Deudney 2007; Slaughter 2005).

Related Topics

Machiavelli, Liberalism, Feminism and the History of Political Philosophy, Environmentalism, Perfectionism, Pluralism, Equality, Freedom, Autonomy, Power, Democracy, Rights, Education

References

Anctil, D. (2009) " La liberté républicaine et la démocratisation du régime international," in *Ateliers de l'éthique* 4: 65–76.

Arendt, H. (1958) *The Human Condition*, Chicago, IL: Chicago University Press.
Baron, H. (1955) *The Crisis of the Early Italian Renaissance: Civic Humanism and Republican Liberty in an Age of Classicism and Tyranny*, *Vols 1 and 2*, Princeton, NJ: Princeton University Press, revised edn. 1966.
Bellamy, R. (2002) "Being Liberal with Republicanism's Radical Heritage: A Comment on Pettit," *Res Publica* 8(3): 269–74.
—— (2007) *Political Constitutionalism: A Republican Defense of the Constitutionality of Democracy*, Cambridge: Cambridge University Press.
—— (2009) "The Republic of Reasons: Public Reasoning, Depoliticization, and Non-Domination," in S. Besson and J. L. Marti, eds., *Legal Republicanism: National and International Perspectives*, Oxford and New York: Oxford University Press.
Berlin, I. (1969) "Two Concepts of Liberty," in *Four Essays on Liberty*, Oxford: Oxford University Press.
Bohman, J. (2007) *Democracy Across Borders: From Demos to Demoi*, Cambridge, MA: MIT Press.
Brennan, G. and Lomasky, L. (2006) "Against reviving republicanism," *Politics, Philosophy and Economics* 5: 221–52.
Carter, Ian, (2008) "Positive and Negative Liberty," *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition)*, Edward N. Zalta (ed.), URL: <http://plato.stanford.edu/archives/fall2008/entries/liberty-positive-negative/>.
Dagger, R. (2006) "Neo-republicanism and the civic economy," *Politics, Philosophy and Economics* 5: 151–73.
Deudney, D. H. (2007) *Bounding Power: Republican Security Theory from the Polis to the Global Village*, Princeton, NJ: Princeton University Press.
Etzioni, A. (1993) *The Spirit of Community: Rights, Responsibilities and the Communitarian Agenda*, London: Fontana Press.
Friedman, M. (2008) "Pettit's Civic Republicanism and Male Domination," in C. Laborde and J. Maynor, eds., *Republicanism and Political Theory*, London: Blackwell.
Galston, W. A. (1991) *Liberal Purposes*, Cambridge: Cambridge University Press.
Geuna, M. (1998) "La tradizione repubblicana e i suoi interperti: famiglie teoriche et discontinuità concettuali," *Filosofia politica* 1: 101–34.
—— (2009) "Quentin Skinner e Machiavelli," in A. Arienzo and G. Borrelli, eds., *Anglo-American Faces of Machiavelli*, Milan: Polimetrica Publisher.
Goodin, R. E. (2003) "Folie Républicaine," *Annual Review of Political Science* 6: 55–76.
Guérard de Latour, S. (2009) *Vers la république des differences*, Toulouse: Presses universitaires du Mirail.
Hankins, J., ed. (2000) *Renaissance Civic Humanism: Reappraisals and Reflections*, Cambridge: Cambridge University Press.
Honohan, I. (2009) "Republicans, Right, and Constitutions: Is Judicial Review Compatible with Republican Self-Government?" in S. Besson and J. L. Marti, eds, *Legal Republicanism: National and International Perspectives*, Oxford and New York: Oxford University Press.
Laborde, C. (2008) *Critical Republicanism. The Hijab Controversy and Political Philosophy*, Oxford: Oxford University Press.
Larmore, C. (2004) "Liberal and Republican Conceptions of Freedom," in D. Weinstock and C. Nadeau, eds, *Republicanism: History, Theory, and Practice*, London: Routledge.
Lazzeri, C. (2001) "Prendre la domination au sérieux: une critique républicaine du libéralisme," *Actuel Marx* 28, 55–68.
Le Goff, A. (2007) "Néo-républicanisme et feminisme. La question de la parité," in V. Bourdeau and R. Merril, eds, *La république et ses demons. Essais de républicanisme appliqué*, Paris: Eres.
Lovett, F. (2010) *A General Theory of Domination and Justice*, Oxford: Oxford University Press.
Maynor, J. W. (2003) *Republicanism in the Modern World*, Cambridge: Polity.
Moulakis, A. (2008) "Civic Humanism," in *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL: http://plato.stanford.edu/archives/fall2008/entries/humanism-civic/>.
Patten, A. (1996) "The Republican Critique of Liberalism," *British Journal of Political Science* 26: 25–44.
Peterson, A. (2009) "Civic republicanism and contestatory deliberation: framing pupil discourse within citizenship education," *British Journal of Educational Studies* 57: 55–69.
—— (2011) *Civic Republicanism and Civic Education. The Education of Citizens*, Basingstoke: Palgrave Macmillan.
Pettit, P. (1996) "Freedom as Power," *Ethics* 106: 576–604.
—— (1997a) *Republicanism: A Theory of Freedom and Government*, Oxford: Clarendon Press.
—— (1997b) "Freedom with honor: a republican ideal," *Social Research* 64: 52–76.

—— (1997c) "Republican Political Theory," in Andrew Vincent, ed., *Political Theory: Tradition, Diversity and Ideology*, Cambridge: Cambridge University Press.
—— (1999) "Republican Freedom and Contestatory Democratization," in Ian Shapiro and Casiano Hacker-Cordon, eds., *Democracy's Value*, Cambridge: Cambridge University Press.
—— (2000) "Democracy, Electoral and Contestatory," *Nomos* 42: 105–44.
—— (2007) "Republican Rights to a Basic Income," *Basic Income Studies* 2.
—— (forthcoming) *On the People's Terms*, Cambridge: Cambridge University Press.
—— and Braithwaite, J. (1990) *Not Just Deserts: A Republican Theory of Criminal Justice*, Cambridge: Oxford University Press.
Philips, A. (2000) "Feminism and republicanism: is this a plausible alliance?," *The Journal of Political Philosophy* 8: 279–93.
Pocock, J. G. A. (1975) *The Machiavellian Moment. Florentine Political Thought and the Atlantic Republican Tradition*, Princeton, NJ: Princeton University Press.
Rawls, J. (1993/1996/2005) *Political Liberalism*, New York: Columbia University Press.
Raz, J. (1986) *The Morality of Freedom*, Oxford: Clarendon Press/Oxford University Press.
Skinner, Q. (1998) *Liberty before Liberalism*, Cambridge: Cambridge University Press.
Slaughter, S. (2005) *Liberty beyond Neo-Liberalism: A Republican Critique of Liberal Government in a Globalising Age*, London: Macmillan Palgrave.
Spitz, J.-F. (1995) *La liberté politique: essai de généalogie conceptuelle*, Paris: Presses Universitaires de France.
—— (2000) *L'amour de l'égalité. Essai sur la critique de l'égalitarisme républicain en France 1770–1830*, Paris: Vrin/EHESS.
Sunstein, C. R. (1988) "Beyond the Republican Revival," *The Yale Law Journal* 97(8), *Symposium: The Republican Civic Tradition*: 1539–90.
Viroli, M. (1990) "Machiavelli and the Republican Idea of Politics," in G. Block, Q. Skinner and M. Viroli (eds.), *Machiavelli and Republicanism*, Cambridge: Cambridge University Press.
Waldron, J. (2004) "Judicial Review and Republican Government," in C. Wolfe (ed.), *That Eminent Tribunal: Judicial Supremacy and the Constitution*, Princeton, NJ: Princeton University Press.
—— (2006) "The core case against judicial review," *The Yale Law Review* 115: 1346–1406.
—— (2007) "Pettit's molecule," in G. Brennan, R. E. Goodin, F. Jackson and M. Smith, eds., *Common Minds: Themes from the Philosophy of Philip Pettit*, Oxford: Oxford University Press.
Weithman, P. (2004) "Political Republicanism and Perfectionist Republicanism," *The Review of Politics* 66: 285–312.
Young, I. M. (2001) "Activist challenges to democracy political theory," *Political Theory* 29: 670–90.

Further Reading

Besson, S. and Marti, J. L., eds (2009) *Legal Republicanism: National and International Perspectives*, Oxford and New York: Oxford University Press.
Weinstock, D. and Nadeau, C., eds. (2004) *Republicanism: History, Theory, and Practice*, London: Routledge.

24
MARXISM AND CONTEMPORARY POLITICAL THOUGHT

Alex Callinicos

Marxism's relationship to politics is a contested one. Of course, a research program regarded by its founder as indissolubly connected to a revolutionary socialist political project is itself political. But it doesn't follow that it has an adequate place for the political in its conceptual system. Indeed, Ralph Miliband notes at the beginning of *Marxism and Politics* that "the available classical writings [of Marx, Engels, and their immediate successors] are simply silent or extremely perfunctory over major issues of politics and political theory" (Miliband 1977: 2). One explanation of this apparent lacuna is that Marx, in giving analytical priority to the economic, reduces political conflict to what Slavoj Žižek calls "a shadow theatre in which events whose proper place is on Another Scene (of economic processes) are played out" (Žižek 1999: 189). This view is shared, not merely by critics, but also by some champions of Marxism. Fredric Jameson argues that the idea that

> the crucial weakness of Marxism was that it structurally lacked a dimension of political (and juridical) theory … was a great mistake … the very force and originality of Marxism was always that it did not have a political dimension of this kind, and that it was a completely different thought system (or unity of theory and practice) altogether.
>
> (Jameson 2009: 299)

One can, however, find plenty of other Marxists willing to challenge the idea that they are committed to reducing the political to a purely epiphenomenal status. Thus Stathis Kouvelakis, writing from a standpoint otherwise sympathetic to those of Jameson and Žižek, argues: "Rather than an Achilles heel, or the sign of a troubling lacuna, politics is, in my opinion, Marx's *strong point*, the point where his work is at its most open and innovative" (Kouvelakis 2003: 351). Kouvelakis bases his view on an interpretation of the development of the young Marx, but those sharing his position can point to the later trajectory of Marxist thought, notably in the writings of Lenin and Gramsci. As Daniel Bensaïd puts it:

> Lenin was one of the first to conceive the specificity of the political field as a play of transfigured powers and social antagonisms, translated into a language

> of its own, full of displacements, of condensations and of revealing slips of the tongue.
>
> (Bensaïd 2004: 121)

Lenin was, of course, the greatest (or most notorious) Marxist practitioner of politics. One might then say that, precisely because of this preoccupation with practice, he was led to highlight the specificity of politics, conceptualizing it as, in Miliband's words, "the pervasive and ubiquitous articulation of social conflict and particularly of class conflict" (Miliband 1977: 6). It doesn't follow that Lenin offered any systematic theorization of politics: famously *The State and Revolution*, his most important text on the subject, was left unfinished as he turned to leading the Revolution of October 1917. In Gramsci's case the contingencies of political struggle allowed him a period of enforced reflection from which there indeed emerged a major, though frequently ill-understood attempt to rethink the Marxist theory of the state (Thomas 2009). But the *Prison Notebooks* stand out as an exception to the situation described by Miliband.

However one explains what he calls "the absence in classical Marxism of systematic theorization of Marxist politics," the past generation has seen a major effort to fill this lacuna (Miliband 1977: 6). This has taken two directions, of very different status—the attempts to develop, on the one hand, a theory of the state, and, on the other, a theory of justice. The differences in status derive partly from the kinds of theory involved: state theory belongs relatively unproblematically to the domain of empirical social science, even if the Marxist version of the latter may vary in some respects from the academic mainstream; developing principles of justice, by contrast, requires invading the territory of normative political philosophy. This brings us to another difference, namely that Marx treated all moral and ethical discourses as forms of ideological misrepresentation, and hence ruled out in principle the possibility of a Marxist theory of justice. The rest of this essay is devoted to reviewing the development of these two areas of Marxist discussion and the difficulties they have confronted.

1. Understanding the State

The renaissance of critical Marxist theory that accompanied and was stimulated by the political radicalization of the 1960s and 1970s had the state as one of its main focuses. In part, this represented a continuation of earlier debates over power in advanced capitalist societies—for example, the disputes between radicals such as C. Wright Mills, who saw post-war America as dominated by an interlocking power elite, and the mainstream school of pluralism in political science, who saw power as fragmented and dispersed as elites compete in different issue-areas (Lukes 2004). One can see the influence of these debates in some of Miliband's early interventions, notably *The State in Capitalist Society* (Miliband 1969, criticized in Balbus 1971).

But one of the main impulses in the revival of Marxism was an effort at intellectual renewal; the thought was that a return to the classical sources could escape the ideological contamination of Soviet orthodoxy. A natural form this took was the close study of Marx's masterwork, *Capital*. The collective project *Reading Capital* by Louis Althusser and his pupils was merely the tip of a much larger iceberg (Althusser and Balibar 1970). But this effort was not intended as a devout exercise but as a step towards renovation by, in part, identifying and seeking to address the silences and aporias that reading the classical texts revealed. This involved addressing issues that, whether by simple neglect or

because of some deeper conceptual difficulty, had not been properly theorized. Among these, as we have already seen, was politics. Thus, at the beginning of the most celebrated work of Marxist political theory to appear in this period, *Political Power and Social Class*, Nicos Poulantzas anticipates Miliband (though in a very different intellectual idiom), noting that Marx, Engels, Lenin, and even Gramsci "did not specifically discuss the region of the political at the level of theoretical *systematicity*" (Poulantzas 1973: 19). This he sought to remedy in a treatise deeply influenced by Althusser's theoretical reconstruction of Marxism.

The development of Marxist political theory in the 1960s and 1970s, in focusing on the state, gave a specific social location to the broader conception of the political (though this largely implicit identification of politics and the state came under challenge, as we shall see). Indeed, debate centered on the problem, in Miliband's words, of the state in capitalist society. (The best overview is provided by Jessop 1982; see also Barrow 1993.) There were good political reasons for this. The liberal democratic state had changed significantly after 1945, developing a capacity for macroeconomic management and providing a much broader range of welfare services than previously. Did this transformation represent the transcendence of its role as an instrument of class domination, which Marx and Lenin claimed was a necessary feature of any form of state rule? Addressing this question was a major preoccupation of the so-called German "capital logic" school (for example, Holloway and Picciotto 1978). Moreover, to the extent that Marxist theory was intended to inform political practice, a better understanding of the capitalist state might lead to the formulation of more effective political strategy: this problem was a major preoccupation of both Miliband and Poulantzas.

The disagreements between these two figures—representatives of the very different intellectual styles of their adopted countries (respectively, the British empirical idiom and French high theory)—formed the focus of one of the two main debates that defined the development of Marxist state theory in the 1970s: if the state in capitalist society is a capitalist state, precisely how is this outcome secured? In other words, what ensures that state institutions and policies serve the interests of the capitalist class? Miliband, initially at least, sought to focus on the empirically ascertainable relationships between economic and political elites. Poulantzas responded by highlighting what he believed to be

> the difficulties that Miliband has in comprehending social classes and the State as objective structures, and their relations as an objective system of regular connections, a structure and a system whose agents, "men," are in the words of Marx, "bearers" of it—*träger*. Miliband constantly gives the impression that for him social classes or "groups" are in some way reducible to inter-personal relations, that the State is reducible to inter-personal relations of the members of the diverse "groups" that constitute the State apparatus, and finally that the relation between social classes and the State is itself reducible to inter-personal relations of "individuals" composing social groups and "individuals" composing the State apparatus.
>
> (Poulantzas 1969: 70)

This criticism reflected the influence of Althusser, for whom Marx conceived history as a "process without a subject": in other words, historical processes are driven by the strains that develop within social structures, while human individuals are formed as apparently coherent and autonomous subjects who in fact help to sustain these struc-

tures and lack any capacity for independent initiative. But it also implied a different way of conceptualizing the relationship between the state and capital, in which the state is a structure within a larger system of structures. Poulantzas argues that "*the state has the particular function of constituting the factor of cohesion between the levels of a social formation*," which it performed by achieving a degree of "relative autonomy" from the economy, thereby enabling it the more effectively to unify the dominant classes and to fragment and divide the dominated classes (Poulantzas 1973: 44). So where Miliband sought to deconstruct the apparent separation of the liberal democratic state from economic relations, Poulantzas highlighted the *distance* between the two—not, however, to reaffirm the legitimacy of the status quo but in order to demonstrate how this distance was required in order to secure the continued dominance of capitalism. For Poulantzas, the separation of economics and politics gives those controlling the state the room for maneuver, and also the legitimacy deriving from the state's apparent autonomy required to ensure that society operates in the interests of the capitalist class.

Miliband responded robustly, accusing Poulantzas of a "structural super-determinism" indifferent to specific forms of state—for example, liberal democracy and fascism (Miliband 1970: 57–8). In fact, in subsequent works Poulantzas turned his attention to the "exceptional form" of capitalist state represented by military dictatorship and fascism (Poulantzas 1974 and 1976b). In all likelihood, this was not an attempt on Poulantzas's part to fill a gap identified by Miliband: the subjection of the former's native Greece between 1967 and 1974 to a right-wing military junta would be reason enough for him to thematize the problem of dictatorship. But this case is illustrative of a more general tendency for the protagonists to approach each other, though less at the level of rhetoric than that of substance. Thus, in subsequent writings Miliband discusses state autonomy extensively and lays explicit emphasis on structural determinations:

> [T]hese three modes of explanation of the nature of the state—the character of its leading personnel, the pressures exercised by the economically dominant class, and the structural constraints imposed by the mode of production—constitute the Marxist answer to the question why the state should be considered as the "instrument" of the "ruling class."
>
> (Miliband 1977: 73–4)

Meanwhile, Poulantzas, in his last book, *State, Power, Socialism* (1978) makes state structures and policies much more dependent on the hazards of social and political struggles than he had appeared previously to allow, thereby retreating from the apparent functionalism of *Political Power and Social Class*. Thus he conceptualizes the state as "the *specific material condensation* of a relationship of forces between classes and class formations," a formula that allows the dominated classes to help shape the state and public policy (Poulantzas 1978: 129). Poulantzas consequently moves away from his earlier, classical Marxist view that "a socialist revolution does not signify only a shift in State power, but it must equally 'break,' that is to say radically change, the State apparatus" towards a strategy much closer to Miliband's, in which social transformation depends on an articulation of direct and parliamentary democracy, rather than, as Lenin thought, a confrontation between the two (Poulantzas 1969: 129; for a critique of his later stance, Barker 1979). So Miliband comes to allow more significance to, or at least to be more explicit about the significance of the structural, and Poulantzas comes to give more weight to the play of the political, including the influence of non-dominant classes.

This asymptotic convergence (Poulantzas's suicide in 1979 means, alas, that we will never know how much closer he might have drawn to Miliband) invites reflection on what else, despite their manifest differences in intellectual style, they might have in common. One problem largely unexplored by both was identified by Ernesto Laclau in a perceptive commentary on the debate: namely the separation of the economic and the political characteristics of liberal democratic societies (Laclau 1977). Miliband seeks to demonstrate the real connections between economic and political elites, arguing, for example, that the "relative distance of businessmen from the state system is markedly reduced by the [bourgeois] social composition of the state elite proper" (Miliband 1969: 59). Meanwhile Poulantzas hypostatizes the separation, making it a structural feature of all modes of production. Neither interrogates it, enquiring into its conditions of possibility. One of the distinctive features of the second major controversy in Marxist state theory during the 1970s, the so-called "state-derivation debate," was that it undertook precisely this form of enquiry. (See, in English, Holloway and Picciotto 1978; Clarke 1991.)

Though, like the Miliband–Poulantzas debate, a transnational discussion, it was otherwise very different—not a confrontation between two individuals, but a series of exchanges mainly among German and British Marxist scholars. Its starting point lay in the efforts by the "capital logic" school that emerged in West Germany in the 1960s, through close study of Marx's economic writings (not just *Capital*, but also the earlier manuscript known as the *Grundrisse*), to reconstruct the conceptual structure of his theory. Because of this intellectual context, the versions of state theory that emerged were indissociably connected with the broader development in this period of Marxist political economy, which was stimulated by the effort to comprehend a very different capitalism from that of Marx's day, a capitalism moreover that was, by the beginning of the 1970s, visibly entering into crisis.

Marx, in *Capital*, conceives the capitalist mode of production as an articulated totality. But the structure of this whole is not revealed in some kind of intuitive flash, but is uncovered through the progressive introduction of concepts that simultaneously specify the properties of particular determinations and explain them by situating them within this structure. So at the start of *Capital* Volume I, Marx looks first at the commodity, then money, and then capital (as money that expands in value), before introducing labor-power, whose transformation into a commodity makes possible the exploitation of the worker that is constitutive of capitalism. Because capitalism is a totality, the existence of an apparently autonomous state is a problem that requires explanation. Thus in a wide-ranging critique of orthodox Communist and social-democratic theorists' "reification of the state," Wolfgang Müller and Christel Neusüss insist that "the state is a particularity of a specific capitalist mode of production." Hence

> this particularization to an existence "beside and outside" of bourgeois society takes place on the basis of bourgeois society, i.e., on the inherently contradictory basis of capitalist production. The real particularization on the basis of this contradiction thus leads to the "wrong," "mystified," idealist conception which juxtaposes the state as independent *vis-à-vis* society.
>
> (Müller and Neusüss 1975: 13, 73)

Now if the separation of the economic and political isn't simply illusory—Müller and Neusüss talk about the "*real* particularization" of the capitalist state—it has to be

explained. The form the explanation took—that of *deriving* the state from capitalist production relations—reflected the paradigmatic role of Marx's method in *Capital*. The state had to be shown to be one among the determinations articulated together to form the totality of the capitalist mode of production. In his early plans for *Capital*, Marx included a volume on the state, but, as with so much else, he never realized this intention. The challenge of the state-derivation debate was to complete Marx's project in a manner consistent with his method (though there has been extensive discussion of the nature and coherence of this method). It is impossible here to summarize a many-sided and wide-ranging series of exchanges, so what follows merely highlights some themes and problems.

In one of the most important contributions, Joachim Hirsch highlighted the issue of form: "This question of what distinguishes the bourgeois state from all previous forms of the exercise of power and domination, is a question of the specific social *form* of the state and not of the particular content of its activity" (Hirsch 1978: 59). This emphasis involved, once again, following a cue offered by Marx. In *Capital* he argues that the main importance of the labor theory of value (which affirms that goods and services exchange in proportion to the socially necessary labor-time required to produce them) is less to provide a means of determining relative prices than to explain why in capitalism the products of human labor take the form of commodities circulating on the market. The answer lies in the fact that capitalism is a system of general commodity production, where the units of production are both autonomous and interdependent, so that the survival of each depends on its ability to sell its own products and thereby to obtain the money it requires to purchase the products of other units. The effect is to subject producers to the imperative of minimizing their costs by raising labor productivity to the levels prevailing in their sector.

This process is also responsible for the phenomenon of what Marx calls commodity fetishism. He argues that an inversion tends to take place under capitalism whereby the social relations among human beings in production appear as relations between the products of their labor. In this way the market is naturalized, as the exchange of commodities is represented as the horizon of social existence. But—and here is a crucial twist—this naturalization isn't purely illusory. As Marx puts it:

> To the producers, therefore, the social relations between their private labours *appear as what they are*, i.e. they do not appear as direct social relations between persons in their work, but rather as material relations between persons and social relations between things.
>
> (Marx 1976: 165–6; italics added)

In other words, social relations are really mediated by the exchange of commodities under capitalism, and therefore, while representations that take this reality for granted lead us to misrecognize the historically distinctive character of the capitalist mode of production, they capture a partial truth about how it functions.

The state-derivation debate rests to a significant extent on the assertion that the capitalist state is itself a fetishized form of capitalist production relations. As such, it shares the duality that we have just encountered in commodity fetishism as such: simply to take the separation of the economic and political at face value is to naturalize the state-form, and thereby to accept the fragmented appearance through which capitalist social relations present themselves; at the same time, the "particularization" of the state

as a distinct and apparently autonomous set of public institutions is a social reality. Marx's value theory also offered a more direct way of explaining the separation of the economic and political itself. He argues that in earlier forms of class society, exploitation required the application of "extra-economic coercion" (for example, the military and juridical power of the feudal lords) to extract surplus-product from the direct producers. But, under capitalism, "[t]he silent compulsion of economic relations sets the seal on the domination of the capitalist over the worker" (Marx 1976: 899). Compared to slaves or serfs, the worker is free legally and politically to dispose of his or her labor-power. But the worker also lacks direct access to all productive resources. To gain access to them, and thereby to survive materially, the worker must sell his or her labor-power to the capitalist, who controls the means of production. Economic pressures rather than direct coercion at the point of production compel the worker to submit to exploitation. This makes possible the differentiation of the economic and political relations and the emergence of the state as a public institution apparently autonomous of the "private" sphere of production (Hirsch 1978: 62ff.; see Wood 1981, for the fullest development of this idea).

This conception of the state is, of course, at a high level of abstraction. One of the main stakes in the state-derivation debate arose from diverging ways of specifying it. Another concerns the move from "form" to "function"—to put it crudely, from the general conception of the state as a "particularization" of the capital-relation to an account of what states actually do. Hirsch warned that

> a strict "derivation" of the functions of the state is still not possible in an investigation which remains at the analytical level of "capital in general" … since concrete state functions come into being only through the mediation of competition and class struggle.
>
> (Hirsch 1978: 83)

But this warning was not always observed, producing a particularly abstract form of functionalism in which the beneficial effects of state institutions and policies on the interests of capital were simply asserted, without any attempt to come up with empirical accounts of the mechanisms through which this relationship was secured and reproduced. There is, moreover, the danger inherent in a research program that seeks to derive the state-form from the capital-relation that, at worst it collapses into a version of economic reductionism, at best the idea of the political as the crystallization of social contradictions common to Miliband and Poulantzas and the implication that politics is a distinct social field with its own specific logic and properties disappear. Finally, the effort to hold together the thought that the state is simultaneously a fetishized form of capital *and* a functioning reality can collapse into a much simpler view of the capital-relation as a form of domination reproduced in all aspects of social life: this, in effect, is the tendency followed by John Holloway and the school of "Open Marxism" in recent decades (for example, Holloway 2002).

Nevertheless, one can see a broader convergence emerge among many Marxist state theorists than that between Miliband and Poulantzas. Maintaining the dual vision of the state as a fetishized form and a functioning reality required coming up with an explanatory account of the apparent autonomy of the state consistent with the state operating in the interests of capital. The same solution seems to have occurred independently to different scholars. Grant that state managers have their own interests dif-

ferent from those of capitalists—let's say in maintaining (and, if possible, increasing) the power of their state against both its subjects and other states. Securing this objective would depend on the state's ability to extract resources, a function of both state capabilities and the size and productivity of the economy. Where capitalist economic relations prevail, this would give state managers an incentive to promote capital accumulation. Should they fail to do so, the movement of capital out of the state, and the negative consequences that would follow for the currency, the state's ability to borrow money, and so on, would provide a sanction sufficient to push them back into line (see Block 1987; Harman 1991; Miliband 1983; Offe and Ronge 1982).

One of the attractions of this explanation is that it offered a way of transcending the Miliband-Poulantzas debate. It posited a mechanism that didn't depend on inter-personal relations between economic and political elites (as Miliband's approach did), but that, in specifying the structural constraints on state managers, at the same time sought to take into account the interests and intentions of individual and collective actors (as Poulantzas's tended not to). Yet this elegant solution emerged at a time when the steam seemed to have run out of Marxist state theory. No doubt this had a lot to do with the political disappointment of a generation of radical intellectuals whose hopes of rapid change were dashed, particularly after the emergence of neoliberalism at the end of the 1970s and its triumph during the subsequent decade. But the focus common in the debates I have been discussing on the state came under formidable intellectual challenge, notably from Michel Foucault. His theory of "power-knowledge" was intended explicitly to displace attention from the state to an understanding of power as a nexus of micro-relations dispersed throughout the social. Thus for Foucault the practices distinctive to, for example, institutions such as the prison and to apparently natural domains of human existence such as sexuality, harbor highly specific power-relations that represented the real locus where domination is formed and reproduced. Macro-institutions such as the state are largely the resultant of these localized and decentralized forms of power, rather than being, as Marxists tended to assume, the main site where capitalist domination is maintained (for example, Foucault 1976 and 1979). In *State, Power, Socialism* Poulantzas offered a vigorous response that sought partially to rebut but also to accommodate Foucault's critique of Marxism. But, despite the anomalies increasingly apparent in the rival construction, the sense of impasse among many Marxist state theorists was evident.

Happily, the story doesn't end there. The revival of Marxist state theory over the past decade can be seen as a response to a presupposition common to the protagonists of both the Miliband–Poulantzas and the state-derivation debates pointed out by Colin Barker at the time, namely that they focused on the problem of *the* capitalist state, ignoring the fact that states exist as a plurality forming the international system (Barker 1978). When serious discussion of state theory resumed in the late 1990s, it was in a very different context from that of the 1960s and 1970s, one defined by the dominant discourse about globalization and the contestation of neoliberalism that developed with the new wave of protest movements starting with the Seattle summit in November 1999. The kind of divisions that emerged in mainstream debates over whether or not economic globalization marked the end of the nation-state were replicated on the radical left. The idea current on the center-left that new forms of "global governance" were developing to steer the world economy was affirmed by various Marxist scholars, notably Michael Hardt, Toni Negri, and William Robinson (Hardt and Negri 2000; Robinson 2004). Hardt's and Negri's theorization of Empire as a transcendence of the nation-state and

its replacement by a new form of transnational sovereignty, helped to stimulate a revival of the Marxist theory of imperialism—though, once again, this development received a mighty push from real-world events, above all the presidency of George W. Bush (see Harvey 2003; Wood 2003; Callinicos 2009). In this context, the assumption unproblematized in earlier debates of a plural state system became a subject of considerable discussion (Rupert and Smith 2002; Callinicos 2007; Anievas 2010). In a changed world, Marxist state theory has experienced a notable renewal of vigor.

2. Constructing Justice

The interruption noted above in discussion of the state (or states) for two decades didn't mean that Marxist political thought went wholly into hibernation. But the most lively developments in the 1980s were in a wholly unexpected direction, namely the place of distributive justice in the socialist critique of capitalism. To understand the issue we must return to Marx's analysis of capitalist exploitation. Here again, we encounter duality. From the perspective of "the sphere of circulation or commodity exchange," "the exclusive realm of Freedom, Equality, Property and Bentham," the bargain between capitalist and worker appears to be a transaction between two free and equal commodity-owners, and Marx assumes that the worker receives the full value of the labor-power she sells for a wage so that equivalents are exchanged (Marx 1976: 280). Once, however, we enter "the hidden abode of production," the aspect changes (Marx 1976: 270). The commodity the worker sells has the unique property that, when it is used as labor in the production of commodities, it creates new value, and moreover more value than is embodied in it. It is this difference between the new created value and the value of labor-power—the surplus-value—that the capitalist appropriates as profits. But Marx insists that this "is by no means an injustice to the seller," because what the worker sold was her labor-*power*, for which she received the full value (Marx 1976: 301). Hence when, for example, workers seek to limit the length of the working day, we have "an antinomy of right against right, both equally bearing the seal of the law of exchange. Between equal rights, force decides" (Marx 1976: 344).

Marx puts it like this partly because he wants to show that socialists such as Proudhon who argued that the ills of capitalism stemmed from distortions of the market by banks, and different forms of monopolies. But he is also committed to a relativist meta-ethics according to which moral principles and conceptions have no universal validity but exist to help reproduce the prevailing relations of production. Behind this lie a variety of considerations—a political rejection of moral reconciliation between the classes as an alternative to revolution, the influence of Hegel's critique of Kantian universalism, and Marx's own materialist theory of history. But the resulting position is uncomfortable: one only has to read chapter 10 of *Capital*, Volume I, "The Working Day," from which the last quotation is taken, to get a strong sense that, to say the least, Marx does not regard the capitalist and the worker as equally justified; indeed, it is hard not to feel that the passion with which he describes and denounces the appalling suffering experienced by workers in Victorian Britain as a moral one. But this then implies that the Marxist theory of exploitation is incoherent.

Whether or not this is the case and, if it were, what the remedy should be became the subject of extensive controversy among Marxist philosophers in the 1970s and 1980s (see the bibliographies in Geras 1985 and 1992). Yet again, the context is important. The appearance of John Rawls's *A Theory of Justice* in 1971 both initiated a major

debate around conceptions of egalitarian justice and demonstrated that liberal political philosophy was not as sterile as Marxists had tended to assume. This development prompted some philosophers who had been attracted to Marxism to reflect on whether they might have a more fruitful relationship to normative political philosophy than Marx himself believed possible. Those who did so tended to have been trained in the analytical tradition—hence the debate about Marx and justice tended, by contrast with the exchanges among state theorists, to take place predominantly in the Anglophone academy. A powerful case could be made for Marx's own position (see, notably, Wood 2004). Nevertheless, in a compelling survey of the debate and of Marx's texts, Norman Geras offered what, in my view, is a definitive solution: "Marx did think capitalism was unjust but he did not think he thought so" (Geras 1985: 70). In other words, his critique of capitalist exploitation relies on the idea that the worker is robbed: thus he refers to "*the theft of alien labour-time, on which the present wealth is based*" (Marx 1973: 705). But Marx is prevented from acknowledging his tacit reliance on an unstated conception of justice by his relativist meta-ethics (since this conception would have to have a validity that transcended capitalism), but also, Geras suggests, by "an extremely narrow conception" of justice that equated it with positive legal norms regulating the distribution of consumption goods (Geras 1985: 70–1).

Geras's resolution of the debate insofar as it concerned the interpretation of Marx posed two questions: What is the content of the required conception of justice? And how much of a reconstruction of Marxism does its formulation entail? On the whole, addressing the latter question came to predominate. This was largely the consequence of the emergence of the school of theorists known as Analytical Marxists, who preferred the Anglophone philosophical style to the Continental traditions that had hitherto been the main reference point in Marxist philosophical discussion. (For a representative compendium, see Roemer 1986.) Substantively, Analytical Marxists sought to restate Marxist propositions in the theoretical idioms of mainstream social science—a trail first burned by G.A. Cohen in his magisterial restatement of historical materialism (Cohen 1978). Not surprisingly, quite a lot of Marxism failed the test. Marx's value theory was a particular target of criticism, though this originated in earlier controversies among left-wing economists that then intersected with and reinforced the trajectory of Analytical Marxism (for example, Steedman *et al.* 1981). The initial effect of acknowledging that Marx's critique of capitalism required the formulation of principles of justice, when combined with the rejection of his value theory, was an attempt to restate the Marxist conception of exploitation in ways that did not depend on this theory (notably Roemer 1982). But this was followed by a more straightforward insertion of leading Analytical Marxists in mainstream controversies among moral and political philosophers in the egalitarian liberal tradition stemming from Rawls (for example, Roemer 1996; Cohen 2008 and 2011; for a more critical survey, see Callinicos 2000).

For some Marxist philosophers, then, acknowledging Marx's reliance on a normative conception of justice led them eventually to abandon Marxism altogether. Whether they were logically impelled to take this course is, however, another matter. One of the problems of Cohen's reconstruction of Marxism lay in the distance it took, in large part because of his hostility to the labor theory of value, from the critique of political economy that was Marx's central intellectual achievement (Callinicos 2006). But a strength of Marxist state theory, particularly in the forms it took in the state-derivation debate, has been its close connection with political economy—something that has been preserved in more recent attempts to theorize globalization. One might indeed think

that, rather than being undermined, the Marxist critique of political economy might be strengthened by a closer engagement with normative questions. Mainstream political philosophy has been marked in recent years by a preoccupation with the question of distributive justice on a global scale—something that ought to be a matter of some political urgency given the extremes of wealth and poverty displayed by the contemporary world. It would seem that the Marxist critique of exploitation and imperialism would have much to gain from philosophical discussions of global justice. So we may see the two hitherto separate currents of Marxist political thought, concerned respectively with capitalism and states and with exploitation and justice, begin to run together.

Related Topics

Marx, Foucault, Needs and Distributive Justice, Global Justice and Politics

References

Althusser, Louis and Étienne Balibar (1970) *Reading Capital*, London: NLB.

Anievas, Alexander, ed. (2010) *Marxism and World Politics*, London: Routledge.

Balbus, Isaac (1971) "Ruling Elite Theory vs. Marxist Class Analysis," *Monthly Review* 23 (1): 36–46.

Barker, Colin (1978) "A Note on the Theory of Capitalist States," *Capital and Class* 4: 118–29.

—— (1979) "A New Reformism? A Critique of the Political Theory of Nicos Poulantzas," *International Socialism* 2 (4), available at http://www.isj.org.uk/?id=294.

Barrow, Clyde (1993) *Critical Theories of the State*, Madison: University of Wisconsin Press.

——, Burnham, Peter, and Wetherly, Paul, eds. (2008) *Class, Power and State in Capitalist Society: Essays on Ralph Miliband*, Basingstoke: Palgrave.

Bensaïd, Daniel (1997) *Le Pari mélancolique*, Paris: Fayard.

—— (2002) *Marx for Our Times*, London: Verso.

—— (2004) *Une lente impatience*, Paris: Stock.

Block, Fred (1987) *Revising State Theory*, Philadelphia: Temple University Press.

Callinicos, Alex (1983) *The Revolutionary Ideas of Karl Marx*, London: Bookmarks.

—— (2000) *Equality*, Cambridge: Polity.

—— (2006) "G.A. Cohen and the Critique of Political Economy," *Science and Society* 70: 252–74.

—— (2007) "Does Capitalism Need the State System?" *Cambridge Review of International Affairs* 20: 533–49.

—— (2009) *Imperialism and Global Political Economy*, Cambridge: Polity.

Clarke, Simon, ed. (1991) *The State Debate*, Basingstoke: Macmillan.

Cohen, G.A. (1978) *Karl Marx's Theory of History—a Defence*, Oxford: Clarendon.

—— (2008) *Rescuing Justice and Equality*, Cambridge, MA: Harvard University Press.

—— (2011) *On the Currency of Egalitarian Justice, and Other Essays in Political Philosophy*, Princeton, NJ: Princeton University Press.

Eagleton, Terry (2011) *Why Marx was Right*, New Haven, CT: Yale University Press.

Foucault, Michel (1976) *Discipline and Punish*, London: Allen Lane.

—— (1979) *The History of Sexuality*, Vol. I, London: Allen Lane.

Geras, Norman (1985) "The Controversy about Marx and Justice," *New Left Review* I/150: 47–85.

—— (1992) "Bringing Marx to Justice," *New Left Review*, I/195: 37–69.

Hardt, Michael, and Toni Negri (2000) *Empire*, Cambridge, MA: Harvard University Press.

Harman, Chris (1991) "The State and Capitalism Today," *International Socialism* 2/51: 3–54.

Harvey, David (2003) *The New Imperialism*, Oxford: Oxford University Press.

Hirsch, Joachim (1978) "The State Apparatus and Social Reproduction," in John Holloway and Sol Picciotto, eds., *The State and Capital*, London: Hodder and Stoughton.

Holloway, John (2002) *Change the World without Taking Power*, London: Pluto.

—— and Sol Picciotto, eds. (1978) *The State and Capital*, London: Hodder and Stoughton.

Jameson, Fredric (2009) *Valences of the Dialectic*, London: Verso.

Jessop, Bob (1982) *The Capitalist State*, Oxford: Martin Robertson.

—— (2002) *The Future of the Capitalist State*, Cambridge: Polity.
Kouvelakis, Stathis (2003) *Philosophy and Revolution: From Kant to Marx*, London: Verso.
Laclau, Ernesto (1977) "The Specificity of the Political," in *Politics and Ideology in Marxist Theory*, London: NLB.
Lukes, Steven (1986) *Marxism and Morality*, Oxford: Clarendon.
—— (2004) [1974] *Power: A Radical View*, Basingstoke: Palgrave Macmillan.
Martin, James, ed. (2008) *The Poulantzas Reader*, London: Verso.
Marx, Karl (1973) [1939] *Grundrisse*, Harmondsworth: Penguin.
—— (1976) [1867] *Capital*, Vol. I, Harmondsworth: Penguin.
Miliband, Ralph (1969) *The State in Capitalist Society*, London: Weidenfeld & Nicolson.
—— (1970) "The Capitalist State: Reply to Nicos Poulantzas," *New Left Review* I/59: 53–60.
—— (1973) "Poulantzas and the Capitalist State," *New Left Review* I/82: 83–92.
—— (1977) *Marxism and Politics*, Oxford: Clarendon.
—— (1983) "State Power and Class Interests," *New Left Review* I/138: 57–68.
Müller, Wolfgang, and Christel Neusüss (1975) [1970] "The Illusion of State Socialism and the Contradiction between Capital and Wage-Labour," *Telos* 25: 13–90.
Offe, Claus and Volker Ronge (1982) [1975] "Theses on the Theory of the State," in Anthony Giddens and David Held, eds., *Classes, Power, and Conflict*, Berkeley and Los Angeles: University of California Press.
Poulantzas, Nicos (1969) "The Problem of the Capitalist State," *New Left Review* I/58: 67–78.
—— (1973) [1968] *Political Power and Social Class*, London: NLB and Sheed and Ward.
—— (1974) *Fascism and Dictatorship*, London: NLB.
—— (1976a) "The Capitalist State: A Reply to Miliband and Laclau," *New Left Review* I/95: 63–83.
—— (1976b) *Crisis of the Dictatorships*, London: NLB.
—— (1978) *State, Power, Socialism*, London: Verso.
Robinson, William (2004) *A Theory of Global Capitalism*, Baltimore, MD: Johns Hopkins University Press.
Roemer, John (1982) *A General Theory of Exploitation and Class*, Cambridge, MA: Harvard University Press.
—— ed. (1986) *Analytical Marxism*, Cambridge: Cambridge University Press.
—— (1996) *Theories of Distributive Justice*, Cambridge, MA: Harvard University Press.
Rupert, Mark and Hazel Smith, eds. (2002) *Historical Materialism and Globalization*, London: Routledge.
Steedman, Ian, *et al.* (1981) *The Value Controversy*, London: Verso.
Thomas, Peter (2009) *The Gramscian Moment*, Leiden: Brill.
Wood, Allen (2004) [1981] *Karl Marx*, London: Routledge.
Wood, Ellen Meiksins (1981) "The Separation of the Economic and the Political in Capitalism," *New Left Review* I/127: 66–95.
—— (2003) Empire of Capital, London: Verso.
Žižek, Slavoj (1999) The Ticklish Subject, London: Verso.

25

FEMINISM AND THE HISTORY OF POLITICAL PHILOSOPHY

Penelope Deutscher

> by reason Opinion is free, and may pass without a pass-port,
> I took the liberty to declare my own opinions
> as other Philosophers do
>
> Margaret Cavendish, *Philosophical Letters*, Preface

Historical accounts of the nature and role of the sexes have often seemed to repeat a limited range of alternatives. Either the sexes are represented as crucially like each other in those respects deemed most relevant to their social, political or moral roles, or it will be stressed that the sexes are significantly different. An emphasis on sexual difference has sometimes produced assertions concerning women's biological, intellectual or spiritual inferiority that have proved decisive in the denial of women's rights (Tuana 1993; Le Doeuff 1990). For example, Nancy Tuana describes the long impact of the belief that woman was "less developed, less moral, less capable of rational thought, and less divine than man *because of* her role in reproduction" (Tuana 1993: xi). But the seeming alternative—the "likeness" arguments—have not always been identified as positive by feminist readers.

Thomas Laqueur has argued that before the eighteenth century women were understood physiologically, and sometimes more generally, as inverted, inferior versions of men. Eighteenth-century understandings of physiological sexual difference were new, and Laqueur places them in the context of a period of social and political change. What might have been the explanation? With the development of universalist theories of political equality perhaps women seemed only too likely to participate in the public sphere. This might provide some context for the fervent interest of the day in asserting women's inferiority by new means:

> The universalistic claims made for human liberty and equality during the Enlightenment did not inherently exclude the female half of humanity. Nature had to be searched if men were to justify their dominance of the public sphere, whose distinction from the private would increasingly come to be figured in terms of sexual difference.
>
> (Laqueur 1990: 194)

One further lesson to be drawn from the overlap of science and politics described by Tuana and Laqueur is that both sameness and difference can be associated with rhetorics of women's inferiority.

For example, consider book five of Plato's *Republic*, the first sustained argument about the possible political role of women to be offered in the history of Western philosophy.

Gregory Vlastos (1994) has discussed the possibility that Plato's argument could be identified as proto-feminist. It offers grounds for the plausibility of women's participation in public and political life. Plato proposes that sex is as irrelevant as baldness in judging aptitude. According to a continuity thesis this might be considered the first of a number of historical arguments to have emphasized sexual equality.

Yet Tuana (1993) and Annas (1996) have argued that scrutinizing the epistemological and metaphysical context of Plato's claims about women's capacities reveals a thinker more importantly *unlike* contemporary defenders of women's equality. The *Republic* describes natural differences in quality and capacity, suggesting that humans can be distinguished as can types of metals: bronze, silver or gold. In the view of Elizabeth Spelman (1988) also, because Plato reveals himself to be committed to a doctrine of natural inequality there is no affinity with later trends in equality arguments.

Spelman directs our attention to the metaphysical foundation for Plato's views. Sex may be irrelevant in judging aptitude, but:

> Paradoxically, [Plato's] argument rests on a metaphysical argument that establishes *inequality* in two ways: (1) People have different natures and in light of that ought to be educated differently to play different roles in the polis (as opposed to being given the opportunity to do anything they want) … [T]he very same arguments meant to establish the equality of some women to some men are also meant to establish or reflect the inequality of some groups of women and men to other groups of women and men.
>
> (Spelman 1988: 25)

By contrast, later and contemporary arguments for the equality of the sexes do not claim that different human qualities reflect an essential difference of one's natural type. Some of the most enduring claims to women's equality, such as those of Harriet Taylor and John Stuart Mill in 1851 and 1869, respectively, and Beauvoir in 1949, are based on their rejection of such a view.

1. The Sexual and the Racial Contract

In turning to feminist readings of the contract tradition, one similarity with the feminist readings of Plato is seen in the preoccupation with how to interpret apparent contradictions concerning the role of women within the texts of the history of philosophy. In *The Sexual Contract* (1988), one of the most influential twentieth-century works in feminist political philosophy, Carole Pateman argues that one does not accurately understand the social contract tradition without identifying its presupposition of an invisible sexual contract: "men's freedom and women's subjection are created through the original contract" (1988: 2). Pateman notes the resulting flaws in the arguments of Locke and Hobbes, for example. In Locke's case, the marriage bond is described as occurring in the state of nature. He does not view the woman as giving up a natural liberty in favor of subordination to her husband (Pateman 1988: 6). Like Locke, Hobbes also depicts women as

subjugated to men in society. Those who form the social contract are described as male heads of households. But women are not subordinate to men in the Hobbesian state of nature. Moreover, he ascribes to women a natural right and lordship over their children in the state of nature, and does not describe the marriage contract as occurring in the natural condition.

In a second component of Pateman's method she reworks the resources of Hobbes and Locke to provide the missing explanations for the subjugation of women. For example, Hobbes makes a point of claiming that in the state of nature adults of both sexes have sufficient wits to kill each other. Women's subordination would not arise from physical inferiority in such contests. So, asks Pateman, how can we understand the fact that when men form the social contract women are already subordinate to them, their rights over their children transferred to their husbands? Hobbes does acknowledge that a master–servant relationship can arise in the state of nature since there are those who "with the conqueror's sword at [their] breast" may prefer subservience to death. Perhaps women's care for their children could be assumed on this account to place them at a disadvantage (or alter their preferences) in the state of nature. Despite the original equality of women with men attributed to this state, maternity might still lead them to that servant-like status presupposed at the point when the male "heads of households" erect a commonwealth.

By contrast, Locke attributes to men a natural sexual right over women, manifest both in the state of nature and in political society. In this sense, his argument is more consistent. Nonetheless, this natural and original subordination of women is not easy to explain in the light of his account. For one thing, this "first Society," as it is described, is said to be a "voluntary Compact" between a man and his wife (Locke, *Second Treatise on Government*, §§77–78, cited in Pateman 1988: 93). Moreover, Locke argues that in the state of nature, "all the Power and Jurisdiction is reciprocal" (Pateman 1988: 53), and that each of us has property in our own person. Furthermore, sons are described as able eventually to withdraw their original subservience to their fathers—why then, Pateman asks, should not women be conceived as able to withdraw their subservience to fathers and husbands?

Some philosophers continued dialogue with Pateman's project by suggesting it had its own oversights. While she alerts us to unexplained sex subordination in the contract tradition, Charles Mills has commented that "Pateman's original 'sexual contract' had little to say about race" (Mills 2007: 165). This was one motivation for his answering work, *The Racial Contract* (1997), which emphasizes that political compacts also occur (theoretically and historically) in colonial or postcolonial contexts between individuals overtly or tacitly supposed to be European and/or are grounded in the prior or concurrent appropriation of the land and work of nonwhites. As Mills developed this analysis, however, he acknowledges that another anomaly arose: "my 'racial contract' had little to say about gender" (Mills 1997: 165). Pateman and Mills have more recently undertaken a collaborative project, *Contract and Domination* (Mills and Pateman 2007), which integrates their work on contract theory from the perspective of gender, race, class and cultural difference. This new project is sympathetic to the perspective of contemporary intersectionality theory—see, for example, Patricia Collins (1998) and Kimberlé Crenshaw (1989). According to this approach, depictions of sex and gender must be understood in the context of their concurrent intersections with race and class. From this perspective an analysis would usually be considered analytically false if it attempted to abstract gender from race and class, or race from gender and class, and so on.

2. Jean-Jacques Rousseau (1712–78) and Mary Wollstonecraft (1759–97)

Within the social contract tradition, Rousseau differentiates himself by having tackled more directly the role of women in political communities and undertaking an elaborate argument defending women's subordination as socially and politically crucial.

The argument is noteworthy for its indirect admissions. Rousseau offers a strident depiction of the perils that will result from women breaking with their traditional role as wives and mothers. His impassioned argumentation undercuts itself with its seeming acknowledgment that radical social transformation is indeed an imminent threat. In the eighteenth century, Rousseau recognized that women were challenging their traditional roles—writing, performing on the stage, holding salons, sometimes attaining roles as effective and influential public figures.

Yet, according to Rousseau, women's role in the home was not incidental, but critical, to the good functioning of political society. There were at least three aspects to this view. (a) Men ran the risk of becoming estranged from nature by their participation in the social and political sphere. Proximity to women in the home (representing the dictates of nature) helped to avert this risk, by ensuring an important proximity to a more natural register (b) Women were responsible for nurturing future citizens and their education ought to equip them for this vital maternal role, and (c) an additional argument was that although a woman was inferior to a man in his domain, she was superior in her own. The sexes were not naturally the "same" or "equal," they were different in the sense of "complementary." This is one of the classic arguments for subordination: that it is a mutually beneficial complementarity. The claim is also that women are not really subordinated if domestic and maternal responsibilities are sufficiently valued.

Now thought of as one of the earliest feminist political philosophers, Mary Wollstonecraft was an important interlocutor of Rousseau and also of Burke. She offers another instance of how feminist interpretation of political philosophy has been attuned to the incoherence (in addition to the injustice or inaccuracy) of its depiction of women.

In *A Vindication of the Rights of Woman* (1792), Wollstonecraft accepted the ideal of a well-ordered society that develops the moral and intellectual abilities of its individuals. She did not disagree with Rousseau's view that women should receive such education as would best equip them, as mothers, to cultivate the qualities of young future citizens. From this perspective, however, she countered that the education of women and men should be identical. Rousseau's own arguments should have led him to this conclusion:

> How should a woman void of reflection be capable of educating her children? How should she discern what is proper for them? How should she incline them to those virtues she is unacquainted with, or to that merit of which she has no idea?
>
> (Wollstonecraft 1995: 169)

Wollstonecraft also identified much in the restricted education proposed by Rousseau that would inhibit a woman from being the good mother and helpmeet—rather she would be coquettish, artificial and deceptive, "beautiful, innocent, and silly" (1995: 169). Though she was not specific about the civil, political rights, or the political participation women ought to enjoy, she decried the fact that women were "groping in the dark," "immured in their families," "den[ied] civil and political rights" (1995: 69), and argued that they ought to have some direct form of political representation and some "direct share allowed them in the deliberations of government" (1995: 237).

3. Susan Moller Okin, Justice and the Family

Rousseau and Wollstonecraft shared an interest in asking what family structures are most appropriate to the ideal functioning of the social and political sphere. They are distinctive in this respect; indeed Okin (1979, 1989) has criticized the political philosophy tradition for repeatedly ignoring the role of the family, which leads to two lines of interrogation. We should assess the history of political philosophy in terms of the extent of its willingness to acknowledge that morally responsible agents cognizant of the principles of justice do not spring up fully formed. Family arrangements, providing their formative conditions, should be a more important theme in political philosophy. The second line of interrogation again concerns the inconsistency of argument evident in its history and traditions when it comes to the depiction of women's role. As Okin argues, philosophers such as Rousseau who do thematize the role of the family may not suppose that it should be organized according to the principles of justice, or that there should be a congruence between the latter and the former.

Like those of Pateman, Okin's re-readings of the history of political philosophy have led to interrogations of present-day suppositions about contractarianism, and about justice, gender and the family. Ruth Sample (2002) provides an overview of feminist critiques of the contract tradition, contemporary contractarianism, and the blindspots in theories of justice with respect to gender and family. Making a case for a contemporary, feminist-friendly contractarianism, Sample acknowledges many objections to the tradition including that it inadequately accommodates participants understood as non-reciprocally dependent. While some have argued that the principles of justice ought to be applied to family arrangements (Ruddick 1995), others have argued that the base tenets of the contract tradition may be inappropriate to non-reciprocal relationships of care and dependency (Nussbaum 2006). Reviewing the debate between feminists favoring so-called "justice" or "care" approaches to these issues, Sample acknowledges that some see these as competing registers, whereas for others they require integration (Sample 2002).

4. The Unconventional History of Western Philosophy

If one thinks back to the opening lines of Margaret Cavendish's *Philosophical Letters* (1664) one notes the gusto with which she espouses reason for its capacity to give passage even to women's opinions: "by reason Opinion is free, and may pass without a passport"—thus she considers herself authorized to "set down several famous and learned Authors opinions" for the purposes of critical debate. Mary Warren is one of a number of feminist philosophers (including Moira Gatens (1991) and Michèle Le Doeuff (1990)) who have reflected on women's contributions to the history of philosophy and diagnosed types of relationships taken up by women philosophers to the existing corpora of prominent male philosophers. Warren (2009) affirms that recognizing the historical contributions of women philosophers leads one to an interestingly "*Unconventional History of Western Philosophy*," a history of women philosophers' dialogues with famous interlocutors: late Pythagorean women philosophers with Aristotle, Hildegard of Bingen with Augustine, Heloise with Abelard, Princess Elisabeth with Descartes, Catharine Macaulay with Hobbes, Lady Masham with Locke, Anne Conway with Leibniz, Wollstonecraft with Rousseau, Anna Maria van Schurman with Kant, Arendt with Heidegger, Jane Addams with John Dewey, Beauvoir with Sartre.

True, most philosophical production has taken the route of dialogue with peers and forebears; one can also tell the history of Western philosophy in terms of Aristotle's responses to Plato, Kant's responses to Descartes and Hume, Sartre to Heidegger and so on. But Le Doeuff (1990) has argued that one of the ways in which historical philosophical dialogue is gendered concerns the sociology of knowledge. Until the twentieth century, the most likely means for intellectual women to access philosophical dialogue was via intimate relationships, such as those possible with lovers, fathers, partners, in household or domestic situations, or by the route of private correspondence. Some religious women could appropriately engage the writings of the fathers of the Church, some wealthy women were benefactors to philosophers, corresponded with or employed them as private tutors. Philosophical production by women was more likely to be over-determined by personal relations and the phenomenon named by Le Doeuff the "Abelard-Heloise complex." In consequence, distinctive dialogues arising in philosophical "duos" have often been a particularly important form for women's contributions to the history of philosophy.

But Cavendish's claiming the right to rebut the "famous and learned Authors opinions" is also characteristic of a distinctive—sometimes exhilarated—tone of self-authorization in some historical women thinkers. Wollstonecraft's later rebuttals of Burke and Rousseau provide an excellent example of this claim to the authority of reason by women of letters—sometimes aristocratic, sometimes of modest means and limited educational advantages, living by their writing—who revisited the arguments of revered male contemporaries in their philosophical production.

Wollstonecraft brings this characteristic vigor to her identification of Rousseau's weak argumentation concerning women's role and nature. But does this exempt her from reproducing the problems in his argumentation? Gatens (1991) suggests not. The very criticisms of many women philosophers may have underestimated "the depth of sexual bias in philosophy" (1991: 24) where they adopted some elements of the philosophical works with which they were in dialogue while aiming to eliminate those elements hostile to women. Consider the differences between Rousseau's and Wollstonecraft's understandings of human reason and character. For Rousseau, the sexes are profoundly differentiated in their rational capacity; Wollstonecraft disagrees. But Wollstonecraft does agree with Rousseau's recommendations for the good organization of civic society, arguing that his account of the appropriate education for the emergent citizen should be extended to girls. She discounts important components of Rousseau's argument: his view that the citizen requires, to offset the denaturing aspects of political and public life, an ongoing proximity with nature ensured on his model by restricting the role of women to the natural domestic hearth to which the man returns (Gatens 1991: 11).

Wollstonecraft isolates from this context the argument that the sexes require a different education, rejects it, and extends Rousseau's educational dicta for men to women. She speaks to women's potential to be at once "good wives and mothers," "rational creatures," and "free citizens" (Wollstonecraft 1995: 275). The plea is founded on her version of perfectibility: women could, with equal education, be like men in intellectual ability and moral qualities. But Rousseau's account of the necessary division of the sexes is premised on his description of a conflict between the natural and political demands on humans, and his view of women's role and familial life as the therapeutic solution to this conflict.

One conclusion is that Wollstonecraft might not have gone far enough when she undertook to offer an egalitarian variation of Rousseau's account of the male citizen (Gatens 1991: 24). Certainly, readers confronting sex (and race) bias in the history of

social and political philosophy have sometimes assumed that these are regrettable components which can be excised, leaving the fundamentals intact and ready to be espoused by the feminist philosopher. Gatens recommends a greater wariness with respect to this "cut and paste" approach.

5. Arguments from Justice: Harriet Taylor and John Stuart Mill

One strategy shared by Wollstonecraft in the eighteenth century and John Stuart Mill and Harriet Taylor in the nineteenth century is their criticism of those who used women's inferiority to justify the denial of rights. Even if women were inferior in some respect—badly educated, for example—surely this demonstrated the importance of their access to better education. Such responses also developed the claim that social environments were capable of distorting significantly human qualities: women would be very different beings if offered different opportunities, education and upbringings. This connected to the claim, espoused by all three, that humans have a right to the circumstances in which capacities such as reason, sense, taste, judgment, benevolence, decency and morality, can develop.

Wollstonecraft, Mill and Taylor all refer to a number of outstanding individals—such as female rulers, actors, and women of letters—to highlight women's capacity for exceptional accomplishments given the proper opportunities. Mill called the bluff of those who considered women inferior for many walks of life. Why not allow them access to adequate training given that any such inferiorities would reveal themselves of their own accord? Why were laws and constraints on such participation needed, any more than laws were needed to restrict a blacksmith's profession to the physically strong? But, important as it was for Mill and Taylor to demonstrate inconsistency in argumentation, Mill's primary argument was also that the legal subordination of women was "wrong in itself" (1989: 119).

Taylor and Mill were intellectual collaborators as well as romantic partners, and Taylor's views are rarely distinguished from those of Mill. Certainly they agreed on most fundamental points—they claimed for women the right to equal education, work and political participation and franchise, as fundamental principles of justice. But Le Doeuff has noted some important differences between their arguments, which become apparent if one compares Taylor's 1851 work *The Enfranchisement of Women* with Mill's *Subjection of Women* of 1869. Although one of his primary claims was the "equal moral right of all human beings to choose their occupation" (Mill 1989: 168), Mill assumed that while unmarried women should have viable work opportunities, married women would be primarily wives and mothers, responsible for child rearing and for the economic management of the household. Additional tasks would add unduly to these duties and inhibit women from performing them well. By contrast, Taylor claimed that paid work opportunities were important for all women, married and unmarried. She rebuts the argument that this will limit men's employment opportunities: at worst "the joint income of the two would be the same as before, while the woman would be raised from the position of a servant to that of a partner" (1998: 96). She goes on to stress how important a protection it can be against a woman "treated in the same contemptuously tyrannical manner" that part of the household income "should be of the woman's earning." She opposed legal marriage more strongly than Mill, though her radicalism on this point was not made clear in her 1851 publication. Another difference is seen in Mill's critical engagement with nineteenth-century debates concerning the biological inferiority of women.

6. Disputing the "Natural": The "Biological Facts" as Contestable

Though no specific theorists are named in *Subjection*, in that work and in his correspondence with Herbert Spencer and August Comte, Mill disputed the claims of his day concerning men's physiological superiority, including the claim that men had larger brains. Mill took such references to physiological and biological difference to be easily answered. If larger brain size was an indication of greater intelligence, why then were elephants not more intelligent than men? Yet, while he rebuts what was proffered as biological evidence, he gives it sufficient credence to take on the challenge of its reinterpretation. A quite different variant of this phenomenon was later seen in the work of twentieth-century feminist philosopher Simone de Beauvoir (1949). Beauvoir rejects causal biological determinism in a chapter of *The Second Sex* which acknowledges many alleged biological facts about sexual difference while developing a radical and existentialist concept of freedom according to which such facts could have no real significance.

By the time of Mill's *Subjection of Women* in 1869 and then again from the 1940s onwards it was common for feminist philosophers to engage with contemporary biological claims about sex difference, with the aim of finding the means to repudiate or reinterpret them. Putative biological data about sex has often been seen as the enemy of the feminist claimant, and there are certainly many instances of biological data (and pseudo-data) being manipulated to serve anti-feminist purposes. By contrast, a number of contemporary trends in feminist theory are noteworthy for their positive engagement with biology and with new directions in the sciences, forming an emergent field sometimes referred to as the new materialist feminisms (Alaimo and Hekman 2008; Coole and Frost 2010).

7. Revisiting the Status of Sexual Difference

Wollstonecraft, Mill, Taylor and Beauvoir also all belong to a tradition in feminism suspicious of praise for sexual difference. Whether laudatory, as in the case of Rousseau, or depreciative, as in the biological or pseudo-scientific accounts rejected by Mill, sexual difference served as the alibi for a limitation on women's opportunities.

By contrast, as Karen Offen (2000) has noted, two periods, the end of the nineteenth century through the early twentieth century, and also the 1980s, saw the advent of a number of feminist methodologies emphasizing sexual difference as a positive means of articulating women's rights claims. The late nineteenth- and early twentieth-century variations saw many feminists in Germany, Switzerland, the United States and Britain, stressing women's reproductive and maternal role as a means of emphasizing their unique character, the specificity of their social contributions—including their literal responsibility for reproduction and their supposed capacity to bring maternal or cooperative or otherwise "feminine" qualities to the public sphere—as the special basis for their claim to political rights.

In the 1980s, a renewed, and less maternally oriented interest in the heuristic of sexual difference was seen in a number of writers associated with the "new French feminisms" (Marks and de Courtivron 1987). One of these, Luce Irigaray, is also associated with an alternative feminist methodology for interpreting the history of social, political and philosophical thought, one with affiliations (and important disputes) with twentieth-century psychoanalytic and deconstructive theory. In her critical analyses of figures

such as Plato, Spinoza, Descartes and Hegel and many other historical figures, Irigaray (1985, 1991, 1993) criticized the history of philosophy for the inadequacy of its representation of femininity and sexual difference. She attempts to cultivate a new and more adequate "imaginary" for sexual difference, to generate a richer pantheon of representational possibilities for the sexes.

8. New Methodologies for Interpreting the History of Philosophy

An important dissensus therefore exists amongst some of the most innovative contemporary feminist approaches to the history of philosophy. The debates turn not only on how best to interpret its representations of sameness and difference, but also on the possible importance of metaphor, symbolism and imagery in the history of philosophy when it is assessed from a gendered perspective. Thus, a feminist approach to the history of philosophy arguably would pay particular attention not only to some of the sociological aspects raised by Michèle Le Doeuff, but also its "literary," in addition to its metaphysical, epistemological and scientific aspects.

Disagreeing with Luce Irigaray, Le Doeuff's position (1990) does not consider the appropriate response to historically subordinating representations of sexual difference to be new, innovative and more positive representations of sexual difference. Nonetheless, Le Doeuff does share Irigaray's focus on the role of language and literary elements (such as metaphor and symbol) in the subordination of women in the history of philosophy. This focus led to Le Doeuff's development of the concept of the "philosophical imaginary" as one means of understanding the role of women and the feminine. Where we have seen a number of historical feminists inclined to identify contradictory or incoherent thought in the representation of women, Le Doeuff—not disagreeing—has argued that philosophy typically turns to components of metaphor and imagery to finesse problematic aspects in argumentation. There is, in Le Doeuff's view, a possible correlation between the tendency to deviate towards incoherence when depicting women and gender, and the tendency to employ imagistic or imaginative argumentative elements in such depictions. To have an interest in the former, may, methodologically, lead to special expertise in the role of the latter in the history of philosophy.

Genevieve Lloyd has also addressed the status of metaphor in the history of philosophy. One of the distinctive elements of her *The Man of Reason* (1993) is its feminist inflected interpretation of a number of philosophers—Descartes is one of the better known examples—in whose texts one does *not* find disparaging remarks about women. *The Man of Reason* more broadly interprets the metaphorical associations of reason and its opposites. Reason might not overtly be associated with masculinity, but it might be represented as transcending sensibility, unreason, unregulated thought, the emotions and the passions. Thus it can, indirectly, be distanced from what was associated historically with femininity even in philosophical contexts that lack overt gender connotations.

9. Recuperative Readings

By what criteria, then, might one most appropriately assess the resources of historical philosophical texts from a feminist perspective? Comparing feminist interpretation of figures such as Descartes and Spinoza might lead one to the conclusion that the philosopher's actual remarks about women's ability might not be the most important consideration.

Spinoza is one of the many philosophers who argued that women's capacities were inferior to those of men. On this basis, he excluded women from a role in political life (Lloyd 2009: 37). His justification, as Gatens explains, "is that women, slaves and servants are not independent beings but rather are under the authority of men and masters … women are under men's authority by nature since 'women do not naturally possess equal right with men'" (Gatens 2009: 204). On the other hand, Gatens notes that the view is odd and inconsistent because Spinoza rejected the ascription of qualities to whole categories or classes of persons. But she asks (as a methodological point), is locating such inconsistency a sufficiently productive task? Not, in her view, if it takes place only under the sign of rebuke. She refers to her agreement with Michael Rosenthal who suggests:

> the task of contemporary theorists is not to censure the writings of past philosophers but to see that "the very tensions in these texts call upon us to resolve them in our philosophical endeavours."
>
> (Gatens 2009: 203, citing Rosenthal 2005)

Notwithstanding his remarks about women's inferior capacities, some recent feminist interpretations have instead turned to Spinoza for a promising historical alternative to models that conceived of reason on a model of sexless minds housed in sexed bodies (Lloyd 2009: 39). Lloyd sees potential in the one substance view, according to which our minds reflect (rather than being independent of) the ways in which our bodies are socialized (2009: 38). The new approaches to the interpretation of Spinoza proposed by Gatens, Lloyd, and Susan James (1997) belong to an alternative tradition within feminist responses to the history of philosophy, for which the most salient question to direct at a historical model concerns the philosophical resources it offers feminist projects, sometimes very broadly construed.

Many recent feminist interpretations of Kant offer further good examples of this methodological divergence between problematizing sex-biased content, and looking to past philosophical frameworks for the most promising theoretical resources for ongoing developments in feminist philosophy. Like *Feminist Readings of Aristotle* (Freeland 1998), *Feminist Readings of Immanuel Kant* (Schott 1997) similarly bears witness to this methodological divide, though she queries it: "today feminists have increasingly come to the recognition that our task is to be both critical and constructive" (Schott 1997: 4).

Though these readings are sometimes identified as a very recent development in feminist philosophy, many writing in the feminist-Marxist tradition asked how his analysis of class subordination could be integrated with the analysis of women's subordination. Marx did not suppose that women are naturally subordinate to men. Yet, he failed to integrate the role of women's responsibility for procreation and, historically, childrearing and housework into his analysis. The feminist reworking of the Marxist tradition which is premised on the integration of such analyses is sometimes known as dual systems theory (Hartmann 1981; Young 1981).

It is true, however, that some recent feminist approaches could be described as going beyond critique, recuperation, and integration. New prisms have been increasingly offered through which to read a historical philosopher reconstructively, sometimes experimenting with strong reconfigurations of philosophical texts to this end. The view that moral philosophy should maximally consider the role of sentiment in moral agency

might provide new grounds to reconfigure Kant so as to foreground such a perspective. Some feminist readings of Hegel have found grounds to favor, particularly against a Kantian model, a subjectivity which is figured by Hegel as situated and intersubjective. Compare this to the arguably more traditional feminist approach which might castigate Hegel for the subordinate role he attributed to women. In fact, Seyla Benhabib has argued that the essentialized depiction of women could be identified as aberrant in Hegel's work, given that a Hegelian model can provide resources for understanding sexual difference as situated, social and constructed (Benhabib 1996). Hutchings (2003) and Alison Stone (2006) have similarly revalued Hegel and Hegelianism from a feminist perspective.

10. Feminist Interpretations of Nietzsche

Because of the extreme hostility of some of his remarks about women, perhaps no starker example of the importance for many feminist commentators of distinguishing between a philosopher's actual depictions of sex and gender, and asking how his philosophy might be fruitful for feminist purposes, has been offered than by the work of Friedrich Nietzsche. From its earliest reception, feminist readers of Nietzsche—such as the late nineteenth- and early twentieth-century Swiss and German feminists Malwida von Meysenbug, Helene von Druskowitz, Hedwig Dohm, and Helene Stöcker—lamented his misogyny but stressed the transformative power of his work (Diethe 1996). Contemporary political philosophers including Wendy Brown (1995) and Rosalyn Diprose (1989) have, from a feminist standpoint, used Nietzsche to challenge some tenets of feminist philosophy including its historical commitment to liberalism. In a direction akin to Lloyd's and Gatens's reading of Spinoza, Diprose finds promising in Nietzsche the role ascribed to embodiment. For Nietzsche the "mind is an idea of the body," mediated by systems of evaluation. A Nietzschean perspective associates the aversion to difference with the "herd-like" or reactive (Diprose 1989: 30–1), and could have the potential, Diprose argues, to affirm new creations of sexual difference. Indeed, Nietzsche's own resources could allow the reader to criticize his own more vitriolic comments about women as those of "*ressentiment*." Brown and Halley (2006) have deployed the Nietzschean critique of *ressentiment* in the powerful cases they make for a post-*ressentiment* and post-reactive mode of feminism.

Many of the feminist methodologies for the interpretation of the history of social and political philosophy reviewed in this chapter have been developed with the assumption that, whether reconstructive or deconstructive, recuperative, or critical, they were to be directed at the canonical male figures of this history: Plato, Aristotle, Augustine and Aquinas, Locke, Hobbes and Rousseau, Kant, Hegel, Nietzsche, Heidegger, Sartre, Merleau-Ponty. But in the wake of recuperative projects ranging from Mary Waithe's four-volume *History of Women Philosophers* (1987–1994) to Karen Warren's *Unconventional History* (2009) the historical philosophical canon has been expanded to include many previously neglected women writers in the history of social and political thought. This offers a new question, and perhaps a new opportunity to the feminist interpretative traditions that have long engaged in developing innovative methodologies for reinterpreting figures from Plato to Heidegger: can these same methodologies now be directed at the history of women philosophers?

Related Topics

Plato's Political Philosophy, Hobbes, Locke, Rousseau, Kant, Hegel, Mill, Late Nineteenth- and Early Twentieth-Century British Thought, Liberalism, Marxism and Contemporary Political Thought, Contractarianism, Contractualism and Political Liberalism, Utilitarianism and Consequentialism, Natural Law and Rights Theory, Multiculturalism, Equality, Rights, Marriage, Sex and the Family

References

Alaimo, Stacey, and Hekman, Susan, eds. (2008) *Material Feminisms*, Bloomington: Indiana University Press.

Annas, Julia (1996) "Plato's Republic and Feminism," in *Feminism and Ancient Philosophy*, ed. Julie K. Ward, New York and London: Routledge.

Beauvoir, Simone de (2010 [1949]) *The Second Sex*, trans. Constance Borde and Sheila Malovany-Chevallier, New York: Alfred A. Knopf.

Benhabib, Seyla (1996) "On Hegel, Women, and Irony," in *Feminist Readings of G.W.F. Hegel*, ed. Patricia Mills, University Park: Pennsylvania State University Press.

Brown, Wendy (1995) *States of Injury: Power and Freedom in Late Modernity*, Princeton, NJ: Princeton University Press.

Cavendish, Margaret (1664) *Philosophical Letters, or, Modest Reflections Upon Some Opinions in Natural Philosophy Maintained by Several Famous and Learned Authors of This Age, Expressed by Way of Letters*, EEBO Editions.

Collins, Patricia Hill (1998) in *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment*, New York: Routledge.

Coole, Diana and Frost, Samantha (2010) *New Materialisms: Ontology, Agency, and Politics*, Durham, NC: Duke University Press.

Crenshaw, Kimberlé (1989) "Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics," *The University of Chicago Legal Forum*: Feminism in the Law: Theory, Practice and Criticism: 139–68.

Diethe, Carol (1996) *Nietzsche's Women: Beyond the Whip*, Berlin and New York: Walter De Gruyter.

Diprose, Rosalyn (1989) "Nietzsche, Ethics and Sexual Difference," *Radical Philosophy* 52: 27–33.

Freeland, Cynthia, ed. (1998) *Feminist Interpretations of Aristotle*, University Park, PA: Pennsylvania State University Press.

Gatens, Moira (1991) *Feminism and Philosophy: Perspectives on Difference and Equality*. Bloomington, IN: Indiana University Press.

——, ed. (2009) *Feminist Interpretations of Benedict Spinoza*, University Park: Pennsylvania State University Press.

Halley, Janet (2006) *Split Decisions: How and Why to Take a Break from Feminism*, Princeton, NJ: Princeton University Press.

Hartmann, Heidi (1981) "The Unhappy Marriage of Marxism and Feminism: Towards a More Progressive Union," in *The Unhappy Marriage of Marxism and Feminism*, ed. Lydia Sargent, London: Pluto.

Hutchings, Kimberly (2003) *Hegel and Feminist Philosophy*, Cambridge, MA and Oxford, UK: Polity.

Irigaray, Luce (1985) *Speculum of the Other Woman*, trans. Gillian C. Gill, Ithaca, NY: Columbia University Press.

—— (1991) *Marine Lover-of Friedrich Nietzsche*, trans. Gillian C. Gill, Ithaca, NY: Cornell University Press.

—— (1993) *An Ethics of Sexual Difference*, trans. Carolyn Burke and Gillian C. Gill, Ithaca, NY: Cornell University Press.

James, Susan (1997) *Passion and Action: The Emotions in Seventeenth-Century Philosophy*, Oxford: Oxford University Press.

Laqueur, Thomas (1990) *Making Sex: Body and Gender from the Greeks to Freud*, Cambridge, MA: Cambridge University Press.

Le Doeuff, Michèle (1990) "Long Hair, Short Ideas," in *The Philosophical Imaginary*, Stanford, CA: Stanford University Press.

Lloyd, Genevieve (1993) *Man of Reason: "Male" and "Female" in Western Philosophy*, rev. edn. Minneapolis: University of Minnesota Press.

—— (2009) "Dominance and Difference: A Spinozist Alternative to the Distinction Between 'Sex' And 'Gender'," in *Feminist Interpretations of Benedict Spinoza*, ed. Moira Gatens, University Park: Pennsylvania State University Press.
Marks, Elaine, and Isabelle de Courtivron, eds. (1987) *New French Feminisms: An Anthology*, New York: Pantheon.
Mill, John Stuart (1989) "The Subjection of Women" [1869], in *"On Liberty" And Other Writings*, Cambridge: Cambridge University Press.
Mills, Charles (1997) *The Racial Contract*, Ithaca, NY: Cornell University Press.
—— and Carole Pateman (2007) *Contract and Domination*, Cambridge: Polity.
Nussbaum, Martha (2006) *Frontiers of Justice: Disability, Nationality, Species Membership*, Cambridge, MA: The Belknap Press, Harvard University Press.
Offen, Karen (2000) *European Feminisms, 1700 1950: A Political History*, Stanford, CA: Stanford University Press.
Okin, Susan Moller (1979) *Women in Western Political Thought*, Princeton, NJ: Princeton University Press.
—— (1989) *Justice, Gender, and the Family*, New York: Basic Books/HarperCollins.
Pateman, Carole (1988) *The Sexual Contract*, Stanford, CA: Stanford University Press.
Rosenthal, Michael (2005) "The Black, Scabby Brazilian: Some Thoughts on Race and Early Modern Philosopbhy," *Philosophy and Social Criticism* 31, no. 2: 211–21.
Ruddick, Sara (1995) "Injustice in Families: Assault and Domination," in *Justice and Care: Essential Readings in Feminist Ethics*, ed. Virginia Held, Boulder, CO: Westview.
Sample, Ruth (2002) "Why Feminist Contractarianism?" *Journal of Social Philosophy* 33, no. 2: 257–81.
Schott, Robin, ed. (1997) *Feminist Readings of Immanuel Kant*, University Park, PA: Pennsylvania State University Press.
Spelman, Elizabeth (1988) *Inessential Woman: Problems of Exclusion in Feminist Thought*, Boston, MA: Beacon Press.
Stone, Alison (2006) *Luce Irigaray and the Philosophy of Sexual Difference*, Cambridge: Cambridge University Press.
Taylor Mill, Harriet (1998) "The Enfranchisement of Women," [1851] in *The Complete Works of Harriet Taylor Mill*, Bloomington: Indiana University Press.
Tuana, Nancy (1993) *The Less Noble Sex: Scientific, Religious, and Philosophical Conceptions of Woman's Nature*, Bloomington: Indiana University Press.
Vlastos, Gregory (1994) "Was Plato a Feminist?" in *Feminist Interpretations of Plato*, ed. Nancy Tuana, University Park: Pennsylvania State University Press.
Waithe, Mary Ellen (1987–94) *A History of Women Philosophers*, volumes 1–4, Dordrecht: Kluwer.
Warren, Karen J., ed. (2009) *An Unconventional History of Western Philosophy: Conversations between Men and Women Philosophers*, Lanham, MD: Rowman and Littlefield.
Wollstonecraft, Mary (1995) *A Vindication of the Rights of Men and a Vindication of the Rights of Woman and Hints (Cambridge Texts in the History of Political Thought)*, ed. Sylvana Tomaselli, Cambridge: Cambridge University Press.
Young, Iris Marion (1981) "Beyond the Unhappy Marriage: A Critique of the Dual Systems Theory," in *Women and Revolution: A Discussion of the Unhappy Marriage of Marxism and Feminism*, ed. Lydia Sargent, Cambridge, MA: South End Press.

26
ENVIRONMENTALISM

Mathew Humphrey

1. Introduction

How should human beings relate to the non-human environment? Whilst other political ideologies may provide answers to this question, environmentalism is marked out by the centrality of the human/non-human relationship to its values and policy proposals. The short answer, on one interpretation at least, is that this relationship should be a *sustainable* one, but what should be sustained, for whom, and over what period of time are matters over which environmentalists can disagree. This chapter will begin with a very brief and inevitably incomplete account of the history of environmental thought, in order to sketch some of the main roots of contemporary environmentalism. We will then review some of the key forms that environmental political thought takes, and examine how different environmental perspectives answer the "what is to be sustained" question. The final section will offer an assessment of the current state of environmental political philosophy, in the wake of recent work by "skeptical" environmentalists, and in the face of suggestions that environmentalism is (or should be) "dead," and that we have moved to a "post-ecologist" age.

Before moving on to the historical story, a brief note on terminology is worth making. There is a common distinction between "environmental" politics and "ecological" politics (see Dobson 2007), which maps roughly onto Arne Naess's (1973) divide between "deep" and "shallow" ecology (see below). According to this distinction "environmentalists" seek to ameliorate environmental problems, addressing key concerns such as climate change and biodiversity loss as discrete problems within the existing industrial-capitalist system. "Ecologists" on the other hand, see that system as an integral part of the cause of our environmental crisis, and seek a radical reconstruction of polity, economy, and society. Calls for deindustrialization, radical decentralization, and participatory forms of democracy emanate from this more radical perspective. In this chapter I will discuss both views, and differentiate where necessary between "environmental" and "ecological" perspectives. Inevitably, however, the boundary between the two is rather fuzzy, and readers should be aware that this distinction is less common in the American literature on environmental politics than it is in the European.

2. Environmentalism in Historical Perspective

Environmentalism is often seen as a recent form of politics, associated with the new social movements that emerged in the 1960s alongside a transition towards "post-material" values (Inglehart 1977). As we shall see there are valid reasons for taking

this view, but this should not disguise the far deeper historical roots to which modern environmentalism is connected. Human beings have long reflected upon their relationship to the natural world, even if this matter has not always been central to political concerns. Although the predominant forms of this thinking were often of a sort that contemporary environmentalists would consider destructive, there have been expressions of concern about local environmental conditions going back to antiquity (Plato complained about deforestation), and particular thinkers who have suggested that the whole human–nature relationship needs to be put upon a new footing. One example of early environmental concern is John Evelyn's complaint about London's air in his *Fumifugium* of 1661:

> It is this horrid Smoake which obscures our churches, and makes our Palaces look old, which fouls our Clothes, and corrupts the Waters, so as the very Rain, and refreshing Dews which fall in the several Seasons, precipitate this impure vapour, which, with its black and tenacious quality, spots and contaminates whatever is exposed to it.
>
> (quoted in Wall 1994: 46)

The size of the human population and the ability of the world to feed it is another long-standing environmental concern, and a famous early text in this regard is "An Essay on the Principle of Population," by the English cleric Thomas Malthus, first published in 1798. Malthus's argument that population increased exponentially (1, 2, 4, 8, 16, 32, 64 etc.), whilst food production could increase only "arithmetically" (1, 2, 3, 4, etc.) thus leading to viciousness and famine unless population was consciously controlled, led to instant notoriety and controversy. "Parson Malthus" was ridiculed by Marx and Engels, and in *Capital* Marx describes the first edition of the *Essay* as "schoolboyish, superficial plagiary." The epithet "Malthusian" is still today attached to works on human population, and it is not normally intended as a compliment.

In the nineteenth century we see the rise of a tradition of environmental thinking in the United States associated with the "transcendentalism" of Ralph Waldo Emerson and Henry David Thoreau, which influenced later campaigners for the natural world such as John Muir. Transcendentalists saw the natural world pantheistically as a manifestation of divine power, and this gave them a reason to campaign for its preservation against the encroachments of civilization. The tendency in American environmentalism to focus on the protection of wilderness (Muir wrote that "None of Nature's landscapes are ugly so long as they are wild") continues to this day, being a central concern of groups such as Earth First! and The Wildlands Project/Network.

What above all distinguishes contemporary environmentalism from earlier forms is that it is now, in addition to continuing local campaigns, a political movement with a strongly global perspective, and it is this form of environmentalism which can be seen as a recent addition to the ideological field, emerging as it did from the 1960s onwards. A series of books and reports in the 1960s and early 1970s served to connect a series of environmental concerns about pollution, resource depletion, and human overpopulation into a unified problem—the apparent unsustainability of industrial society. At this time we see the publication of Rachel Carson's *Silent Spring* (1962), the *Limits to Growth* report of the Club of Rome (1972), *The Population Bomb* by Paul Ehrlich (1968) and The Ecologist's *Blueprint for Survival* (1972). *Limits to Growth* in particular drew a compelling systems analysis of a series of interrelated global problems (resource depletion,

pollution, human population, industrialization, and food production) that appeared to demonstrate that, under any conceivable scenario, the industrial system was headed for collapse sometime before 2100. These arguments did not go unanswered, and there was a trenchant critique of this environmental literature from economists such as Julian Simon, in his *The Ultimate Resource* (1981). Nonetheless, these environmental studies, and many more like them, were part of the trigger for the creation of activist environmental groups such as Greenpeace (founded in Canada in 1971) and Friends of the Earth (founded in the USA in 1969 when David Brower split from the more traditional Sierra Club). We have also seen the rise of green parties in various western democracies (with some success in those countries with proportional electoral systems).

Environmentalism is not a unitary political movement and, as with any other ideology, there are a variety of currents that flow around a set of core beliefs, in this case regarding the importance of the human–nature relationship and the desirability of putting this on a sustainable footing. What that means in terms of political structures and policies can vary quite radically, and as noted at the outset, the particular set of values that environmentalists seek to sustain also varies quite widely. The next section will cover some of the main currents of environmentalism, examining in each case the structure of the ideational commitments and the kinds of policy proposals that are taken to be concomitant with them.

3. The Anthropocentric Domain of Eco-Authoritarianism and Ecological Modernization Theory

One of the initial reactions to the wave of environmental literature of the 1960s and 70s, and in particular to the ideas of *Limits to Growth*, was the perception of a global environmental crisis that required rapid, effective, and comprehensive policy change. On this view the industrialized west must rapidly curb its industrial output and give up its addiction to consumer goods. Furthermore the developing world cannot see the western industrial model as the path to rapid economic growth, as ecological disaster lies down that road. A major problem for the west under these circumstances is its liberal-democratic political system. Western publics are committed to the "wrong" values—particularly economic growth and liberal freedoms. They cannot be expected to give these up voluntarily for an abstract and future-orientated problem such as ecological collapse. Liberalism is a politics for a "time of plenty" and in times of crisis we need to turn back to the forceful Leviathan of Hobbesian politics (Ophuls 1977). There is insufficient time to cultivate the required value-change among western publics, if this is possible at all, and so we require an authoritarian state that can coerce us into making the necessary changes in *behavior*. The authors who pursued this line of reasoning (in addition to Ophuls, see Heilbroner 1974; Hardin 1968) have become known as the "eco-authoritarians" or "survivalists," as they placed the preservation of the environmental conditions necessary to sustain human societies over the liberal-democratic freedoms that these societies currently enjoyed. What was required, in Garrett Hardin's well-known phrase, was "mutual coercion, mutually agreed upon." Rule by ecological experts was necessary for at least a temporary and transitional stage, until such times as the conversion to a sustainable economic system had been accomplished. These were self-avowedly reluctant authoritarians, but the survivalist imperative necessitated the employment of drastic means.

For other environmental thinkers the dangers of this approach were all too clear (see in particular Eckersley 1992); it demonstrates what may be a naive faith in the

ability of the state to achieve its objectives against the values and wishes of its population, and it is not clear what will keep such a non-democratic state operating within the envisaged constitutional constraints. The eco-authoritarian literature does demonstrate very nicely, however, the sense of extreme urgency that followed the first wave of environmental literature. There was no time to change the values and dispositions of the people, they were deeply wedded to the "wrong" values already, and so an urgent program of top-down economic and political transformation was seen as the only viable solution. We should note that this kind of thinking has not disappeared from view; a number of ecological thinkers still despair at the inability of liberal democracies to get to grips with a range of serious environmental problems (Westra 1998; Shearman and Smith 2007), and the fault is taken to lie with the uninformed preferences of the masses. On this Schumpeterian view of democracy people are rationally ignorant and possibly rationally irrational, and so our salvation lies in a move toward "expertocracy."

Another manifestation of an environmental, as opposed to ecological response to the perceived environmental crisis is "ecological modernization theory." Despite its name, ecological modernization theory (EMT) belongs (at least in its "weaker" forms) on the "environmental" side of the divide sketched above. EMT (Mol and Sonnenfeld 2000) seeks to employ the incentives of the market and taxation in order to foster the development of environmentally benign technology and behavior, whilst not giving up on the aspiration for economic growth. In its pursuit of benign growth EMT ties in closely with ideas about sustainable development, which is most usually described, following the UN's Brundtland Report of 1987, as development that "meets the needs of the present without compromising the ability of future generations to meet their own needs." The substitution of existing employment taxes (such as National Insurance in the United Kingdom or Social Security Tax in the United States) with carbon taxes would be a typical EMT policy prescription, using the taxation system to encourage necessary behavioral change whilst relieving taxes on another "good"—employment. Another strategy that fits well with the aspirations of EMT is road-pricing, which would allow governments to capture the full social cost of road use in the direct charge to the motorist. On this analysis part of the reason why environmental problems persist is that the environment represents a free public good into which economic agents can either dump waste products or overexploit resources, thus leading to a "tragedy of the commons." To the extent that we can use market mechanisms to force economic agents to internalize these negative externalities and pay the full social cost of their activities, we can move the market economy towards a sustainable footing (see, for example, the work of David Pearce and his collaborators (1989), and for an advocacy of 'free market environmentalism' see Anderson and Leal (2001)). The overall aim is to, as far as practicable, "decouple" economic growth from growth in environmental "bads" such as pollution and resource depletion. Empirically, evidence of "decoupling" is viewed with suspicion when industrial jobs have been moving from western nations to rapidly developing countries such as China and India. Unless the imputed ecological footprints of imported goods are also taken into account, "decoupling" at the national level can appear as something of a mirage as developed nations move towards having economies dominated by service and high-tech sectors whilst importing basic manufactured goods. This is one of the ways in which the development of a global economy complicates the attribution of important indicators such as national levels of greenhouse gas emissions.

4. "Radical" Forms of Environmental Thought: Ecocentrism and Deep Ecology

At the same time as the eco-authoritarians were promoting the suspension of democracy, other thinkers inspired by the ecological problems of industrial societies sought to work on exactly the kinds of dispositions and values that Heilbroner *et al.* thought too inaccessible to rapid change to be the key to salvation. This is particularly true of an approach to environmental thought known as ecocentrism, and of deep ecology, a related but distinct "ecophilosophy." Ecocentric thinkers set themselves up in self-conscious opposition to "anthropocentrism," or "human chauvinism," which they take to be the view that either the sole, or the greatest, amount of moral value lies with human beings. This ethical hierarchism is associated with the view that non-human nature has value only insofar as it is instrumentally useful for human beings. In the absence of such instrumental value there are no ethical constraints on the human treatment of the natural world. When nature is seen as a valueless "waste" (as John Locke, for example, called "unimproved" land) then we should not be surprised at its destruction at the hands of humanity. That said, the identification of these *ethical* deficiencies does not, for some, get to the heart of what has gone wrong with human attitudes to nature. If we ask what these ethical views are based on, then we come to metaphysics, the (often implicit) understanding that people have about the nature of the universe. Deep ecologists take the "popular" conception of nature to be atomist and reductionist; in other words people hold a belief, based on their understanding of natural science, that each element of the natural world is separable and that the most basic understanding of natural processes comes at the molecular level (see Eckersley 1992 for an example). This in turn is taken to found two further views, that (a) nature is infinitely manipulable at the micro-level, and so human ingenuity and technology can always be expected to resolve our environmental problems, and (b) the human beings are autonomous creatures whose lives, especially in urban settings, are separated from the natural environment—the view and circumstance that leads to anthropocentric ethics. Against this view ecocentric thinkers offer a holistic metaphysics: all things in nature are connected, and collective entities such as ecosystems have "supervenient" characteristics that are lost in the reduction to micro-level analysis. The whole really is more than the sum of its parts. They urge a "metaphysical reconstruction" upon us, as we have to learn to see the world anew. Once we understand that we are merely "knots" in a "biospherical net" (to use a metaphor popular with deep ecologists) then we will understand that we cannot impact upon "nature" without impacting upon ourselves as well. Furthermore, if we have intrinsic value, and we are part of the biospherical web, does not the whole of the web have such value also? If so then nature possesses intrinsic, not merely instrumental, value, and must be respected as the holder of this value in the same way that human beings have been so treated in traditional forms of ethics. The logical culmination of this line of thought is taken to be "biospherical egalitarianism," a *prima facie* assertion of moral equality across all natural entities, including entire ecosystems, incorporating their non-living features such as rivers and mountains. (On biospherical egalitarianism see Naess 1973.)

Deep ecology rests on the same metaphysical bedrock as ecocentrism, a holistic view of the universe, but in some of its variants it is not an ethical theory at all, it is rather a philosophy of self—a psychological account of human development and the human capacity for an extended sense of selfhood. In the original 1973 article by Arne Naess

that launched the concept of deep ecology upon the world, "deep" ecology was opposed to "shallow" ecology in that it sought not merely to ameliorate environmental problems within the existing political, economic, and ethical system, but to probe to the "roots" of the environmental crisis which was taken to lie in a defective, atomistic worldview. However, one aspect of the deep ecological response is to urge not merely changing our views about the nature of the universe, but changing our very *sense of self* so that we come to identify, in a literal sense, with the natural world around us. If we are successful in this endeavor, then our dispositions towards, and actions within, the natural world will not need an ethical mediation (Fox 1995). We will act to protect the natural world in the same way that we would act to defend our narrow corporeal self against attack. Acts of self-defense are not taken to need ethical justification, they just flow naturally from our sense of self. (See Fox 1995, ch.11, for extensive discussion of this point.) Thus we see here ecocentric metaphysics being employed in order to develop a non-ethical defense of environmental activism.

Ecocentrism and deep ecology have their critics from within the environmental movement, few more trenchant than the late "social ecologist" Murray Bookchin, whose book *Re-enchanting Humanity* (1995) takes sustained aim at what he sees as the "ecomysticism" of deep ecology and ecocentrism. Bookchin offers in their place an evolutionary form of teleological humanism, in which humans are seen as by far the most complex creations of nature to date, to the extent that we are qualitatively distinct from the rest of the natural world due to our capacities for self-consciousness and reason. Once humanity appears on the scene, nature has bifurcated into "first" and "second" forms, with second nature having to assume an ethical responsibility for first nature. This responsibility, to be exercised according to an "ethics of complementarity," is just unavoidable: it is the natural outcome of a dialectical evolutionary process, and we can no more escape it than we can escape our self-conscious state. For these views Bookchin is charged in turn with wanting to "seize the helm of evolution" by ecocentric critics such as Robyn Eckersley. It is worth noting, contra the eco-authoritarians, that for Bookchin ecological crisis was never associated with too much democracy, but rather with too little, and of the wrong sort. Bookchin consistently advocated "libertarian municipalism," a proposal for small-scale, politically autonomous communities in which forms of participatory democracy could be practiced, connected to the wider society through a confederal system. In the period since the eco-authoritarian turn, the call for political decentralization and small-scale communal living has been a commonplace in much environmental thought, finding expression in particular in "bioregionalism," which holds that such communities should be defined by natural topographic features such as rivers and watersheds, rather than artificially constructed political boundaries.

Another important intellectual current in environmental thought has been ecofeminism. Ecofeminists make a series of connections between the historical oppression of women and human attitudes towards the natural world (Plumwood 1993; Salleh 1997). The key ideological force here is *patriarchy*, which consists in a series of value-laden binaries, which place men/rationality/culture above women/emotion/nature. Thus women and nature are placed on the same disvalued side of the ideological divide, neither fit to share in the exclusive world of (inevitably white, middle-class) male privilege. The upshot of this analysis is that both men and women have to overcome this deep-rooted patriarchal perspective. Patriarchy not only leads to the oppression of nature and the subjugation of women, it also damages men, distorting their understanding of the world and inculcating the idea that power and privilege are their natural right. A

change in consciousness is required, as part of the process of moving to an ecologically sustainable and egalitarian society. Patriarchy is often taken by extension to apply to indigenous peoples as well, as they are also placed on the side of women/emotion/nature in patriarchy's value dualism, thus some ecofeminists seek to make common cause with movements to further the interests of indigenous communities.

5. Environmental Activism

Environmentalism is not, of course, completely captured by philosophical texts. It is also an active political movement. This activism takes at least two distinct forms, electoral participation and direct action politics. Electorally, part of the 1970s reaction to the discovery of global ecological crisis was the creation of "ecology" or "green" parties in various democratic nations. One of the first of these was the UK's "PEOPLE," formed in 1973, which became the "Ecology Party" in 1975 and "The Green Party" in 1990. Other European green parties were formed in the 1970s and 1980s, and where proportional electoral systems are in operation, they have regularly returned representatives, and have participated in government in various countries including Germany, France, Belgium, Finland, and Ireland. In general European green parties are seen to belong on the political left, and a commitment to environmental protection is often allied with another three "pillars"—social justice, non-violence, and grassroots democracy (although more recently some green parties such as the UK's have moved towards having more conventional leadership roles). Participation in government by green parties has not been uncontroversial. The major split in the German green movement in the 1980s between "realos" and "fundis" was over whether to participate in a coalition cabinet, with the "fundis" promoting strict term limits, in accordance with their commitment to "grassroots democracy," and opposing close co-operation with mainstream parties. Within the party the "realos" won the debate over strategy and many of the "fundis" left the party over the next few years.

The electoral route is not the only method, of course, by which environmental objectives have been pursued by activists. The environmental movement is particularly associated with various forms of direct action politics. Part of the raison d'être for the establishment of groups such as Greenpeace and Friends of the Earth was to provide a framework for the support of forms of direct action, when environmentalists such as David Brower had come to see traditional environmental champions such as the Sierra Club as too staid and conventional. In turn a new wave of activists in the late 1970s and 1980s came to see groups such as Friends of the Earth and Greenpeace as too cautious in terms of what actions they would consider. This led Paul Watson, for example, to create the Sea Shepherd Conservation Society in 1977, to pursue a more aggressive anti-whaling policy than Greenpeace had been willing to countenance. Dave Foreman and others broke away from Friends of the Earth to establish a yet more radical environmental direct action movement "Earth First!" in 1979–80, whose slogan "No compromise in defense of mother Earth!" made their position clear. Other radical direct action groups have emerged in recent years, one of the most notorious being the "Earth Liberation Front," responsible for the arson attack at Vail Ski Resort, Colorado, in 1998. The justification of direct action of any sort lies in the perceived urgency of the environmental crisis and the irreversibility of many of its consequences, taking us back to the question of how best to respond to perceived environmental problems. Whilst the less radical groups such as Friends of the Earth and Greenpeace focus on particular environmental

issues as well as broader political themes such as social and global justice, they do not overtly challenge the prevailing liberal-capitalist economic system. The early version of Earth First! was very much associated with campaigns to preserve wilderness areas in the USA, although the UK version of Earth First! always operated with a concomitant concern for social justice. In the USA an increasing concern with anarchism and questions of social justice caused a split in Earth First!, with the more wilderness orientated founders (including Dave Foreman) going on to form the Wildlands Project. The most politically radical of the groups listed above has been the Earth Liberation Front, often seen as a "sister" organization to the Animal Liberation Front, and deeply influenced by anarcho-primitivism, which seeks the overthrow of "civilization" to be replaced by a "rewilded" world. For anarcho-primitivists the causes of our environmental crisis can be traced back to the agricultural revolution and the abandonment of hunter-gatherer lifestyles in many parts of the world. This leads to settled communities having to fight for land, and protect their territory, which in turn leads to the division of labor, as a specialist class of soldiery comes into existence, along with a bureaucracy to organize the transfer of resources from producers to the non-productive classes. To focus on the state, or capitalism, or organized religion as the vehicle of oppression and environmental destruction is to miss the bigger picture, as these are all part of the structures and ideology of civilization. The only viable way to resist this "megamachine" is via radical direct action as more mainstream approaches will always be recuperated back into the logic of the system. The Earth Liberation Front was declared America's number one domestic terror threat by the FBI in 2001, and although it has no formal leadership, and is organized along cellular lines to make it difficult to penetrate, it has come under pressure in recent years due to FBI activity, in particular "Operation Backfire" which led to a number of arrests of ELF activists between 2005 and 2009.

6. Environmental Skepticism, Post-Ecologism, and the "Death of Environmentalism"

Recent years have seen something of a "backlash" against environmentalism (Rowell 1996). This has particularly focused on some of the wilder claims of early green literature about the looming prospect of resource scarcity, mass starvation due to overpopulation, and the collapse of the existing economic and political order. As human societies have proved more resilient in the face of environmental problems than these early predictions forecast, the claims of environmentalists have been increasingly called into question, and one book that made a significant impact in terms of casting a critical eye over the "litany" of environmentalist complaints is *The Skeptical Environmentalist* by Bjørn Lomborg.

Lomborg is careful to distinguish himself from anti-environmentalists, claiming that he was a card-carrying Greenpeace member who initially sought to test the hypotheses of economist Julian Simon, whose work against the idea of limits to growth was noted above. (Simon famously won a bet with demographer Paul Ehrlich on the future prices of a series of commodities, between 1980 and 1990. The price of every commodity went down rather than up, indicating lower relative scarcity, just as Simon had predicted.) What Lomborg found was that Simon was broadly right; on most measures such as air and water quality, the environmental condition of the world is getting better not worse. "We are not running out of energy or natural resources. There will be more and more food per head of the world's population. Fewer and fewer people are starving" (Lomborg

2001: 4). Where we do face serious problems, such as with climate change, the suggested "green" responses are far too costly and inappropriate. Kyoto-style agreements are both ineffective and expensive, and we should look instead to massively increasing research and development into alternative technologies. It will often be the case that adaptation to climate change provides a better return on scarce resource use than mitigation of greenhouse gas emissions. Lomborg brings us back to an anthropocentric and broadly utilitarian ethical approach to environmental problems. Where they exist we should conduct careful cost-benefit analysis in order to determine the most appropriate policy response and we are seeking to sustain nature as a necessary resource for continuing human flourishing. What we should not do, on this view, is reduce levels of economic growth significantly through misguided environmental radicalism, as it is the very proceeds of economic growth that will allow us to overcome environmental problems in the future. The best approach to tackling our environmental problems is for governments to ensure that at least $100 billion a year is devoted to research and development into low-carbon energy systems.

Ingolfur Blühdorn (2000) focuses on the very nature of our "environmental crisis" and the way in which this crisis discourse is socially constructed. For Blühdorn it is not the material existence of environmental problems that drives waves of environmental concern, but rather the *representation* of environmental change as a moment of crisis in the discourse of environmentalism. The rise of ecologism represented a final surge of modernist politics, as the attempt was made to reconstruct society according to demands generated by new developments in scientific knowledge. In the wake of the failure of that project, we are now in an era of "post-ecologism," where green politics has become a politics of "simulation" in which we have two simultaneous but contradictory discourses. One of these continues to pay homage to the ecological ideal, demanding moves by government to set economy and society onto a sustainable footing. However, there is a second discourse that amounts to "hands off consumer capitalism!" People are deeply attached to consumerist lifestyles and the sense of identity formation that is now available to us in the sphere of consumption rather than production—"the allegedly superficial pleasures and satisfaction granted by material consumption remained much more tangible and attractive than the *real and lasting fulfillment* promised by ecologists" (Blühdorn 2000: 21). This second discourse may be less overt than the former, but on Blühdorn's view it has more political traction. Thus any politician foolish enough to take the apparent environmental "demands" of western democratic societies seriously would find themselves punished electorally in the future, as populations are hostile to the kind of changes to taxation that would be necessary to make environmentally damaging activities more expensive, in order to reflect their true social costs. Thus western democracies find themselves stuck with a combination of pro-environmental rhetoric that is not matched by toleration for any significant changes in policy.

There is some survey evidence to support Blühdorn's view. Even in surveys where people report a willingness to pay for environmental improvements in the abstract, they are often unwilling to pay for any specific policy changes that would be effective. As Iain McLean puts it, "feeling that something must be done is disconnected from feeling in favour of doing anything" (McLean 2007: 10). However, it is less clear whether this actually represents a transition to a "post-ecologist" era or merely highlights a long-standing difference between the views of environmental activists and the view of the general public. There may be widespread support for environmental objectives amongst western publics, but if these pro-environment preferences are held weakly then they

will not survive in the face of significant costs and demands for substantive behavioral change. This seems to be what lies behind the apparent "American Paradox" identified by Dale Jamieson, whereby the American public indicate an abstract willingness to pay to mitigate climate change, but refuse to support any specific measure that entails noticeable costs (Jamieson 2006).

Weak environmental preferences relate to another view which seems to indicate an "end" to environmentalism, in Ted Nordhaus and Michael Shellenberger's essay "The Death of Environmentalism" (2005) and subsequent book, *Break Through* (2007). Here they raise the question of the relationship between the core values espoused by the environmental movement, and what they see as a technocratic approach to campaigning for environmental change. They argue that environmentalism as it is currently practiced in the USA is effectively "dead" in that the campaigning models that worked for pollution abatement in the 1960s to 1980s are no longer effective for a world threatened by global warming. Indeed that approach should be allowed to "die" in order to be replaced by a new approach that has some chance of galvanizing public opinion around a new set of core environmental values. Nordhaus and Shellenberger pick up here on a very important but rather under-explored theme in environmentalism: the relationship between environmental philosophy on the one hand and environmental practice on the other. Nordhaus and Shellenberger argue, from an environmentalist perspective, for the "death" of a single-issue, technocratic approach to environmental campaigning, in which "the environment" is seen as a discrete entity, separate from other policy areas, and in need of protection against human predation. Environmentalists have cast themselves in an entirely negative relationship with human potentiality, always seeking to "limit," "restrict," or "reverse" human impacts on the natural world. Instead we should on this view embrace the capacity of humankind to use its ingenuity and creative power to overcome the existential crises we face today. It is a "powerful indictment" of environmentalists that they "so often consider our long life spans and large numbers as terrible tragedies rather than extraordinary achievements" (2007: 151). Rather like Blühdorn they wonder if "the pleasure we get from buying a trinket at the mall is any less *innate* than the pleasure we get from walking through an ancient redwood forest" (2007: 143), the point being that environmentalism needs to learn to run with the grain of social values, not constantly fight against it. And with Lomborg, Nordhaus and Shellenberger point out that ecological thinking is parasitic upon economic prosperity, and favor major investment in research and development of low-carbon technologies as the most promising environmental strategy.

7. Conclusion

Environmentalism has become an entrenched part of modern political life, and it is to some extent a tribute to the activities of environmental thinkers and activists that concern about the environment has become such a "routine" part of politics in democratic nations. What the above discussion shows, however, is that we are still very uncertain as to how to proceed in the face of major ecological problems such as climate change. We are still developing the theoretical tools that will allow us to factor in the risks and uncertainties involved in thinking about the possibility of environmental catastrophe, that will help us to think clearly about our obligations to future generations and how they can be discharged, and without which we cannot make moral sense of our relationship with non-human nature. It might be right to be skeptical about the "environmental

litany," and it could be that environmentalism in its current form will have to "die" so that something more profoundly optimistic can take its place, but we won't know if that's true without more careful thought about the nature of our environmental problems and how we might best respond to them. The continuing free exchange of ideas about what we should sustain, how, and for what purpose is an essential part of that process.

Related Topics

Anarchism, Feminism and the History of Political Philosophy, Intergenerational Distributive Justice

References

Anderson, Terry L. and Donald R. Leal (2001) *Free Market Environmentalism*, rev. edn, Basingstoke: Palgrave.

Blühdorn, I. (2000) *Post-Ecologist Politics*, London: Routledge.

Bookchin, Murray (1995) *Re-enchanting Humanity*, London: Cassell.

Carson, Rachel (1962) *Silent Spring*, Boston, MA: Houghton Mifflin.

Dobson, A. (2007) *Green Political Thought*, 4th edn, London: Routledge.

Eckersley, R. (1992) *Environmentalism and Political Theory*, London: UCL Press.

Ecologist, The (1972) *A Blueprint for Survival*, Harmondsworth: Penguin.

Ehrlich, P. (1971) *The Population Bomb* [1968], London: Pan Books.

Fox, W. (1995) *Toward a Transpersonal Ecology: Developing New Foundations for Environmentalism* [1990], Dartington: Green Books.

Hardin, G. (1968) "The Tragedy of the Commons," *Science* 162: 1243–8.

Heilbroner, R. (1974) *An Inquiry into the Human Prospect*, New York: W. W. Norton.

Inglehart, R. (1977) *The Silent Revolution: Changing Values and Political Styles Among Western Publics*, Princeton, NJ: Princeton University Press.

Jamieson, Dale (2006) "An American Paradox," *Climatic Change* 77: 97–102.

Lomborg, B. (2001) *The Skeptical Environmentalist: Measuring the Real State of the World*, Cambridge: Cambridge University Press.

Malthus, Thomas (1970) [1798] *An Essay on the Principle of Population*, Harmondsworth, Penguin.

McLean, I. (2007) "Climate Change and UK Politics, from Brynle Williams to Sir Nicholas Stern" at http://www.nuffield.ox.ac.uk/Politics/papers/2007/Climate.pdf, accessed 05/18/11.

Meadows, D.H., D.L. Meadows, J. Randers and W.W. Behrens III (1974) *The Limits to Growth* [1972], London: Pan Books.

Mol, A.P.J. and D.A. Sonnenfeld (2000) "Ecological Modernisation Around the World: Perspectives and Critical Debates," Special Issue of *Environmental Politics* 9 (1).

Naess, A. (1973) "The Shallow and the Deep, Long-range Ecology Movement: A Summary," *Inquiry* 16: 95–100.

Nordhaus, T. and M. Shellenberger (2005) "The Death of Environmentalism," *Grist Magazine* 1/13/05.

Nordhaus, T. and M. Shellenberger (2007) *Break Through: From the Death of Environmentalism to the Politics of Possibility*, Boston, MA: Houghton Mifflin.

Ophuls, W. (1977) *Ecology and the Politics of Scarcity*, San Francisco, CA: W. H. Freeman.

Pearce, David, Anil Markandya and Edward Barbier (1989) *Blueprint for a Green Economy*, London: Earthscan.

Plumwood, V. (1993) *Feminism and the Mastery of Nature*, London: Routledge.

Rowell, A. (1996) *Green Backlash: Global Subversion of the Environment Movement*, London: Routledge.

Salleh, A. (1997) *Ecofeminism as Politics: Nature, Marx, and the Postmodern*, London: Zed.

Shearman, D. and J.W. Smith (2007) *The Climate Challenge and the Failure of Democracy*, Westport, CT: Praeger.

Simon, J. (1981) *The Ultimate Resource*, Oxford: Martin Robertson.

Wall, D. (1994) *Green History: A Reader in Environmental Literature, Philosophy and Politics*, London: Routledge.

Westra, L. (1998) *Living in Integrity: A Global Ethic to Restore a Fragmented Earth*, Lanham, MD: Rowman & Littlefield.

Further Reading

Bookchin, Murray (1991) *The Ecology of Freedom*, Montreal: Black Rose Books. The most comprehensive articulation of "social ecology."

Dryzek, John (2005) *The Politics of the Earth: Environmental Discourses*, Oxford: Oxford University Press, 2nd edn. Provides an excellent overview of different approaches to environmental politics.

Goodin, Robert (1992) *Green Political Theory*, Cambridge: Polity Press. Goodin seeks to separate theories of natural value from questions of political organization and practice, arguing that the two are often conflated to the detriment of green politics.

Mellor, Mary (1997) *Feminism & Ecology*, Cambridge: Polity. An examination of the connections between feminism and green politics.

Naess, Arne (1989) *Ecology, Community, and Lifestyle*, Cambridge: Cambridge University Press. The classic exposition of deep ecology.

Porritt, Jonathon (1984) *Seeing Green: The Politics of Ecology Explained*, Oxford: Blackwell. This is a classic exposition of green ideology.

Zerzan, John (ed). (2005) *Against Civilization: Readings and Reflections*, Los Angeles, CA: Feral House. A collection of readings in anarcho-primitivism.

Part III

NORMATIVE FOUNDATIONS

27
CONTRACTARIANISM

Claire Finkelstein

1. The Structure of Contractarian Theories

The term "Contractarianism" refers to the school of political philosophy that advances a consent-based approach to the problem of State authority. Emerging most prominently with pre-Enlightenment and Enlightenment thought, the central philosophical figures associated with Political Contractarianism are generally identified as Thomas Hobbes (1588–1679), John Locke (1632–1704), and Jean-Jacques Rousseau (1712–78). The core theoretical commitment of Contractarianism is the claim that the legitimacy of the major rules and institutions of a given society can best be established by the agreement of its political subjects. The different varieties of contractarian thought pertain to the nature of the agreement the theory hypothesizes. Because it privileges the voluntary nature of political governance, contractarian political thought provides a particularly attractive foundation for democratic theory, especially insofar as the latter rests on a Lockean conception of rights (Simmons 2008).

Moral and Political Contractarianism Compared

In more recent writings, Contractarianism has often emerged not as a political doctrine, but as an approach to the central questions of moral philosophy (Gauthier 1986; Scanlon 1998). As a thesis in moral philosophy, Contractarianism holds that the ethical obligations of individuals in civil society are a product of the commitments they assume when they leave the state of nature. These commitments are rationally made, given that human beings can expect to fare better under the constraints of mutual moral obligation than in their unconstrained condition, or, on some versions of the claim, they are commitments to accept general normative principles that no one could reasonably reject (Scanlon 1998).

While much has been made of the difference between "political" and "moral" Contractarianism, the relationship between these two types of theories remains unclear. Having written his most significant work on the topic of Moral Contractarianism (Gauthier 1986), David Gauthier went on to write an essay entitled "Political Contractarianism" (Gauthier 1997), whose ambition was to extend the principles developed in his moral work to the question of political legitimacy. The aim of that work, he said, was to inquire into "the relationship between the normative claims of a society and the normative stances of its members" (Gauthier 1997: 132). His thought was that the theory he developed in *Morals By Agreement* articulated the likely normative commitments of individual members of society. To establish the rationality of individual moral

principles such as these, however, does not yet identify the normative commitments of society as a whole. Unlike Moral Contractarianism, then, Political Contractarianism must provide a framework for society to speak with one voice to its members. Political Contractarianism requires a unanimity that Moral Contractarianism need not achieve.

The foregoing thought is captured in Gauthier's further remark that "the contractarian perspective … [i]s the only one available to persons who may neither expect nor require their fellows to share their own orientation to values and norms" (Gauthier 1997: 132). In other words, people who cannot reach agreement about substantive morality can, nevertheless, agree to procedures for selecting the basic institutions of their society. It is here that Contractarianism makes its most distinctive advantage felt: whether in their political or moral incarnation, contractarian theories are compatible with multiple and conflicting conceptions of the good. But their defense of basic political and moral principles in formal terms rests on an underlying substantive commitment to the value of consent, a concept whose broad appeal is clear from its manifestation in most normative theories.

Nevertheless, the relationship between Moral and Political Contractarianism remains puzzling, insofar as Moral Contractarianism, which identifies both the mechanism by which moral norms can be developed in society and the content of those norms, would appear to leave no room for doubt about the constraints that apply to the pursuit of selfish aims. Although people may have different conceptions of the good, and thus make widely divergent life choices for themselves, in a properly constructed contractarian moral theory, they will not end up with radically different ethical constraints on the pursuit of their individual aims. If everyone were to adhere to those norms of morality established by contractarian agreement, there would be little need for a theory of political norms, since the multiplicity of private judgments that creates the demand for centralized normative authority would not obtain. If this is true, however, Political Contractarianism, unlike its moral counterpart, is likely to be a theory of the second best.

Nevertheless, it is important to note that even if agents were unanimous in the moral norms to which they subscribed, the State would often play a crucial coordinative role, even with regard to matters that are morally indifferent. No agreement with regard to private morality, for example, would eliminate the need to have a centralized decision-maker to tell us whether to drive on the right or on the left. Moreover, once a decision has been made and individual members of society can relate to one another in an organized way, political subjects newly acquire a *moral* obligation to adhere to the coordinative decisions society has made. In a world in which everyone was moral, then, society would need only the most minimal of governments, given the normative force exerted by the sheer logistical demands of social interaction (Nozick 1974).

Despite these differences, contractarian theories in ethics and politics display similar structures. Each agent starts with an injunction to seek to enhance his welfare according to rational principles of maximization. Each agent is also subject to an injunction to place restrictions on his maximization and to subject himself to an obligation to other members of society by entering into an agreement to develop relevant normative principles of constraint. Whether those constraints are individual or social, moral or political, their relation to the injunction to be rational remains the same: In both domains, contractarian principles maintain that self-constraint is necessary to achieve the highest available levels of personal welfare.

But how do we know which states of personal welfare are available and which are not? Let us say that any outcome of collective decision-making must be a "feasible" outcome, and accordingly that contractarian agreement is subject to a "feasibility" condition. A given level of personal welfare is a feasible outcome of collective decision-making if it is possible for an individual rational maximizer to achieve that level of welfare in a collective setting, compatible with others behaving as rational maximizers as well. Thus the fact that other rational agents are motivated in the way that I am restricts what I can hope to achieve by way of personal maximization in a group setting. Collective personal maximization thus renders infeasible some distributive states that might otherwise be available for adoption, say, under a system of social maximization instead. That I have everything and you have nothing is not a feasible contractarian arrangement, even if it would maximize social welfare to impose a social distribution of this sort on society.

Contractarian Approaches by Comparison

One of the clearest ways to articulate the central contribution of the contractarian tradition is by contrast with the political and ethical theories that flank it. For many years the core positions in both moral and political writings have been divided between those that defend a deontological approach to normative questions and those that defend a utilitarian, or economic, approach. The former conception conceives of the State as justified insofar as it protects individual rights. Deontological views have the defect, however, that they cannot easily make sense of the need to balance protection for such rights against considerations of social utility. Utilitarian theories, by contrast, provide a metric for considerations of social welfare, but they have difficulty incorporating intuitions of justice into that metric.

Contractarian theories, insofar as they are truly grounded in agreement or consent, hold out the promise of providing an alternative to these two historic rivals. This is because the deontological notion of a right is not a necessary safeguard if an individual cannot be subject to a measure of social control without his consent. That is, consent in a contractarian theory can take the place of the notion of a right in deontological theory, provided that individuals consent only to rules and institutions that are reasonably likely to enhance their welfare and protect other social norms. On the other side, the concept of social utility is indirectly incorporated into contractarian accounts in the idea that voluntary exchanges are agreements for mutual advantage. Contractarianism thus allows political theory to avoid the major weaknesses of the two traditional approaches, at the same time that it captures the benefits of each. The central aim of this chapter is to demonstrate that a particular kind of contractarian theory presents an attractive alternative to both deontological and utilitarian accounts. This is a significant task, as those who fail to accept the central contractarian premise will not be convinced to adopt it by negative arguments designed to show the weakness of deontological or utilitarian accounts.

2. Early Contractarian Arguments

Although Contractarianism is most prominently associated with the political philosophy of the Enlightenment, the first contractarian treatment of State authority occurs in one of the early Platonic dialogues. In *Crito*, Socrates discourses with a former student in his prison cell, while he awaits his execution for the crime of "corrupting the young." Crito has

come to implore Socrates to submit to a plan to secure his escape from prison, which Socrates steadfastly resists. Socrates' argument, in brief, is that he entered into an agreement with the State to abide by its laws, in exchange for which he enjoyed all of the benefits of Athenian citizenship, such as begetting, rearing, and educating his children in Athens. Although his consent to this agreement has been more implicit than explicit, Socrates nevertheless believes that he gave sufficient indication of his assent to the authority of the State, and that he did so in exchange for a significant set of benefits that he enjoys. Having accepted these benefits, and registered no objection, he cannot now, at the end of a long life, reject the authority of Athenian laws in favor of his own opinion of what is just. Thus by accepting the benefits of Athenian society, and in particular, by choosing to remain in Athens all of his life, Socrates has manifested his acceptance of the burdens of Athenian citizenship as well.

From this tacit manifestation of his acceptance of the laws of Athens, Socrates says it follows "[t]hat if you cannot persuade your country you must do whatever it orders, and patiently submit to any punishment that it imposes, whether it be flogging or imprisonment" (*Crito* 51b3–5). The obligation to abide by a juridical verdict, Socrates explains, is like the duty to do military service for one's country: "Both in war and in the law courts and everywhere else you must do whatever your city and your country command, or else persuade them in accordance with universal justice" (*Crito* 51ac).

Unlike most of the contractarian theories that followed, Socrates' argument in *Crito* treats the required consent as actual, tacit consent. For this reason, it would be reasonable to think of Socrates' argument in *Crito* as one version of a bargaining approach to political agreement. This places it in sharp contrast with later contractarian accounts that treat the relevant consent as hypothetical, rather than actual consent, and the grounds for agreement as normative in nature. By the same token, the account offered in *Crito* differs markedly from the versions of Contractarianism that seek to establish the relevant consent by relying on deep-seated intuitions of justice held by all rational agents. In the latter sort of account, any actual agreement among members of society would go no distance towards establishing political legitimacy. Whether actual or hypothetical, however, what identifies an account as contractarian is the particular justification offered for political (or moral) legitimacy, namely that the justification for asserting the prescriptive principles in question appeals to the willingness of rational agents to accept the normative authority of those principles.

3. Varieties of Contractarianism

The general description of Contractarianism offered above leaves many aspects of this approach open to interpretation. As a result, contractarian accounts vary significantly from one another. Not only do they differ with respect to whether the required consent is actual or hypothetical, but they also differ over whether the persons whose agreement is sought or hypothesized are informed of facts about their situation and their position in society, or whether they are ignorant of those details; whether individuals seeking one another's agreement are conceived of as rational maximizing agents or in some other way.

Normative Contractarianism

To make an initial attempt to identify the different types of Contractarianism, we will focus on a central distinguishing feature, namely the distinction between contractarian

theories that use the Social Contract to expose intuitions of justice, and those that, like Socrates, regard the agreement as in some sense the product of actual bargaining, or at least actual positions relative to others in society. The former type of Contractarianism usually has normative aims. While philosophers in this tradition make extensive use of the notion of a contract, all agreement is hypothetical, and the characteristics of the parties are also hypothetical, in that the parties are stripped of any individual characteristics. The device of the Social Contract in such theories mostly serves as a heuristic to expose our deepest intuitions about justice in a liberal society.

Normative Contractarianism descends from Immanuel Kant, but the most influential of the various forms of normative Contractarianism in recent years has been that of John Rawls (Rawls 1971). The point of Rawls' version of Contractarianism is to attempt to tap into intuitions about justice in a liberal society by asking what principles would be selected by individuals entering into a political agreement with one another regarding society's basic institutions. Rawls assumes that in this "original position" of choice, contractors would select these principles behind a "veil of ignorance," meaning that they choose without any knowledge of the particular circumstances in which they expect to find themselves once they enter society. The point of this exercise, Rawls explains, is not to determine what rational agents would actually bargain for if placed in negotiation with one another for the creation of a social contract. It is, rather, to imagine that all human beings have certain moral characteristics, as well as certain highly motivating preferences, and to ask what individuals with those characteristics and preferences would regard as a fair projection of their commitments in the political arena. This is typical of normative approaches to Contractarianism generally: they start from intuitions of justice, and work backwards to identify circumstances that would be required for those intuitions to be reflected in the resulting theory of justice.

Another important feature of such accounts is that they do not generally seek to justify appeal to the interests of actual human agents, or even of putative agents who, one might imagine, are concerned to advance a characteristic set of human aims. Instead, they are content to hypothesize how agents with intuitions of justice would regard various social arrangements. It is, in short, an exercise in considering how political institutions fare in light of certain normative commitments and the human characteristics of the agents who bear those commitments. I shall refer to such accounts as "contractualist" rather than "contractarian," given that the notion of an actual contract does little work in theories of this sort.

Rational Choice Contractarianism

The second type of Contractarianism has normative aims as well, but aims of a decidedly different character. Rational Contractarian accounts should, in theory, be willing to accept whatever agreement emerges from the designated process. There should be no backward reasoning that would require the theory to be fashioned to suit pre-conceived conclusions and no initial intuitions to constrain or guide the deliberative process. If the method is applied in its purest form, there is indeed no position outside the preferences of the parties from which to judge the content of the Social Contract, and no set of independent requirements that any contract must satisfy, with the exception of those initial procedural constraints that frame the terms, but not the content, of the bargain. In what follows I shall limit my attention to Rational Choice Contractarianism, as this is the approach that places greatest emphasis on the process of actual agreement.

Rational Choice Contractarianism descends from Thomas Hobbes (1994 [1651]). It asks what form of social organization rational agents seeking to maximize their own welfare would choose if they were to try to improve their positions relative to their pre-social baselines. In its view of human nature, Rational Choice Contractarianism shares the basic presuppositions of law and economics, namely that human beings are rational maximizers whose preferences obey certain rationality conditions (Finkelstein 2004). Thus instead of holding normative commitments, the parties hold a series of rationally ordered preferences—preferences that can be ordinally ranked and that obey the basic axioms of rationality, such as completeness, transitivity, and reflexivity. Agents with preferences such as these are available for rational bargaining. If agreement to the Social Contract requires unanimous consent, then each agent must compare how she would fare under various possible social arrangements to how she would fare under her pre-social baseline, and then agree to that social arrangement that most increases her welfare, relative to her baseline.

The trouble, however, is that it is not feasible to maximize welfare in all situations. As we saw earlier, a rational agent must take the preferences of other agents as a constraint on the degree to which she can enhance her own welfare. She must accept, in short, that social arrangements might not be maximally to her benefit. While one social arrangement might be beneficial, as compared with the status quo, there could be alternative social arrangements that would advantage the agent yet more. Still, as long as a proposed social arrangement provides for a welfare enhancement, and thus leaves an agent better off than she would be in its absence, a rational agent has reason to accept it, even if her welfare is not improved as much as it might be under another feasible alternative.

Rational Contractarianism, Economics and Utilitarianism

Although contractarian theories appear in the political philosophy canon as far back as Socrates' argument, Contractarianism did not fully enter philosophical consciousness until Hobbes, and even then, contractarian theories in the rationalistic tradition were few and far between. Most of the later contractarians who descended from Hobbes ended up writing in the normative tradition. Thus the contractarians that descend most directly from Hobbes and nevertheless betray their intellectual debt to Lock and Rousseau, eventually finding clearest expression in the Kantian tradition exemplified by Rawls. Rational Contractarianism also descends from Hobbes, but it follows a different trajectory. The aspect of Hobbesian thought that Rational Contractarianism developed was Hobbes' theory of human nature, in particular the account of rational agency that gave rise to Adam Smith, on the one hand, and to the utilitarian tradition, on the other.

A point often overlooked in discussions of rationality is the following: The theory of rational agency on which Hobbes, and modern economics, both rely is markedly different from the foundational assumptions on which philosophical utilitarians rely. The economic theory of rational agency does not, therefore, make an appearance in the utilitarian tradition, despite important affinities between the two lines of thought. Furthermore, the basic assumptions of economic reasoning and those of utilitarian normative theory are, in important respects, at loggerheads with one another.

Economic analysis relies on fairly sparse assumptions. Its central theoretical commitment is a rather non-controversial postulate about human nature, namely that human beings are rational maximizers who reason instrumentally toward the fulfillment of their

preferences or attainment of their ends, whatever those happen to be. A second crucial assumption is that the preferences of human beings are self-promoting, in the sense that the preferences of other individuals enter their maximizing calculations only indirectly (as, for instance, when a parent's preferences encompass the well-being of his child). Together, these two assumptions form the basis for the most common approach to rational agency, which, in its most extreme form, is typically referred to as "psychological egoism."

The foregoing is the standard portrayal of rational agency in the economic tradition, and one accepted by many different schools of thought in political philosophy. It became the key ingredient in contractarian theories that developed in the rationalistic tradition, at least when is the assumption of rationality was combined with other crucial assumptions. The first such assumption is that of moderate material scarcity, which, in combination with psychological egoism, strongly suggests the existence of competition and strife among human beings. A second crucial assumption is that despite the existence of natural competition, human beings can expect to fare better if they are able to overcome this natural competition and engage in cooperative ventures with one another, given that cooperation produces a surplus that can, in turn, be distributed among members of society. Cooperation, when possible, thus leaves human beings better off than they would be if they attempted to enhance their well-being and protect their survival alone. The task of Rational Contractarianism is thus to explain why rational agents should regard it as in their interest to agree with their fellow men to cooperate, rather than compete. Since rational agents, on this view, care first and foremost, about increasing their security and well-being, any political arrangement that enhances those goals should, in theory, be rationally sustainable.

Not all theories that treat psychological egoism as their point of departure, however, produce contractarian results. In particular, contemporary economic theories combine psychological egoism with the idea that the purpose of social rules is not to maximize individual utility, but to maximize *social* utility instead. On this view, the desire individuals have to maximize their own well-being is not of *normative* importance in establishing the foundations for social relations. Instead, what is of greatest importance is enhancing social utility, even at the cost of personal welfare.

Left to their own devices, rational maximizers would not favor social utility maximization as the principle to follow. They would, instead, favor individual utility maximization, and social maximization would be securely supported only where the results of individual maximization happened to coincide with social utility maximization. Since much of the time individual and social interests will diverge, the emphasis on social utility will seem irrational to the individual maximizer, and the insistence on individual maximization will appear unduly self-serving from the standpoint of concerns about social welfare.

It is crucial to notice that utilitarians do not share the economist's central descriptive thesis about human nature. On the contrary. At least the early utilitarians were clear about the fact that philosophical egoism does not provide a terribly good foundation for utilitarian moral theory. In his *Methods of Ethics*, for example, Henry Sidgwick (1907: IV, I, 1) wrote as follows:

> The difference … between the propositions (1) that each ought to seek his own happiness, and (2) that each ought to seek the happiness of all, is so obvious and glaring, that instead of dwelling upon it we seem rather called upon to

> explain how the two ever came to be confounded, or in any way included under one notion. … [C]learly, from the fact that everyone actually does seek his own happiness we cannot conclude, as an immediate and obvious inference, that he ought to seek the happiness of other people.

Bentham (1879: 63) likewise was painfully aware that the normative ideals of utilitarian theory could place considerable psychological strain on ordinary human beings, and that the impulse to maximize social welfare might have to be inculcated through a laborious process of education. In order for the ideal of utilitarian moral theory to be met, then, human beings would most likely have to be altruistic in nature—the opposite of what the narrowly conceived egoistic self-dealers economic theory assumes.

An important conclusion can be drawn from the foregoing discussion: The theory of individual rationality to which economists are committed is neither entailed by nor entails the theory of value that economists inherited from the early utilitarians. There is, in fact, no intrinsic relation between personal utility maximization and social utility maximization. Strictly speaking, Utilitarianism does not even maintain that subjects of a utilitarian regime must possess the capacity to reason in a maximizing way about their own utility. We could, for example, ask what the best life for cows would be and seek to maximize their utility by providing them with grassy fields and plenty of water. But we need not think cows capable of reasoning on their own behalf about what would maximize their own utility, still less of engaging in anything resembling instrumental reasoning in order to be fitting subjects of utilitarian normative theory. This is, of course, well recognized in the history of Utilitarianism, and it is both the source of creative uses of utilitarian moral theory as well as the basis for objections to it. Utilitarianism, for example, has been used to defend more humane treatment of animals (Singer 1975). It has also been reviled because it appears to suggest that we should give up our egocentric pursuits and take up animal husbandry, given that we would bring about the largest increases in total social utility by doing so.

Now consider the claim that a maximizing individual psychology does not entail a commitment to the normative theory of social utility maximization. This point has traditionally been less obvious to rational choice theorists than the previous one, but the logic is just as clear once one reflects on the point. If agents are individually rational, it is highly unlikely they would be social welfare maximizers, given that the state of affairs that maximizes an individual's personal utility will almost invariably be different from the state of affairs that would maximize society's welfare. This suggests that not only does rational actor psychology not entail the utilitarian theory of value; the two are actually in some tension with one another. The tension stems from the fact that when we maximize social utility, we inevitably sacrifice the welfare of some individuals for the sake of achieving greater gains in social welfare overall (though this assumes the possibility of interpersonal comparisons of utility). Alternatively, we might say that when we maximize social utility, we can only be assured that the upshot would allow the winners to compensate the losers, not that resulting distributions fail to make anyone worse off than they were. In other words, we can only be assured that outcomes are *Kaldor-Hicks efficient*, not that they are *Pareto efficient*. This is, of course, another reason why the normative theory of Utilitarianism has been so fraught with controversy: In the process of maximizing social utility, we must often override considerations of individual welfare, including considerations moral philosophers think of as protected by *individual rights*. As has often been noted, traditional Utilitarianism is indifferent to distributions of utility that do not affect total value.

What this suggests is that a normative political theory that subscribes to psychological egoism at the same time that it assumes utilitarian normative theory has some explaining to do. It must explain why the normative theory it inherits from the utilitarian tradition is not fundamentally at odds with its assumption about individual human psychology. Applied economists, particularly legal economists, do have a proposal for reconciling the two, though they rarely, if ever, state the point explicitly. They admit that the gap between the goal of social welfare maximization and individual rationality is real and regrettable. This gap, however, can be easily closed by the judicious use of social rules. When social rules are correctly formulated, they ensure that when individual actors maximize their own utility, they will be maximizing social utility as well. Public rules and institutions are able to accomplish this convergence of personal and social utility by restricting individual maximizing to socially useful bounds. This reconciliation of personal and social utility is justified in contractarian terms when the constraint on individual maximization is one to which rational individuals would give their consent.

4. Contractarian Jurisprudence: Testing the Theory

To the extent the contractarian tradition has been incorporated into legal theory, it has been nearly all of the normative variety. Legal theorists have tentatively explored the application of Rawlsian-style Contractarianism to international law, punishment theory, contract law, and even to specialized subjects such as bankruptcy. For this reason, the more straightforward project for the present work would have been the application of normative Contractarianism to problems in legal theory. But arguably there is reason to eschew this branch of Contractarianism in favor of its rationalistic cousin, despite the popularity of the former as compared with the latter.

Consensual Coercion?

As we have seen, the justificatory force for the use of State coercion lies in the notion of *consent*. The same applies to the legal variant of Political Contractarianism. What makes Contractarianism a significant and potentially superior alternative to utilitarian and deontological legal theories is that, at least in principle, contractarian theories seek their justificatory force in the consent of legal subjects. This implies that the legal institutions that appear to be coercively organized are in fact the products of choice on the part of the governed, and therefore they are not fundamentally coercive. The more voluntary a legal organization, the easier it becomes to justify the imposition of the rules of that institution on presently unwilling subjects. But what kind of framework precisely is needed before we can think of a legal institution as consensual? When we attempt to apply a contractarian approach to justification to legal institutions, we put pressure on the claim that it is possible to justify social institutions by showing them to be consensual. For law is a coercive social institution. It is a particular challenge to justify a coercive institution as consensual and to see its dictates as voluntarily imposed. How can we justify forcing someone to pay damages for failure to satisfy the terms of a contract, or imposing a civil damages remedy, or forcibly depriving someone of his liberty, in terms that represent the recipient of such coercive measures to have consented to such treatment?

Let us consider in greater detail what a contractarian approach to a particular legal institution. The case of criminal punishment is particularly useful for considering

Contractarianism in application. For the more coercive the institution, the more difficult it ought to be to justify in contractarian terms. There are, however, several ways of conceiving of punishment as voluntarily imposed, despite the fact that the criminal does not welcome such treatment at the moment it is delivered.

A Contractarian Approach to Criminal Law

Assuming that a given punishment scheme has at least moderately strong deterrent efficacy, a rational agent would agree to live in a regime that furnishes such deterrence to serious crimes, since he would prefer it to one in which such deterrence is absent. The attraction of a contractarian approach is that the justificatory burden for consensual arrangements is particularly low. This implies that if a contractarian account of punishment is possible, it should be substantially easier to justify the infliction of punishment on such an account than on any other.

Suppose a group of contractors in an initial bargaining position is trying to decide how much and what kind of protection they should institute for private property. They have already selected a series of rules establishing a system of ownership, and they now seek a means of enforcement. They must weigh the following considerations. On the one hand, they would like the maximum deterrence feasible for violations of ownership rights. On the other hand, they also want to protect their own personal freedom and would like to maximize independence of choice without interference from others. Maximizing independence of choice would leave no protection for ownership, while maximizing protection for private property would sharply curtail personal liberty.

In balancing security and liberty, each person should ask himself: Would I be better off in a society that established penalties for theft and other violations of property norms than I am at my current baseline welfare? In answering this question, each agent might weigh the benefit she would receive from increased deterrence against the loss she would suffer in the worst case scenario—that is, the balance of gains and losses she might experience under the rule under the most unfavorable circumstances. The worst case scenario is clearly the case in which the agent is most disadvantaged by the rule: She is herself liable to suffer the increased penalty but has little property to protect. Thus she must ask whether she would be advantaged on balance from penalties for theft in a world in which she was herself subject to such penalties, as compared, say, with a baseline in which there was no protection for private ownership at all. If the penalties for theft are too low, the deterrent effect will be insignificant and private property will not be protected. If the penalties are too high, agents receiving the penalty would be worse off than they would have been in the absence of private property and the requirement of benefit would not be satisfied. (See Finkelstein 2011.)

Although a fully worked out theory of punishment along Rational Contractarian lines has not yet been achieved, the merits of such an account, if available, are clear. Contractarian accounts are potentially able to provide a justification for coercive institutions that is addressed *to* the recipients of such coercive treatment. This seems a necessary feature of any moral justification for the use of coercion. A justification such as deterrence purports to justify punishment, but it does not provide a reason for the use of coercion that its recipient need care about. Deterrence, after all, purports to justify inflicting punishment on one individual for the sake of deterring others. To offer such a reason for its infliction fails to appeal to anything the agent might see as warranting his approval of such a scheme.

5. Final Thoughts

Rawls says (1971: 12): "Our social situation is just if it is such that by [a] sequence of hypothetical agreements we would have contracted into the general system of rules which defines it." He continues (1971: 12):

> No society can, of course, be a scheme of cooperation which men enter voluntarily in a literal sense; each person finds himself placed at birth in some particular position in some particular society, and the nature of this position materially affects his life prospects. Yet a society satisfying the principles of justice as fairness comes as close as a society can to being a voluntary scheme, for it meets the principles which free and equal persons would assent to under circumstances that are fair. In this sense its members are autonomous and the obligations they recognize self-imposed.

It should be clear that the sense in which obligations are "self-imposed" in Rawls' scheme is highly attenuated, since the original position involves neither actual agents nor actual agreement, and so a fortiori the individuals restrained by a system of justice have not in any sense agreed to be so restrained. Rawls conceived of the members of the original position as either "representatives" of actual persons, or as embodiments of certain intuitions of justice. This would serve to explain why their "consent" might be binding for real subjects. But, as is often pointed out by critics of Rawls' original position, it is not clear why hypothetical creatures lacking in all human characteristics should be thought of as an adequate substitute for, or a representation of, actual persons.

Rawls has in effect responded that the issue is not about representation—it is about fairness. Each actual person should recognize the rules under which he is punished as legitimate because they correspond to his deepest sense of the fairness of basic institutions, elicited through the thought experiment of the original position. But recognizing certain rules as fair does not by itself mean a person would consent to be governed by them. Fairness might, of course, ultimately justify imposing those rules on him, regardless of whether he accepts them. But that is a different story, and it is not, at any rate, a contractarian story.

The rationalist approach to contractarian justification depends on our ability to conceive of even coercive social institutions as voluntarily agreed to by rational agents. I have suggested that even highly coercive institutions can potentially be thought of in this way. A fuller account of contractarian justification of actual institutions in the rationalistic tradition would require far more detail than I have been able to provide here. But there is no reason, in principle, to think such details cannot be fully articulated on another occasion.

Related Topics

Hobbes, Locke, Hume and Smith on Justice, Contractualism and Political Liberalism, Utilitarianism and Consequentialism, Rational Choice Theory

References

Bentham, Jeremy (1879) *Principles of Morals and Legislation*, Oxford: Clarendon Press.

Finkelstein, Claire (2004) "Legal Theory and the Rational Actor," in Alfred R. Mele and Piers Rawling, eds., *The Oxford Handbook of Rationality*, Oxford: Oxford University Press, pp. 399–416.

—— (2006) "A Contractarian Argument Against the Death Penalty," *New York University Law Review* (October): 1283.
—— (2011) "Punishment as Contract," *Ohio Journal of Criminal Law* 8: 319.
Gauthier, David (1986) *Morals By Agreement*, Oxford: Oxford University Press.
—— (1997) "Political Contractarianism," *Journal of Political Philosophy* 5: 132–148.
Hamilton, Edith and Huntington Cairns (eds) (1961) *Crito* in *The Collected Dialogues of Plato*, Hugh Tredennick trans., Princeton: Princeton University Press.
Hobbes, Thomas (1994 [1651]) *Leviathan*, ed. Edwin Curley, Indianapolis: Hackett.
Nozick, R. (1974) *Anarchy, State and Utopia*, New York: Basic Books.
Rawls, John (1971) *A Theory of Justice*, Cambridge, MA: Belknap Press of Harvard University Press.
Scanlon, Thomas (1998) *What We Owe to Each Other*, Cambridge, MA: Harvard University Press.
Sidgwick, Henry (1907) *The Methods of Ethics*, 7th edn., Chicago, IL: University of Chicago Press.
Simmons, John (2008) *Political Philosophy*, New York: Oxford University Press.
Singer, Peter (1975) *Animal Liberation: A New Ethics for our Treatment of Animals*, New York: Random House.

28

CONTRACTUALISM AND POLITICAL LIBERALISM

Aaron James

Contractualism and political liberalism are central ideas of the political philosophy of John Rawls. The two ideas are also potentially inconsistent. Considering Rawls's way out of the inconsistency highlights fundamental questions about the aims and subject matter of political philosophy that are as yet unresolved in the social contract tradition.

1. The Potential Inconsistency

As we will understand it, *contractualism* is the following philosophical thesis about the nature of justice:

> Justice is, by its very nature, a matter of what is justifiable to all, in light of what each could reasonably agree to or accept, under certain idealized conditions.

So, for example, Rawls's "idealized conditions" involve choice from the Original Position: we ask what principles each member of society would self-interestedly accept in ignorance of his or her social position (Rawls 1971: 17–22). This tells us what social arrangements are, in fact, justifiable to all involved. Each is fairly represented, on a footing of equality, so when the chosen principles are implemented in practice, no one—not even the least well-off—can reasonably complain of the result.

The assumed link between justifiability and idealized acceptance is explicit in T. M. Scanlon's contractualism about "what we owe to each other." According to Scanlon, what is justifiable to others, and in that sense right or just, is a matter of what principles of conduct "no one could reasonably reject," when fully informed, unforced, and seeking principles of conduct that are so justifiable to all (Scanlon 1998: 4). Scanlon's "idealized conditions" moralize motivation, rather than restricting information, in part because Scanlon is more concerned with interpersonal moral motivation than with justice in society. Yet the underlying contractualist idea is much the same. (Rawls, however, preferred the name "contractarian" in part to distinguish his view from Scanlon's project (Freeman 2007b: 36).)

As we will understand it, political liberalism is about a particular issue of justice, the legitimate exercise of political power. The main idea is expressed in Rawls's *principle of legitimacy*, which may be put (in our own terms) as follows:

> The exercise of basic political power is legitimate only when its grounds are publically available, available, that is, from within every reasonable subject's point of view.
>
> (Rawls 1993: 137, 217)

In domestic society, the principle applies in "constitutional essentials" and matters of "basic justice" and requires officials (judges, legislators, administrators, campaigners, and voters) to base their decisions only in "political" values that all citizens can accept. On the international stage, the principle requires that societies use force and coercion (in war, coercive diplomacy, economic sanctions, and so on) only on grounds available to all societies (in effect, a list of intervention-legitimating human rights) (Rawls 1999a: 64–70). While Rawls speaks of "common" or "public" reason in both contexts, it is not sufficient that the ideas have been aired in public life; the relevant grounds of power must be available, as we are putting it, "from within the point of view" of those subject to the form of power in question. Nor will all reasonable points of view necessarily converge. For, Rawls explains, it is a general fact of life, especially under free institutions, that different fully reasonable people will have to make up their minds about philosophy, morality, and even the relative weight of "political" values given very different experiences, upbringings, trusted traditions and authorities, and in light of conflicting evidence and open-ended matters of interpretation. The "burdens of judgment," as Rawls calls them, normally generate divergence rather than convergence in belief (Rawls 1993: 55–7).

But now consider the *thesis of public unavailability*:

> The philosophical thesis of contractualism, along with its substantive principles (e.g. Rawls's egalitarian principles), are not available to every reasonable (adequately developed, educated, informed) subject of political power from within their respective points of view. This is true of societies, in international relations, but also true of citizens within societies, even or especially democracies.

If this thesis is correct, it follows that political power (in its relevant "basic" forms) cannot be legitimately grounded in either contractualism or its substantive principles. In short, the contractualist/egalitarian philosophical outlook is just too controversial; the background of agreement needed for its legitimate enforcement just isn't there.

Why might that be so? Leaving aside substantive and potentially controversial principles for the moment, might the general thesis of contractualism at least be available in the relevant sense? Perhaps, but this can hardly be assumed. As a purely psychological matter, we certainly can't expect all or most or many people to agree that justice is *by its very nature* a matter of justifiability to others, even upon careful examination—and perhaps even after extensive cognitive psychotherapy. That is surely true for many religious citizens and societies, but perhaps true for many non-religious people as well. A good number of people just won't see justifiability *to* others as a central moral idea, in contrast, say, with what is owed to God or what is best overall for people taken in the aggregate. Still others might accept a general contractualist conception but deny that Rawls's egalitarian principles (or other contractualist principles) are the best way to understand what justifiability to all requires. And again, much as Rawls explains, this spread of opinion need not reflect unreasonable religious or philosophical outlooks (e.g., fundamentalisms which are immune to reasoned revision). Reasonable people naturally disagree about such things.

None of this is yet to say that contractualism or its principles are false or mistaken. That does follow, however, when contractualism is read as a *normative* thesis, about how society ought to be governed. On this reading its principles present decisive reason for political action, even or especially on "basic" matters of justice in the domestic or international setting. But this normative thesis seems just wrong if it turns out to be *illegitimate* to enforce contractualist principles, as according to the reasoning above. We would then have decisive reason *not* to guide society according to contractualist principles, since truly illegitimate exercises of state power are presumably out of the question.

Rawls accepts all of this as a theoretical possibility. His way out of the potential conflict is to deny the thesis of public unavailability. When "public" grounds for political power are understood in the right way, he argues, we see that the background of agreement needed for his contractualist proposals is indeed to be found. It is found in the generally, if implicitly, understood political culture of modern constitutional democracies, and possibly elsewhere, though perhaps not in some religious or "hierarchical" societies which do not assume that all are "free and equal." Rawls's position thus ultimately depends heavily on social interpretive assumptions about prevailing political culture at a certain stage of history. It isn't an argument from pure reason, in the style, say, of Kant. If the picture is Kantian, we might say, it is of a Hegelian Kant.

2. The Force of the Problem

Before developing Rawls's position, we might consider alternative ways of escaping our problem. Our problem has force because several of the main alternative routes of escape leave a lot to be explained.

A first route is to reject contractualism about justice. Yet Scanlon's elaboration of the underlying contractualist idea shows that it has considerable intuitive force, in both interpersonal and institutional settings, quite aside from Rawls's specific use of it. When I am considering some action, I will naturally take pause if it seems that I could not justify that action to you, on grounds you could reasonably refuse to accept. If you suffer at my hand, and then ask me why I acted as I did, it seems especially important if I cannot give a reply which you have to accept on pain of being unreasonable. And if you could indeed reasonably reject my operative reasons for action, it doesn't matter if you never have a chance to call me out: the mere thought that my conduct is unjustifiable to you, that you could reasonably reject it (or its operative grounds), is itself a sufficient reason for me to have done something else, or to now have feelings of self-reproach.

Scanlon readily admits that "what we owe to each other" is not the whole of morality. Nor, presumably, is it the whole of anything rightly called "justice." Yet it does seem applicable beyond interpersonal morality, and indeed Scanlon's most powerful illustration of justifiability to others is institutional. He writes:

> In the 1950s many Americans believed, naively, that their institutions were uniquely justifiable; that America was free of class barriers, and that it was a society in which benefits were fairly earned. They therefore felt that they could enjoy these benefits in the comforting confidence that the institutions through which they had acquired them, though not perfect, were closer than any other sort being ones that no one could reasonably object to. The combined blows of the civil rights movement and the movement that arose in reaction to the war in

> Vietnam shattered these illusions beyond repair. … What [the] reactions had in common was a deep sense of shock and loss; both testify, I believe, to the value people set on the belief that their lives and institutions are justifiable to others.
> (Scanlon 1998: 163)

The example is more than suggestive, since Rawls has shown how far the same underlying contractualist idea goes as an account of social justice. Even if the result is not the "whole of justice," as Rawls admits, it need only, and arguably does, capture one of its central and perhaps politically most important parts.

A different way out of our inconsistency problem is to deny that contractualism advances a normative thesis. Its claim, one might say, is merely about standards of evaluation, akin to the standards for a good or bad watch, basketball player, or colleague: the theory only says when the evaluative concepts "just" and "unjust" apply, without implying that justice and injustice have any essentially decisive force for politics and social institutions. Contractualism would then be seen not as a mistaken normative thesis, but as a true evaluative thesis about what full justice requires over and above distinctive concerns of legitimacy (perhaps seen as either a subset of, or precondition for, full justice).

Rawls, however, clearly has stronger normative aspirations. He writes, with a ring of commonsense, that "laws and institutions no matter how efficient and well-arranged must be reformed or abolished if they are unjust" (Rawls 1971: 3). He never adequately explains the normative force of justice (does it simply reflect his Natural Duty of Justice, or an essentially collective obligation, borne by societies, or the international community, which is separable from the duties of individuals?). But Scanlon's account does develop and qualify the relevant normative "must." It shows how the idea of justifiability to others could support a normally conclusive reason for action, partly by explicating its force on its own terms, as suggested above, and partly by absorbing the force of potentially competing values. That is, because associated principles need only be *normally* conclusive for action, there is ample room for extenuating circumstances; and because contractualist reasoning admits various grounds for reasonable complaint—including liberty, welfare, efficiency, arbitrary treatment, and so on—various values can shape what is and is not justifiable to each and to all without changing its basic normative force. We can plausibly say, at least presumptively, that laws and institutions must be reformed or abolished if they are unjust.

A still different way out of our problem is to relax the principle of legitimacy. Legitimacy might be said to depend on what is *reasonably acceptable* quite aside from what would be accepted *from* so many points of view. In that case, however, the challenge is to explain why *legitimacy* is the concern. Why won't a liberal society that enforces a conception of justice be in the same illegitimate boat as a society that enforces an official religion? Both, we may say, act in the name of the truth. But it won't do *simply* to say that the liberal outlook *is* true—as though the real issue in the religious wars of Europe was whether Protestant or Catholic Christianity was closest to being right. The beginning of wisdom—and the historical beginning of liberalism—lies precisely in the thought that bare truth will not suffice.

If legitimacy must be in *some* way sensitive to public perceptions, the required sensitivity might be shaped by several factors. First, if justice is to be done, or even doable, there will have to be large-scale institutions and social practices which command support over time, in part because enough people regard them as legitimate. As Rawls puts the idea, in terms of "stability," it is "futile to try to realize an unstable conception

[of justice]" (Rawls 1999b: 185). Second, to the extent such institutions exercise political *authority*, as opposed to pure coercion, they must purport to tell people what to do on grounds that are in principle authorizable from within their own perspectives. That will be true in domestic law but also in international law and practice, which will have to be, in principle, authorizable from the point of view of every society involved (as reflected in established treaties or, perhaps, in mere emergent practice). Third, the demands of legitimacy may be shaped by positive democratic ideals. As Rousseau might put it, insofar as terms of cooperation reflect the "general will," those addressed count as both the subjects and the authors of law—at once forced and free. That ideal has special resonance within democratic society, but also applies in international relations and perhaps non-democratic societies in the ideal that shared relations are to be governed collectively, based on mutual respect (Rawls 1999b: 62).

3. Reasoning from Agreement

For roughly these reasons, Rawls takes a different route of escape from our inconsistency problem: he rejects the thesis of public unavailability. According to Rawls, in both the domestic and international contexts, both contractualism and at least some of its specific principles are indeed available from within the point of view of each reasonable subject of political power. For, Rawls argues, both are indeed part of the best understanding of the social practice and political culture in which those subjects are themselves involved. In that sense, Rawls constructs bases for reasonable political agreement, not from pure reason, but from agreement itself.

While this is especially clear after Rawls's "political turn" (see, e.g., Rawls 1993: 13–15), it is also quite explicit in his most Kantian phase, where contingent societal understandings would seem least at home. In "Kantian Constructivism in Moral Theory," just after formulating a "Kantian constructivism" which reasons solely from ideas about the nature of persons as such, Rawls explains that "On the Kantian view that I shall present, conditions for justifying a conception of justice hold only when a basis is established for political reasoning and understanding within a political culture" (Rawls 1999a: 305). As Rawls would later explain, he here follows Hegel. "[O]ne of [Hegel's] great contributions" is to see "the deep social rootedness of people within an established framework of their political and social institutions. … The concepts of person and society fit together; each requires the other and neither stands alone" (Rawls and Herman 2000: 366). Accordingly, Rawls describes the "real task" of political philosophy as "The search for reasonable grounds for reaching agreement rooted in our conception of ourselves and in our relation to society." Or as he elaborates:

> The task is to articulate a public conception of justice that all can live with who regard their person and their relation to society in a certain way. What justifies a conception of justice is not its being true to an order antecedent to and given to us, but its congruence with our deeper understanding of ourselves and our aspirations, and our realization that, given our history and the traditions embedded in our public life, it is the most reasonable doctrine for us. We can find no better basic charter for our social world.
>
> (Rawls 1999a: 306–7)

How then does what is "most reasonable for us" depend on "how we conceive of our-

selves," as somehow informed by "our history and the traditions embedded in our public life"? Rawls's general answer is a constructive method of justification, which has the following three argumentative steps. (The order is irrelevant within a broader "reflective equilibrium" methodology.)

The first stage is *individuation*: we single out a social practice which provides an object of social interpretation and moral evaluation (e.g., the "basic structure of modern constitutional democracy," or "international law and practice"). We do so at first in relatively uncontroversial sociological terms, with reference to various interpretive data points or source materials that any further conception of the practice should take into account and explain (or "explain away"). We also identify a general moral concept—justice, fairness, or legitimacy—which has currency in political argument in the context in question, by explicating that concept's distinctive practical role. Thus justice, seen as a kind of fairness, is to serve as a basis for public governance of a society's basic structure, while legitimacy governs the workings of domestic government and the international use of force and coercion. This is generally a plausible step because it is not yet to specify any particular interpretation of either the practice in question or any "conceptions" of the operative concepts. This step brackets those more controversial issues.

The second stage is *framing characterization*: we work up a general characterization of the nature of the practice, in light of its distinctive structure and generally assumed purposes. Although this is not yet to offer substantive moral principles, the characterization can be moralized in at least three ways. First, it can cite and develop moral elements of the practice (e.g., stated or implicit rights, goals, or principles), so long as they are part of an *interpretation* of the practice, which comports with familiar interpretive constraints of consistency, coherence, explanatory power, simplicity, and so on. Those constraints allow or even require "constructive interpretation" of the practice in the sense associated with legal or artistic interpretation (Dworkin 1986). Second, because the proposed characterization is to frame and structure moral evaluation, it can draw from very general ideas associated with the moral concept in question (e.g., "formal constraints on the concept of right") as well as particularly well-entrenched moral judgments (that slavery is unjust, that state coercion in the name of religion is illegitimate). Third, the plausibility of an interpretation can depend on the possibility of seeing the practice's purposes and corresponding organization as having a potentially legitimate general point. We don't seek an internal justification of plainly illegitimate forms of social activity, as though "the practice of slavery" could be internally just (though we might ask whether slavery could be a just "system of property" and answer "no").

The third stage is *substantive reasoning*: we engage in substantive moral reasoning about what the various relevant reasons in play support, as guided by the specified framing conception and our more general understanding of contractualist moral reasoning. We make substantive moral judgments, which we take to be defensible upon reflection, subject to the constraint that they must be reflectively available to reasonable parties involved. As Rawls explains:

> The real task is to discover and formulate the deeper bases of agreement which one hopes are embedded in common sense, or even to originate and fashion starting points for common understanding by expressing in a new form the convictions found in the historical tradition by connecting them with a wide range of people's considered convictions: those which stand up to critical reflection.
>
> (Rawls 1999a: 306)

In these terms, Rawls's "framing characterization" takes the form of certain "model-conceptions" of how we are to conceive of ourselves and our society (Rawls 1999a: 307–8). In democratic society, we conceive of persons as "free and equal," and society as a "fair scheme of cooperation," which is publicly governed and exists for the sake of mutual benefit. We thus rephrase the question of social justice in more specific terms, which lend themselves to more informative reasoning, clearer competing answers, and more convincing reasons and verdicts. That is, we rephrase and so "frame" the question, "What is a just society?" as the question, "What public terms of cooperation are fair to each person, seen as free and equal?" It is in response to this more contentful question that Rawls then "works out" the rest of his theory—as a matter of what we are calling "substantive reasoning"—including his list of "primary social goods," the Original Position's design (e.g., who is represented, how much the parties know, etc.), and the resulting argument for his egalitarian principles.

For our purposes, the crucial question is about the assumed conceptions of person and society. Where do they come from? We are suggesting that Rawls's answer (still part of what we are calling "framing characterization") involves a "constructive interpretation" of ideas implicit in democratic society. The idea that society is a publicly governed scheme of cooperation for mutual benefit, and so open to assessment as fair or unfair, makes the best overall interpretive sense of a democratic society's history, public texts and speeches, avowed or assumed principles, specific institutions, legal and judicial practices, and perhaps more. Although Rawls tends to simply put forward his proposed framing conceptions in their most plausible form, rather than ruling out rival interpretive conceptions, they are in principle open to thorough interpretive defense (James 2005).

While all of this is clear in Rawls's most Kantian phase, Rawls would later drop certain Kantian emphases (e.g., "persons" become "citizens"). As he gradually became less sanguine about whether an agreement among reasonable people would be reached (given the "burdens of judgment"), he became more concerned to emphasize that he is drawing from established political culture rather than ideas about the bare moral nature of persons. Rawls continued to work with some ostensibly Kantian ideas, for instance, that people are "self-originating sources of valid claims," but only insofar as he took these to be open to philosophically neutral interpretation: people are "self-originating sources of valid claims" just in the familiar sense that they have standing to address claims, and are owed justification for how they are treated, whatever the ultimate ground for this might be (whether it be the rational nature, as for Kant, or a divine grant of personal sovereignty, as for Locke). That idea, along with the moral significance of the reasonable, is supposed to be uncontroversial enough to command agreement from all reasonable people. Other ideas are more challenging. *Political Liberalism* focuses on the harder case of religious belief, while *The Law of Peoples* focuses on what is acceptable to "non-liberal" but "decent" societies. In each case, Rawls conjectures that his contractualist proposals can indeed provide a stable basis for social cooperation—a "reasonable overlapping consensus"—given the general facts of human social life. The argument, in each case, works by imagining mutual adjustment between culturally grounded reasonableness and fragmented and pluralistic "points of view" (barring, that is, unreasonable fundamentalisms that are unresponsive to reasons).

We might speculate about Rawls's increasing concern with what is available within specific points of view, as opposed to what is merely "out there" in a public culture (e.g., in an honored but obscure text). It may reflect a deep shift in interpretive methodol-

ogy, roughly, a shift from a "protestant" to a more "catholic" approach (Postema 1987). A "catholic" interpreter assigns special authority to how participants in a social practice *themselves interpret what they are doing together*, and so will be especially sensitive to both shared understandings and disagreement in interpretations among participants. A "protestant" interpreter gives no special authority to participant interpretation. How participants interpret their practice, and whether they agree or disagree, represents only so much initial data or source material from which the interpreter must construct his or her own best characterization of the practice overall (e.g., Dworkin 1986). In these terms, Rawls's early writings are "protestant": he simply works out his own proposed best understanding of domestic societal practice without great concern for people in the practice who might disagree. The later Rawls is more "catholic": he accepts a special burden to take disagreement into account and addresses it on its own terms. He thus offers not simply a constructive interpretation of overall public practice, but a constructive interpretation of what is available to its (reasonable, informed, educated) participants, given the (perhaps implicit) understandings they can be assumed to have by being so involved.

In effect, then, we are suggesting that Rawls's "political turn" was a mere shift of emphasis. It is neither a fundamental revision nor a complete change in topic. Some have misread Rawls as having fundamentally revised *A Theory of Justice*, by weakening his socio-economic egalitarianism. On one version of this reading, Rawls's Difference Principle becomes an optional member of a "family" of liberal conceptions. But this ignores the fact that Rawls did change his focal question, from justice to legitimacy: the Difference Principle still characterizes justice, even if other *incorrect* principles are also legitimate (Estlund 1996).

Yet it is also wrong to characterize this as a complete change in topic, as though the new concern with legitimacy is discontinuous with the concerns that animated *A Theory of Justice* (see Weithman 2011 for emphasis on continuity). Rawls himself says that the new concern with legitimacy seeks to correct a "serious internal problem" in that book, namely, that the account of stability (in Part III) is "unrealistic" (Rawls 1993: xix). The agreement needed for a just society to last (to reproduce itself across generations and to recover from external shocks) must come from participants' own sense of justice, as shaped by their upbringing in that very socio-cultural milieu. Rawls had suggested that people attach special importance to their own personal Kantian autonomy, as realized in just social arrangements, but came to find this too utopian. He replaces this with what he takes to be the more modest, more realistic conjecture that people can be brought to abide by and support fairness, in part because it is personally beneficial in various ways, and in part for its own sake (see Rawls 1993: 202–3 on the "two moral powers.")

Still, are not concerns of realistic stability quite distinct from questions of justice? Not for Rawls. Both are aspects of the agreement-sensitive problem of justice in the first instance. What Rawls's later Kantian period helped him to see was that many Kantian ideas were too controversial to be consistent with his original aspiration, in *A Theory of Justice*, to work out an ideal of "for a democratic society" from "widely accepted but weak premises." The role of agreement appears in the early pages of that work, where Rawls is setting out what we are calling the "individuation" stage of the constructive method. He characterizes the *concept* of justice in terms of "the role which … different principles, different conceptions [of the concept], have in common," where this amounts to the quite substantial, shared assumption that "institutions are just when no arbitrary dis-

tinctions are made between persons in the assigning of basic rights and duties and when rules determine a proper balance between competing claims" (Rawls 1971: 5). So much is *taken as agreed* if open to divergent interpretations. Rawls calls such moral agreement a "prerequisite for a viable human community," in addition to the agreement required for coordination, efficiency, and stability. And because such forms of agreement also depend on shared conceptions of justice, we cannot assess a conception of justice "by its distributive role alone" (1971: 6). Because the question of justice is *about* justifiable social practice in the first instance, justice and practicality aren't fundamentally distinct concerns.

4. Practices and Political Philosophy

Rawls is therefore not a state of nature theorist. He isn't asking what people would or ought to agree to in a socially unorganized "state of nature" in the style of Hobbes, Locke, or Kant. Like Rousseau, he begins from established social and political relationships, and constructs justice for and in the light of the sorts of agreement or understandings involved. This thought raises large, as yet unresolved, questions. We mention a few, along with some possible answers.

One question is, What sort of "agreement" is required? Rawls had always said that his general concern is with the structure of a "social practice" (Rawls 1999a: 47), where obligations of individuals are assumed to be a secondary, practice-sensitive matter (Rawls's Natural Duty of Justice, for example, depends on an independent characterization of just institutions). A "social practice" presumably requires more "agreement" than de facto behavioral coordination, say, at least some understood purposes which shape the structure and maintenance of coordination over time. It perhaps also requires *mutuality* in purpose and organization, or even established (if potentially decentralized) *political* arrangements. Rawls mainly suggests that his focal contexts—major domestic institutions and international law and practice—meet sufficient conditions for a social practice, but without characterizing the relevant social preconditions in more general terms. We are left asking, Do global capital markets, or the global economy generally, qualify as part of an international social practice of some kind? Is emergent human rights practice, beyond its role in regulating outright coercion and force (e.g., in "jawboning" and "naming and shaming"), a social practice in the appropriate sense? Even if we can characterize sufficient conditions in these cases as well, the more general necessary and sufficient conditions applicable to all subjects of social justice remain as yet obscure.

A deeper question is, Why social practices at all? Why not just specify our preferred ideal distributional state of affairs (e.g., an equally good life for all)? If the state of affairs is a worthy impersonal ideal, that is just to say that all agents have reason to organize themselves to bring it about, even if that means leaving a state of nature. Why then should any social organization be necessary? A partial answer is as follows. Insofar as politically central concepts of justice are part of "what we owe to each other," this is already to say they are essentially concerned with social or interpersonal relations rather than impersonally valuable distributional states of affairs. In that case, a mere worthy ideal is insufficient. We don't yet have an idea of distributive *justice*, in the relevant sense, until a proposed ideal is attached to some governance expectations concerning others, expectations which are normally conclusive for action and grounded, in the first instance, in "personal" rather than "impersonal" considerations. This is only a partial answer, however, because Scanlon's account of interpersonal

morality doesn't necessarily take us out of the state of nature. Without supplementary argument, it doesn't tell us why social and political relationships should be established. Even when they are established, it doesn't explain how to reason in light of the relationships they are, let alone why we should worry about their publicity and realistic stability.

Indeed, Scanlon's account might pose a significant explanatory challenge. For while Rawls often mentions the vital role of "social unity," "mutual recognition," and "mutual respect," Scanlon's theory shows precisely that such ideas do plausibly apply in a state of nature. By acting in ways I can justify to you, even in the absence of any special relationship, I thereby stand in a valuable relation of "recognition" with you, a relation (to adapt Mill's phrase) of "unity with my fellow creature." If there is a distinct, more deeply social idea of "social unity" or "mutual recognition," the question is what these ideas are and why they should be of central concern to political philosophy.

One answer is the intrinsic force of Rawls's master ideal of the "well-ordered society." Much as Rousseau imagined, people publicly and stably govern their shared arrangements from agreement upon the same principles and in that sense mutually recognize one another (Cohen 2002: 129–31; Freeman 2007a: 184–8). Alas, Rawls's "political turn," in effect, admits that such agreement is too much to hope for and so seeks a feasible "overlapping consensus" marked by a "family" of principles in its stead.

A different (not incompatible) answer is Rawls's Hegelian concern with "reconciliation to our social condition" (Rawls 1999b: 128). Political philosophy, Rawls says, can give us grounds for reasonable hope, encourage efforts of reform, and address dark feelings about human existence itself, simply by describing a "realistic utopia." The values of "social unity" and "mutual recognition" take an essentially social form, then, because the "great evils of past and present" are attributable specifically to political injustice, and because any sufficiently "utopian" but also plausibly "realistic" story about how such evils might be overcome involves stable domestic and international institutions and practices (1999b: 7). We can't reconcile ourselves to a dystopian state of nature, or to the familiar cruelty and callousness of grave political evils. But neither are we reconciled to our social condition by describing a social world beyond the horizon of future human possibility; we need an "achievable social world" (1999b: 6). A radically different global order—for example, a global state, instead of a perhaps deeply reformed international system—might be unavailable to us if only because its prospects are so uncertain (James 2012: ch. 4; Risse 2012).

Although this existential function of political philosophy mainly gains prominence in Rawls's later writings, it harkens back to themes in Rawls's undergraduate thesis (Rawls 2009), and so was arguably at work all along. Still, the goal of reconciliation never supplanted Rawls's own narrower task of addressing more specific impasses within prevailing political argument and culture (Rawls 1999a: 307). And why should it? Political philosophy can intervene at different levels of argument, ranging from concrete matters of policy, to larger institutional design, to master political ideals, to how we should feel about the human social condition. Everything depends on what question we are asking.

In that case, however, political philosophy can, equally, take up the purely intellectual task of *characterizing justice*, in light of our basic convictions, where that might in theory leave aside all such matters of practice and practicality (Cohen 2008). This poses one final question: Is there any such purely intellectual rationale for focusing on realistic social practice?

One starts from the normativity of justice. To situate justice within what we owe to each other already implies a kind of realistic availability. What we owe to others will be limited to what we can do, given not just the general human condition, but the specific available forms of information, reason, and experience that shape our regulatory powers. To be sure, if we focus solely on individual conduct, this arguably excludes important large-scale patterns of interaction and distribution—in domestic society, the global economy, or global climate change—which are beyond the regulative powers of any particular agent taken by him or herself. Must such matters simply be left to fate? Not necessarily. For, it may be argued, we already have domestic and international social practices which give us a good measure of regulatory power over matters of large and even global scale, and which do give rise to collective (non-distributive) responsibilities concerned with what the organized "we" owes to all affected.

Even so, however, realistic availability remains essential. The practices in question can't be purely hypothetical, or otherwise unavailable to us from our current historical position. If they were, applicable principles would not be addressed as moral responsibilities to *us*. If they apply in an ideal world, but are unavailable to us, they would not say what *we* owe to each other in the social relationships we have and could reasonably hope to establish in the conditions of human life as we know them. Accordingly, the natural approach is much as Rawls suggests, following Rousseau. We start from the "political world as we see it" (1999b: 83). We take "men as they are, and laws as they might be," even as possible laws are bound by the "fixed laws and tendencies" of our available social worlds (Rawls 1999b: 11, 17).

Related Topics

Liberalism, Contractarianism, The Difference Principle, The Capability Approach (and Social Justice), Human Rights and Cosmopolitanism, Global Justice and Politics

References

Cohen, G.A., 2008. *Rescuing Justice and Equality*, Harvard University Press.

Cohen, J., 2002. For a Democratic Society. In S. Freeman, ed. *The Cambridge Companion to Rawls*, Cambridge University Press.

Dworkin, R., 1986. *Law's Empire*, Belknap Press of Harvard University Press.

Estlund, D., 1996. The Survival of Egalitarian Justice in John Rawls's Political Liberalism. *Journal of Political Philosophy*, 4(1), pp. 68–78.

Freeman, S., 2007a. *Rawls*, new edition, Routledge.

Freeman, S., 2007b. The Burdens of Public Justification: Constructivism, Contractualism, and Publicity. *Politics, Philosophy & Economics*, 6(1), pp. 5–43.

James, A., 2005. Constructing Justice for Existing Practice: Rawls and the Status Quo. *Philosophy & Public Affairs*, 33(3), pp. 281–316.

James, A., 2012. *Fairness in Practice: A Social Contract for a Global Economy*, Oxford University Press, USA.

Postema, G.J., 1987. "Protestant" Interpretation and Social Practices. *Law and Philosophy*, 6(3), pp. 283–319.

Rawls, J., 1971. *A Theory of Justice*, Harvard University Press.

Rawls, J., 1993. *Political liberalism*, Columbia University Press.

Rawls, J., 1999a. *Collected Papers*, Harvard University Press.

Rawls, J., 1999b. *The Law of Peoples*, Harvard University Press.

Rawls, J., 2001. *Justice as Fairness: a Restatement*, ed. Erin Kelly, Belknap Press of Harvard University Press.

Rawls, J., 2009. *A Brief Inquiry into the Meaning of Sin and Faith: With "On My Religion,"* Harvard University Press.

Rawls, J. and Herman, B., 2000. *Lectures on the History of Moral Philosophy*, Harvard University Press.
Risse, M., 2012. *On Global Justice*, Princeton University Press.
Scanlon, T.M., 1998. *What We Owe to Each Other*, Belknap Press of Harvard University Press.
Weithman, P., 2011. *Why Political Liberalism?: On John Rawls's Political Turn*, Oxford University Press, USA.

Further Reading

Dreben, B., 2002. On Rawls and Political Liberalism. In S. Freeman, ed. *The Cambridge Companion to Rawls*, Cambridge University Press.
Hill, T.E., Jr., 1994. The Stability Problem in *Political Liberalism*. *Pacific Philosophical Quarterly*, 75, pp. 332–52.
Larmore, C., 1999. The Moral Basis of Political Liberalism. *The Journal of Philosophy*, 96(12), pp. 599–625.
Waldron, J., 1994. Disagreements about Justice. *Pacific Philosophical Quarterly*, 75, pp. 372–87.

29

UTILITARIANISM AND CONSEQUENTIALISM

Dale E. Miller

1. Consequentialism: An Overview

Consequentialism is a view about evaluation. As a first approximation, consequentialism can be described as the view that evaluands—the objects of evaluations—are to be judged solely on how well they promote value. "Promoting value" here means promoting the existence of whatever is intrinsically valuable or good for its own sake and the absence of whatever is intrinsically disvaluable or bad for its own sake. Consequentialist views are sometimes also referred to as "teleological" (e.g., Rawls 1999: 21–2), since they appraise evaluands based on how well they promote certain ends, although this invites confusion with the Aristotelian view that evaluands should be appraised on how well they realize their own natural ends.

The moral theory sometimes labeled "classical utilitarianism" exemplifies consequentialist evaluation. The classical utilitarian believes (1) that nothing is intrinsically valuable except happiness and nothing is intrinsically disvaluable except unhappiness, where happiness and unhappiness are defined in terms of pleasure and pain, and (2) that in any situation and for any agent, an action is wrong if there is something else the agent could do instead that would yield more "total net happiness" (Shaw 1999: 10–11). If the "net happiness" that an individual would experience if an action were performed is understood as the amount of future positive happiness that she would enjoy from that point in time "minus" the amount of future unhappiness, then total net happiness is the aggregate or sum of everyone's net happiness. Classical utilitarianism essentially evaluates potential actions twice, first ranking them in terms of how well they promote total net happiness and then judging the action that emerges at the very top of this ranking—or actions, in the case of a "tie for first place"—morally right and the rest morally wrong.

This brief characterization of consequentialism stands in need of further elaboration. Six points in particular are worth noting.

First, one might have a consequentialist view about the evaluation of anything: actions, rules, motives, practices, institutions, cocktails, or umbrellas. The basic consequentialist idea is in this respect entirely general (Parfit 1984: 25; Kagan 2000: 151; Pettit and Smith 2000: 121–2).

Second, nothing has been said so far about what things are good or bad for their own sakes. Consequentialism is compatible with nearly any "axiology" or "theory of the good." I will say more about these theories in the next two sections.

Third, the consequentialist's conception of consequences is much broader than that of causation (Shaw 1999: 13–14). An economical way of expressing the consequentialist conception is to say that an evaluand's consequences include the entire future history of the world that would unfold subsequent to the evaluand's presence. A consequentialist theory of the evaluation of actions, for example, may appraise a given action based on how much that is intrinsically valuable and disvaluable the entire future history of the world would contain were it to be performed. Anyone troubled by this inclusive conception of consequences should bear in mind that consequentialists are generally interested in how the consequences of different evaluands compare. This both necessitates an inclusive conception of consequences and renders harmless some of the implications of this conception that might seem problematic. Suppose that I can either save a drowning swimmer or continue reading by the pool. We cannot compare the results of these two actions without taking into account that the swimmer dies if I keep reading, even though my reading would not be the cause of her drowning. On the other hand, suppose that a swimmer is about to drown on the other side of the world and that nothing that I can do now would prevent this. Since this drowning is unavoidable so far as I am concerned, it will be among the consequences of *all* of my possible actions, and hence it makes no difference to how their consequences compare. Of course, since unavoidable events like this make no difference to consequentialist evaluation, consequentialists have the option of forgoing the economy of "entire future history" talk and framing their conception of consequences in a way that excludes them. The result will be a conception that comes closer to our ordinary one but is still considerably broader than that of "causal consequences."

Fourth, a distinction is sometimes drawn between "direct" and "indirect" consequentialism. Suppose that we are ultimately interested in evaluating actions. A direct consequentialist theory for evaluating actions will appraise individual actions based on their consequences. An indirect theory, in contrast, will appraise evaluands of some other kinds based on their consequences, and will then use the results of this appraisal in setting up criteria for evaluating actions. One might, for instance, compare different sets of rules based on the consequences of their general acceptance, and then morally evaluate actions based on whether they are required, permitted, or forbidden by the set of rules whose general acceptance would maximize value. In light of this distinction, the characterization of consequentialism with which this discussion began needs to be amended. We can say that consequentialist evaluation of evaluands that belong to a class Φ involves judging evaluands that belong to a class Ω based on how well they promote value, where Ω and Φ may or may not be identical. A consequentialist theory of Φ will be one that says that evaluands of that class are to be judged solely on the basis of consequentialist evaluation.

Fifth, "objective" and "subjective" consequentialism are also sometimes distinguished. Objective consequentialism is also called "actual consequence" consequentialism, and subjective consequentialism "expected consequence." An objective consequentialist will think that the best action for an agent to perform is the one that will actually best promote value, for instance, while a subjective consequentialist will think that it is the one with the highest "expected value." The expected value of an action can be thought of as the weighted average of the amounts of value that every possible future history that might follow the action would contain, with the weight attached to each history determined by the probability that it would occur. Except perhaps where objective probabilities are involved, e.g., in cases involving quantum indeterminacies, these probabilities

are someone's subjective probabilities. When the evaluands in question are actions, this someone will typically be the actual agent or an idealized, e.g., better informed or more rational, version of the agent. Otherwise it might be an arbitrary reasonable individual. I will refer frequently to the ideas of promoting or maximizing value in the course of this chapter, and when I do "value" should be read as shorthand for "actual value in the case of objective consequentialism and expected value in the case of subjective."

Finally, consider an "egoistic" moral theory that says (1) that nothing is intrinsically valuable except happiness and nothing is intrinsically disvaluable but unhappiness, and (2) that in any situation and for any agent, an action is wrong if there is something else the agent could do instead that would yield more net happiness for herself. Is this theory consequentialist? The literature is divided on this question. There is a straightforward sense in which the theory evaluates actions based on their consequences, and some philosophers would call it a consequentialist theory on this basis (e.g., Shaw 1999: 17). However, the theory directs agents not to promote value *simpliciter*, but only to promote it within their own lives. The end that the theory directs an agent to promote cannot be characterized without making essential reference to the agent herself, and it is therefore "agent relative" rather than "agent neutral." Some definitions of "consequentialism" incorporate the requirement that the value that agents are to promote be agent neutral (e.g., Howard-Snyder, 1993: 271). Notice that if this egoistic view does not count as properly consequentialist for this reason, neither will views that say that agents are not obligated to promote the well-being of anyone but their compatriots.

2. Theories of the Good

"Consequentialism" and "utilitarianism" may sometimes seem to be synonymous, but as these terms are usually employed today "utilitarianism" refers to a particular family of consequentialist views. Part of what distinguishes utilitarian views from other consequentialist positions is that utilitarians subscribe to a general theory of the good known as "welfarism." According to welfarism, only welfare or well-being is valuable for its own sake, where your level of welfare reflects how well your life is going for you. While not all welfarists are utilitarians, all utilitarians are welfarists.

I call welfarism a "general" theory of the good because there are competing conceptions of welfare. What might be called "pure" conceptions of well-being are commonly divided into three categories: "hedonistic" conceptions, according to which how well off you are depends on how much pleasure you enjoy; "desire-fulfillment" conceptions, according to which your desires—or, in some cases, an idealized version of your actual desires, such as the ones that you would have if you were fully informed—determine what makes your life go better for you; and "objective-list" conceptions, according to which having certain things in your life makes your life go better or worse for you irrespective of your desires (Parfit 1984: 493–502; Griffin 1986; Sumner 1996).

There are stock objections to each of these pure conceptions of welfare. Robert Nozick, for instance, famously criticizes hedonism by pointing out that most people would not want to spend their lives hooked up to a virtual-reality "experience machine," even if they were thinking only about what they would want for themselves (and ignoring how this might affect their families, etc.), despite the fact that someone attached to an appropriately programmed machine that offered the illusion of a life in which her deepest desires were satisfied might enjoy a great deal of pleasure (Nozick 1974: 42–5). As much as I might enjoy a life in which I am wrongly convinced that my deepest

desires are satisfied, one wants to say, such a life would not in fact go well for me. But consider John Rawls's example of a talented mathematician whose overriding desire is to count blades of grass (Rawls 1999: 379–80). It is tempting to say that a life dedicated to acting on this desire (which might not be uninformed or even irrational) could not go well for the mathematician, since it would lack certain things that a life worth living must include whether they are desired or not. Objective-list views, though, can leave us wondering whether having any number of "objective goods" in her life is enough to make a person's life go well *for her* if she neither wants them nor enjoys them. Welfare is inherently "subject-relative," L. W. Sumner notes, and it is hard to explain why unless welfare is subjective (Sumner 1996: 42–80).

Despite these objections, each of these pure conceptions of welfare has its defenders, including Fred Feldman in the case of hedonism (Feldman 2004), John Harsanyi and Richard Brandt in the case of desire fulfillment (Harsanyi 1998: 285–9; Brandt 1998: 88–129), and Brad Hooker in the case of the objective-list approach (Hooker 2000: 37–43). Another way of addressing some of these objections may be to adopt a "hybrid" conception of welfare. For example, Parfit considers the notion that welfare might be identified with pleasure derived from the possession of things that are objectively good (Parfit 1984: 501–2; see also Kagan 2009). Perhaps Martha Nussbaum's version of the "capabilities" conception of well-being could also qualify as a hybrid conception, since she acknowledges that her list of the essential human capabilities would need to be revised if people's informed desires did not converge on it (Nussbaum 2000: 111–66, esp. 151–2), although the justificatory role that she assigns to desire in this account is, by her own admission, so "ancillary" that the account is frequently construed on objective-list lines (e.g., Adler and Posner 2008: S256).

One prominent consequentialist who is not a welfarist is G. E. Moore. (That Moore calls himself a utilitarian reflects the fact that this term was sometimes used more broadly in the early twentieth century than it is today.) The highest goods, according to Moore, are "personal affection and the appreciation of what is beautiful in Art or Nature" (Moore 1903: 188). Moore conceives of these goods as "organic unities" that combine cognitive states, emotions, and the objects of affection or appreciation themselves (Moore 1903: 189–225). Insofar as these wholes involve pleasure, Moore insists, the pleasure itself is of limited value. In fact, Moore maintains that a beautiful world that no one ever saw or enjoyed would still possess some intrinsic value (Moore 1903: 83–4). A more recent non-welfarist consequentialist is Hooker, who takes virtue to have intrinsic value in addition to well-being (Hooker 2000: 33–7).

3. Further Questions about Welfare

There are two questions that anyone who counts welfare among the things that are good for their own sakes has to confront, whether or not she is a welfarist. One is whether equal amounts of welfare are always equally valuable. The other is whether it is possible to make comparisons between the levels of welfare enjoyed by different individuals.

Utilitarians maintain that the answer to the first question is yes, i.e., that a given "amount" of welfare is equally valuable no matter who enjoys it. J. S. Mill typifies this kind of thinking when he writes that "equal amounts of happiness are equally desirable, whether felt by the same or by different persons" (Mill 1969: 258). Different people may derive entirely different amounts of welfare from a given quantity of some instrumental good such as money, and instrumental goods will tend to exhibit "diminishing

marginal utility": the more of a good someone has, the less that giving her an additional given amount will tend to raise her well-being. But utilitarians believe that the value of a given amount of welfare itself is the same no matter who enjoys it or how well off she already is. In consequence, utilitarians believe that the goodness of a state of affairs or an entire future history depends entirely on the total amount of welfare it contains. While some people's use of the term may deviate from this rule, to call someone a utilitarian usually just means that she is a consequentialist who is a welfarist and takes equal amounts of welfare to be equally valuable.

The "prioritarian," in contrast, maintains that equal amounts of welfare may differ in value depending on who enjoys them. Typically, prioritarians assert that "Benefitting people matters more the worse off those people are" (Parfit 1997: 213). In other words, welfare itself has diminishing marginal value; the additional value generated by a given increase in an individual's level of welfare is inversely proportional to the level of welfare that she already enjoys.

Utilitarianism and prioritarianism both presuppose that it is possible to compare people in terms of how much well-being they are enjoying. Whether it is possible to make these comparisons, and if so how, will depend on the particular conception of well-being at issue, although on almost every conception there are epistemic hurdles that would appear to confound attempts to make them with any precision. In general, economists have been more skeptical than philosophers about the possibility of making interpersonal comparisons. In economics, "utility" has come to be understood as an index of preference satisfaction. Economists frequently seem to regard utility so understood as a specific conception of welfare of the desire-fulfillment variety (Broome 1991). The received view, however, is that individuals' utility functions must either be ordinal and unquantifiable or must be Von Neumann-Morgenstern cardinal functions with arbitrary units, neither of which allow for interpersonal comparisons (Robbins 1945: 136–41; Barrett and Hausman 1990; Hausman 1992: 57–62). Some economists, most notably Harsanyi, do think otherwise (Harsanyi 1955, 1982: 49–52; see also Ng and Singer 1981).

4. Consequentialist Moral Theories

When consequentialism figures in discussions of social and political philosophy, it is usually in the form of a consequentialist moral theory. Many defenders of such theories maintain that consequentialism supplies the criteria that distinguish right from wrong actions for every situation and agent. A moral theory that does this will obviously have implications for how individuals should live together in society and how public officials or states ought to behave.

The most straightforward consequentialist moral theories are versions of act consequentialism, according to which an action is right if it is "optimific" or value maximizing, i.e., it would better promote value than anything else the agent could do instead. Otherwise, it is wrong. Classical utilitarianism epitomizes act consequentialism, but despite its name, which if any of the major historical figures in the utilitarian tradition actually held the view is far from clear. The utilitarians J. J. C. Smart (1956, 1973), R. M. Hare (1981), and Peter Singer (1972) are among the most prominent contemporary proponents of act consequentialism.

Given act consequentialism's prominence as a moral theory, it is no surprise that numerous objections have been raised against it. One obvious criticism is that the results

of our constantly trying to calculate the consequences of our actions would be very bad. Not only would we frequently face the paralysis of analysis, most of us are not well-equipped to make these calculations. Sophisticated act consequentialists point out in reply that the claim that the right action is the one that will have the best consequences does not entail that agents should invariably decide what to do on the basis of explicit calculations. They recommend instead that we employ a very different "decision procedure" in most cases, such as following a set of rules that are shown by experience to steer agents to actions that will usually have good consequences (e.g., Hare 1981: 25–64).

"Epistemic" objections to act consequentialism charge that the near impossibility of knowing whether an action would maximize value or not, even after the fact, somehow constitutes an obstacle to the theory (e.g., Howard-Snyder 1997; Lenman 2000; D. E. Miller 2003). Subjective act consequentialism, according to which the right action is the one with the highest expected value, is less vulnerable than objective act consequentialism to epistemic objections. Of course, it is fair for the act consequentialist to ask why we should suppose that it should always be possible for us to know what is right; perhaps some moral uncertainty is just part of the human condition.

One further line of objection to act consequentialism is that its implications are sharply at odds with our "considered moral judgments" or pre-theoretical moral convictions. Critics of the theory have charged that it can require unjust actions, as in a scenario in which offering false testimony against a person known to be innocent of a crime is necessary to prevent gross public disorder—panic, lynch mobs, etc. (McCloskey 1965). They have also argued that with so many people in the world in dire poverty, act consequentialism demands extreme personal sacrifices that our "ordinary morality" tells us are supererogatory, not obligatory (e.g., Mulgan 2001: esp. 3–49). While the act consequentialist can quibble with some of the particular examples that are used to illustrate this point, perhaps contending that they are highly unrealistic and that we cannot trust our pre-theoretical convictions about cases entirely outside of our experience, ultimately it seems that she must deny that strict conformity to our considered moral judgments is a necessary feature of an acceptable moral theory (Smart 1973: 67–73; Singer 1974; Hare 1981: 10–15).

Rule consequentialism is not the only alternative to act consequentialism within the family of consequentialist moral theories—"motive" consequentialism (Adams 1976) and "virtue" consequentialism (Driver 2001) are among the other possibilities—but it is the alternative that has received the most attention from philosophers. Rule consequentialism is a type of indirect consequentialism. The rule consequentialist evaluates moral codes or sets of moral rules based on their consequences—or, more precisely, the consequences of their being generally obeyed or generally accepted—and evaluates actions by reference to the "ideal" moral code, i.e., actions that this moral code would forbid are wrong (cf. R. Miller 2009). Brandt (1979), Harsanyi (1998), and Hooker (2000) are among the best known contemporary rule consequentialists. Brandt and Harsanyi are rule utilitarians.

While act consequentialists can recommend that we employ a rule-based decision procedure, they must regard an action that fails to maximize value as wrong whether the agent properly follows recommended decision procedure or not. (No decision procedure that mere humans could employ could possibly guide an agent to the optimific action in every situation.) In contrast, rule consequentialism holds that an action that conforms to the best moral code is right regardless of the particular act's upshot.

The objections to act consequentialism that were discussed above may well have less force against rule consequentialism. It might be easier to assess the actual consequences

of the general acceptance of a given rule than of an individual action, at least as long as we can arrive at a probability distribution of the different possible results of adherence to the rule in individual cases, and Hooker in particular has called attention to the close conformity of rule consequentialism—especially his own version of rule consequentialism—to our considered moral judgments (Hooker 2000: 9–19). But other objections have been leveled against rule consequentialism.

One of these objections is sometimes labeled the "collapse" objection. It says that act and rule consequentialism are "extensionally equivalent," i.e., that they will always agree about any given action's moral standing. The thought behind this objection is that the moral code that it would maximize value for people generally to obey will require the value-maximizing action in every situation, either because this code will comprise a specific rule for virtually every conceivable situation, a rule that requires the value-maximizing action (Lyons 1965: 62–139), or because this code will comprise only a single rule like "always maximize value" or "always produce the best consequences" (Smart 1973: 10–12). This objection has little force against versions of rule consequentialism that say that the ideal moral code is the one that it would be best for people to accept as opposed to the one that it would be best for them to obey. Our psychology is too limited for us to accept a code that contains a massive number of rules or individual rules that are exceptionally complicated. And while we might be able to accept the code that contains only a rule that tells us to produce the best consequences, for a variety of reasons (our epistemic limitations, problems of interpersonal coordination, etc.) the consequences of our accepting a code with more specific rules would almost certainly be better (Brandt 1998: 271–7; Hooker 200: 94–5).

Another objection to rule consequentialism is the so-called "incoherence" or "rule worship" objection. This objection claims that there is an inconsistency between rule consequentialism's account of right and wrong and a necessary premise in any argument for that account. As Smart writes, with rule utilitarianism in particular in mind: "[T]he rule-utilitarian presumably advocates his principle because he cares about human happiness; why then should he advocate abiding by a rule when he knows it will not in the present case be most beneficial to abide by it?" (Smart 1973: 10; see also D. E. Miller 2011). This is a serious objection. Hooker attempts to deal with it by arguing for rule consequentialism based on its close conformity with our considered moral judgments, as opposed to an "overarching commitment to maximize the good" (Hooker 2000: 99–102), but this leads to a different set of worries (D. E. Miller 2000).

One common feature of consequentialist moral theories is that they account for the rightness or wrongness of actions in terms of the promotion of value. This is what is meant by the claim that in consequentialist moral theories "the good is prior to the right," as opposed to deontological theories in which "the right is prior to the good" (e.g., Rawls 1999: 26–30). Because they make the good prior to the right, consequentialist moral theories cannot, on pain of circularity, have axiologies that refer to the rightness or wrongness of actions, e.g., that say that the performance of right actions has intrinsic value. They must, as Rawls puts it, "specify the good independently of the right."

5. Consequentialism and Public Morality

Even if no consequentialist theory can plausibly claim to govern the entire moral sphere, consequentialist thinking about morality might still be important in the context of social and political philosophy. Some moral theorists who deny that consequentialism

can account for "private" morality still maintain that it offers the best account of "public" morality, i.e., the morality of the state and/or public officials. Robert E. Goodin, who is perhaps the best known exponent of this way of thinking, describes utilitarianism as "truly compelling" in the role of "a guide to public … conduct," adding that in this role "virtually all of its vices—all of the things that make us wince in recommending it as a code of personal morality—loom instead as considerable virtues" (Goodin 1995: 8; see also Nagel 1978).

Goodin offers several reasons why utilitarian considerations loom especially large in public morality. Foremost among these is the impersonality of political decision making. States are free of personal attachments, and we expect public officials to set their personal attachments aside while they are acting in an official capacity. Further reasons that he adduces include the fact that we expect political decisions to be made on the basis of calculation rather than sentiment and on the basis of consequentialist reasoning rather than adherence to precepts whose observance yields no benefits.

In contrast, Stuart Hampshire, writing near the end of the Vietnam War era, contends that while the general acceptance of a utilitarian public morality once contributed to social progress, it has become "an obstruction" (Hampshire 1978a: 1): "Persecutions, massacres, and wars have been coolly justified by calculations of the long range benefit to mankind … The utilitarian habit of mind has brought with it a new abstract cruelty in politics, a dull destructive political righteousness" (Hampshire 1978a: 4; see also Hampshire 1978b).

Of course, one might question whether the problems that Hampshire points to, if real, are inherent in a utilitarian public morality or are, rather, the product of the misapplication of such a morality, i.e., of poor calculations. Indeed, Goodin claims to share Hampshire's "dismay" at utilitarianism's "current tendencies" and to be embarked on a project of rehabilitating the view. This project involves, in part, an argument meant to show that rule utilitarianism provides a more tenable foundation than act utilitarianism for public morality (Goodin 1995: 16–18, 60–77).

While there are different ways of characterizing the so-called "dirty hands problem," one of the best known presupposes that public morality in particular is utilitarian or at least consequentialist in nature. On this characterization, the essence of this problem is that public officials are subject to special moral requirements in virtue of their roles, requirements that they may sometimes be able to satisfy only by violating other moral requirements that apply to everyone, all of the time. While their special role obligations require public officials to produce good outcomes, at least for the members of their own political community, the universal requirements are deontological moral rules that prohibit certain courses of action *tout court*. When a prohibited course of action would yield a better outcome than would any that are permitted, the public official purportedly faces a dilemma in which everything she can do is wrong. As Michael Walzer writes,

> [A] particular act of government … may be exactly the right thing to do in utilitarian terms and yet leave the man who does it guilty of a moral wrong. … If on the other hand he remains innocent, … he not only fails to do the right thing (in utilitarian terms), he may also fail to measure up to the duties of his office (which impose on him a considerable responsibility for consequences and outcomes).
>
> (Walzer 1973: 161)

Note that this way of characterizing the dirty hands problem rejects the commonly held view that in every situation at least one action that it is possible for the agent to perform must be no worse than any other, morally speaking, and that any such action cannot be wrong. According to Gerald Gaus, this way of formulating the problem "commits one to an unorthodox meta-ethics" and "undermines the action-guidance function of moral duties" (Gaus 2003: 173).

6. Consequentialism, Justice, and Rights

Justice has always been a central concept in social and political philosophy. Since the modern period it has been closely linked with another concept, that of moral (as opposed to legal) rights. Principles of justice are commonly thought of as moral rules that entail that people have certain especially important rights. Consequentialist theories differ in how they can account for the existence of justice and rights.

There is a significant sense in which act utilitarianism cannot account for their existence at all. If act utilitarianism is the correct theory of morality then justice and rights as we usually think of them, as considerations that have some—indeed considerable—moral weight, are chimeras. Rights, as Ronald Dworkin has famously observed, serve as "trumps" over collective goals (Dworkin 1977: xi). Act utilitarianism, however, leaves no room for such trumps. Maximizing total net happiness is, in Dworkin's sense, a collective goal, and act utilitarianism entails that an action is justified if and only if it makes a greater contribution to the realization of this goal—even an infinitesimally greater contribution—than any competing action. So according to act utilitarianism rights can never trump the maximization of happiness but are invariably trumped by it. And if principles of justice are moral rules that entail the existence of rights, then the same reasoning shows that the act utilitarian must hold that an agent is justified in violating a principle of justice whenever doing so would yield even slightly more total net happiness than could be produced through complying with the principle. Rights this weak are no rights at all, and likewise principles of justice.

This is not quite to say that rights and justice can play no role in act-utilitarian thought whatsoever, but at most, it seems, they can serve as heuristic devices, i.e., as parts of a decision procedure. Obeying certain principles of justice and thereby treating people as if they have certain rights might turn out to be a useful strategy for selecting actions that would maximize total net happiness (Hare 1981: 147–64). Relegating justice and rights to a decision procedure means depriving them of any moral weight of their own, however; on occasions when violating someone's rights would maximize total net happiness, the act utilitarian will not regard the rights-violation as a reason not to perform the action or a reason to consider the action wrong.

In contrast, many forms of indirect consequentialism can more easily account for the existence of justice and rights. Even rule utilitarianism, which shares act utilitarianism's welfarist axiology, can do so. The rule utilitarian is not committed to the proposition that every moral rule that entails that someone has a certain right ought to be overridden whenever this would maximize total net happiness. For the rule utilitarian, in other words, an action that denies someone what she wishes or imposes a loss on her might not be justified even though it is happiness maximizing. A related example is the utilitarian theory of justice that Rawls considers in *Theory of Justice*, according to which "society is rightly ordered, and therefore just, when its major institutions are arranged so as to achieve the greatest net balance of satisfaction summed over all the individuals

belonging to it" (Rawls 1999: 20). While he takes Justice as Fairness to offer a better conception of the concept of justice, Rawls does not deny that this utilitarian view is a theory of justice.

Nothing has been said so far in this section about forms of act consequentialism other than act utilitarianism. It may appear that any act-consequentialist theory will be in the same boat with act utilitarianism, since any such theory will say that actions that violate moral rights and principles are justified whenever they would promote value even slightly better than any alternatives, and this renders rights and justice otiose. But while there is considerable truth in this thought, an act consequentialist could in principle find a place for rights and justice in her axiology. This would give them a kind of moral significance that they lack when they are relegated to the status of heuristic devices. Robert Nozick describes, albeit only to illustrate what he takes to be a mistaken way of thinking about rights, a view according to which the right action is the one that "minimizes the total (weighted) violations of rights"; he understandably refers to this view as a "utilitarianism of rights," although strictly speaking it is not utilitarian at all (Nozick 1974: 28). Fred Feldman defends an act-consequentialist theory in which the goodness of outcomes depends upon both how much happiness or unhappiness people are enjoying in them and the extent to which their happiness or unhappiness is deserved (Feldman 1995). He describes this theory's axiology as "justice adjusted hedonism."

7. The Implications of Consequentialism

The discussion in the previous section was entirely formal, saying nothing about what substantive claims about justice and rights an indirect consequentialism would underwrite. Neither this question nor the broader question of which it is a part, the question of what substantive implications consequentialist moral theorizing has in general, admits of any easy answer. Even if we ignore the philosophical differences between consequentialists and limit our attention to one particular consequentialist theory, including one particular axiology, there could still be significant disagreements over the implications of this theory because of empirical disagreements about what the consequences of different evaluands (actions, rules, institutions, etc.) would be.

Some of the most important of these empirical disagreements concern human nature, e.g., disagreements about how much of the fruits of their own labors people must be allowed to keep in order to incentivize work or about the degree to which wealth exhibits "diminishing marginal utility." Because of empirical differences like these, there are consequentialists of every ideological stripe.

Many supporters of the free market base their views on the desirable consequences of the unfettered operation of the "invisible hand," for example. This is especially, if not exclusively, true of free-market thinkers who come from the field of economics, e.g., Milton Friedman and his followers in the Chicago school or Friedrich Hayek and other "Austrian" economists (see Barry 1987: 19–80). Other consequentialists reach much more egalitarian conclusions. Goodin, for instance, argues not only that all citizens should receive "basic income" payments—an idea to which some free-market thinkers also give grudging support (e.g., Murray 2007)—but also (at least *ceteris paribus*) for policies that would aim at producing more equal distributions of goods that satisfy "relative needs" by preventing anyone from having too much of them (Goodin 1995: 228–61). Examples of such goods include housing, certain kinds of education, legal advice, and even food and clothing. Still others have been led by consequentialist reasoning all

of the way to socialism, e.g., J. A. Hobson (see Weinstein 2007: 164–94). And Peter Singer has notoriously argued that utilitarianism demands that states and individuals alike make far greater efforts to eliminate radical poverty in the Third World than have been undertaken previously (e.g., Singer 1993: 218–46).

Another example of the political diversity of consequentialist thinking can be found in nineteenth-century Britain, where James Fitzjames Stephen's *Liberty, Equality, Fraternity* represents a conservative rebuttal to Mill's defense of liberalism in *On Liberty* but shares the latter's utilitarian foundations (Stephen 1993; Mill 1977).

Related Topics

Mill, Late Nineteenth- and Early Twentieth-Century British Political Thought, Social Choice Theory, Rational Choice Theory

References

Adams, Robert (1976) "Motive Utilitarianism," *Journal of Philosophy* LXXIII: 467–81.

Adler, Matthew and Eric A. Posner (2008) "Happiness Research and Cost-Benefit Analysis," *Journal of Legal Studies* 37: S253–292.

Barrett, Martin, and Hausman, Daniel (1990) "Making Interpersonal Comparisons Coherently," *Economics and Philosophy* 6: 293–300.

Barry, Norman (1987) *On Classical Liberalism and Libertarianism*, New York: St. Martin's.

Brandt, Richard (1998) *A Theory of the Good and the Right*, Amherst, NY: Prometheus [originally published Oxford: Oxford University Press, 1979].

Broome, John (1991) "Utility," *Economics and Philosophy* 7: 1–12.

Driver, Julia (2001) *Uneasy Virtue*, Cambridge: Cambridge University Press.

Dworkin, Ronald (1977) *Taking Rights Seriously*, Cambridge, MA: Harvard University Press.

Feldman, Fred (1995) "Adjusting Utility for Justice: A Consequentialist Reply to the Objection from Justice," *Philosophy and Phenomenological Research* LV: 567–85.

—— (2004) *Pleasure and the Good Life: Concerning the Nature, Varieties, and Plausibility of Hedonism*, Oxford: Clarendon.

Gaus, Gerald (2003) "Dirty Hands," in R. G. Frey and Kit Wellman (eds.) *A Companion to Applied Ethics*, Oxford: Blackwell, pp. 169–79.

Goodin, Robert E. (1995) *Utilitarianism as a Public Philosophy*, Cambridge: Cambridge University Press.

Griffin, James (1986) *Well-Being: Its Meaning, Measurement and Moral Importance*, Oxford: Clarendon.

Hampshire, Stuart (1978a) "Morality and Pessimism," in Stuart Hampshire (ed.) *Public and Private Morality*, Cambridge: Cambridge University Press. Originally published in the *New York Review of Books* 19 (January 25, 1973), pp. 26–33.

—— (1978b) "Public and Private Morality," in Stuart Hampshire (ed.) *Public and Private Morality*, Cambridge: Cambridge University Press.

Hare, R. M. (1981) *Moral Thinking: Its Levels, Method and Point*, Oxford: Clarendon Press.

Harsanyi, John (1955) "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility," *The Journal of Political Economy* LXIII: 309–21.

—— (1982) "Morality and the Theory of Rational Behaviour," in Amartya Sen and Bernard Williams (eds.) *Utilitarianism and Beyond*, Cambridge: Cambridge University Press. Originally published in *Social Research* 44 (1977).

—— (1998) "A Preference-Based Theory of Well-Being and a Rule-Utilitarian Theory of Morality," in Werner Leinfellner and Eckehart Köhler (eds.) *Game Theory, Experience, Rationality: Foundations of Social Sciences, Economics and Ethics*, Dordrecht: Kluwer, pp. 285–300.

Hausman, Daniel (1992) *The Separate and Inexact Science of Economics*, Cambridge: Cambridge University Press.

Hooker, Brad (2000) *Ideal Code, Real World: A Rule-Consequentialist Theory of Morality*, Oxford: Clarendon Press.

Howard-Snyder, Frances (1993) "Rule Consequentialism Is a Rubber Duck," *Philosophical Quarterly* 30: 271–8.

—— (1997) "The Rejection of Objective Consequentialism," *Utilitas* 9: 241–8.
Kagan, Shelly (2000) "Evaluative Focal Points," in Brad Hooker, Elinor Mason, and Dale E. Miller (eds.) *Morality, Rules, and Consequences*, Edinburgh: Edinburgh University Press, pp. 134–55.
—— (2009) "Well-Being as Enjoying the Good," *Philosophical Perspectives* 23: 252–72.
Lenman, James (2000) "Consequentialism and Cluelessness," *Philosophy and Public Affairs* 29: 342–70.
Lyons, David (1965) *Forms and Limits of Utilitarianism*, Oxford: Clarendon.
McCloskey, H. J. (1965) "A Non-Utilitarian Approach to Punishment," *Inquiry* 8: 249–63.
Mill, John Stuart (1969) *Utilitarianism*, in J. M. Robson (ed.) *Essays on Ethics, Religion and Society: Collected Works of John Stuart Mill* vol. X, Toronto: Toronto University Press, pp. 205–59.
—— (1977) *On Liberty*, in J. M. Robson (ed.), *Essays on Politics and Society: Collected Works of John Stuart Mill* vol. XVIII, Toronto: Toronto University Press, pp. 213–310.
Miller, Dale E. (2000) "Hooker's Use and Abuse of Reflective Equilibrium," in Brad Hooker, Elinor Mason, and Dale E. Miller (eds.) *Morality, Rules, and Consequences*, Edinburgh: Edinburgh University Press, pp. 156–78.
—— (2003) "Actual-Consequence Act Utilitarianism and the Best Possible Humans," *Ratio* XVI: 49–62.
—— (2011) "Mill, Rule Utilitarianism, and the Incoherence Objection," in Ben Eggleston, Dale E. Miller, and David Weinstein (eds.) *John Stuart Mill and the Art of Life*, New York: Oxford University Press, pp. 94–116.
Miller, Richard (2009) "Actual Rule Utilitarianism," *The Journal of Philosophy* CVI: 5–28.
Moore, G. E. (1903) *Principia Ethica*, Cambridge: Cambridge University Press.
Mulgan, Tim (2001) *The Demands of Consequentialism*, Oxford: Clarendon Press.
Murray, Charles (2007) "Guaranteed Income as a Replacement for the Welfare State," Foundation for Law, Justice and Society policy brief, available online at http://www.fljs.org/uploads/documents/Murray.pdf, visited 7 November 2010.
Nagel, Thomas (1978) "Ruthlessness in Public Life," in Stuart Hampshire (ed.) *Public and Private Morality*, Cambridge: Cambridge University Press, pp. 75–91.
Ng, Yew-Kwang, and Singer, Peter (1981) "An Argument for Utilitarianism," *Canadian Journal of Philosophy* 11: 229–39.
Nozick, Robert (1974) *Anarchy, State, and Utopia*, New York: Basic.
Nussbaum, Martha (2000) *Women and Human Development: The Capabilities Approach*, Cambridge: Cambridge University Press.
Parfit, Derek (1984) *Reasons and Persons*, Oxford: Clarendon.
—— (1997) "Equality and Priority," *Ratio* n.s. X: 202–21.
Pettit, Philip and Smith, Michael (2000) "Global Consequentialism," in Brad Hooker, Elinor Mason, and Dale E. Miller (eds.) *Morality, Rules, and Consequences*, Edinburgh: Edinburgh University Press, pp. 121–33.
Rawls, John (1999) *Theory of Justice*, rev. ed., Cambridge, MA: Harvard University Press.
Robbins, Lionel (1945) *An Essay on the Nature and Significance of Economic Science*, 2nd ed., London: MacMillan.
Shaw, William H. (1999) *Contemporary Ethics: Taking Account of Utilitarianism*, Malden, MA: Blackwell.
Singer, Peter (1972) "Is Act Utilitarianism Self-Defeating?," *Philosophical Review* 81: 94–104.
—— (1974) "Sidgwick and Reflective Equilibrium," *The Monist* 58: 490–517.
—— (1993) *Practical Ethics*, Cambridge: Cambridge University Press.
Smart, J. J. C. (1956) "Extreme and Restricted Utilitarianism," *Philosophical Quarterly* 6: 344–54.
—— (1973) "An Outline of a System of Utilitarian Ethics," in *Utilitarianism: For and Against*, Cambridge: Cambridge University Press, pp. 3–74.
Stephen, James Fitzjames (1993) *Liberty, Equality, Fraternity*, Indianapolis: Liberty Fund.
Sumner, L. W. (1996) *Welfare, Happiness, and Ethics*, Oxford: Clarendon.
Walzer, Michael (1973) "Political Action: The Problem of Dirty Hands," *Philosophy and Public Affairs* 2: 160–80.
Weinstein, David (2007) *Utilitarianism and the New Liberalism*, Cambridge: Cambridge University Press.

Further Reading

Henry Sidgwick's *The Methods of Ethics* (1874) contains one of the most important and influential defenses of consequentialist morality. Bernard Williams's "A Critique of Utilitarianism" (1973) raises a number of much-discussed objections to the theory, including the worry that from a utilitarian standpoint it might be

desirable for the state to be run by elites who accept utilitarianism while the general public does not. In the introduction to the volume *Utilitarianism and Beyond* that Williams wrote with Amartya Sen, this possible upshot of the theory is labeled "Government House utilitarianism." R. M. Hare's *Essays on Political Morality* (1989) contains a selection of readings on a variety of political topics by the twentieth century's leading act-utilitarian theorist. Also noteworthy is his essay "Reasons of State" in *Applications of Moral Philosophy* (1972). The essays collected in Richard Brandt's *Morality, Utilitarianism, and Rights* (1992) offer insights into the social-political implications of rule utilitarianism. Peter Singer's *The President of Good and Evil* (2004) is something unusual: a detailed analysis by a philosopher of the actions and rhetoric of a specific political figure (President George W. Bush). D. H. Hodgson's *Consequences of Utilitarianism: A Study in Normative Ethics and Legal Theory* (1967) explores some of the implications of utilitarian thinking for the law.

30
PERFECTIONISM

Steven Wall

The label "perfectionism" attaches to a family of views in political philosophy. The members of the family differ in substantive content, but they are united in their commitment to two very general, and at first pass, quite plausible claims. The first claim is one in value theory. Some human lives are better lived lives than others. On this claim, a maximally well lived human life, if it were possible, would be a perfect life, even if few or no human beings achieve it. (Some perfectionists deny that a perfect human life is possible. The realization of some excellences precludes the realization of others, but they still accept that some human lives are better lived than others.) The second claim is one in political morality. The state should promote, deliberately and actively, well lived human lives over less well lived human lives. Acceptance of this second claim carries with it the rejection of the currently popular liberal principle of state neutrality—the principle that the state should aim to be neutral among rival conceptions of the good life, at least insofar as these conceptions themselves are consistent with neutral principles of justice. The acceptance of the second claim also, as we will see, poses a challenge to those who favor a principled rejection of state paternalism. The two claims raise a number of important issues, and each is subject to a range of objections. Some of these issues and objections will be discussed in this chapter.

1. The Human Good

Political perfectionists hold that the state should promote well lived human lives. These are lives that consist of valuable experiences, states of being, attitudes, activities, goals and accomplishments. But how should the value of these elements themselves be understood? Unsurprisingly, a number of different answers to this question have been defended by perfectionist writers, several of which will be briefly discussed in this section.

A traditional perfectionist view of the human good appeals to the development of human nature. Such a view was originally articulated within a teleological metaphysics. If human nature has a telos (a normative end), then the good for human beings plausibly consists in becoming what they ought to be. Yet Aristotelian Perfectionism, as it may be termed, can be freed from teleological doctrines. The human good can be identified with the development of a set of properties that are taken to define human nature (Hurka 1993). The greater the development of these properties, the more valuable the human life is. Those who achieve excellence in activities that manifest the properties of human nature lead better lives than those who achieve little in these activities.

Aristotelian Perfectionism offers an account of a well lived human life. But if it is freed from teleological doctrines, then it may have difficulty accounting for the good of

those who lead valuable lives. A well lived human life, or so it may be thought, should be good *for* the person whose life it is. And a life very high in achievement and self-development could be very low in well-being, even if achievement and success in one's efforts at self-development, contribute to well-being. The concept of human well-being (what is good for a human being) is often construed in subjective terms. A human being is well off if he has pleasurable experiences or if he gets what he wants out of life, but objective conceptions of the concept are also available. These conceptions hold that certain states of being, activities and pursuits enrich a human being's life not simply because they are desired by him, but because they are worthy of being desired. A number of perfectionist writers have offered objective accounts of human well-being. We need not discuss the details of such accounts here. The key idea is that certain goods are taken to have objective value and a human life high in well-being must be one that engages with these goods in the right way.

Objective accounts of human well-being need to explain why certain activities, pursuits, etc. are worthwhile and others are not. They also need to explain how these activities and pursuits can enrich the lives of those who engage with them. Many such explanations have been offered. For example, Stephen Darwall, who is not a political perfectionist, but is a proponent of an objective account of human well-being, has argued that "the best life for a person (in terms of welfare) is one involving activities that bring her into appreciative rapport with various forms of agent-neutral value, such as beauty, the worth of living beings, and so on" (Darwall 2002: 17). On this view, it is not just the realization of objective value by a person, but the felt appreciation by that person that she is engaging with objective value that accounts for its benefit to her.

The political perfectionist thus appears to confront a choice. Granted: the state should promote well lived human lives. But, as we have seen, these lives can be characterized either in terms of self-development and the achievement of excellence, as versions of Aristotelian Perfectionism suggest, or in terms of what makes a life go well for the person whose life it is, as objective accounts of human well-being suggest. The choice may be a false one, however. A well lived human life may be a choiceworthy human life, where a choiceworthy human life designates one that is led in a way that gives due regard both to self-development and the realization of excellence and to the goodness of the life for the person in question. This formulation is not terribly precise, as it leaves open what can qualify as "due regard" for each of the two dimensions. But the general idea is clear enough. A life high in well-being could be a less well lived human life than one that was lower in well-being and a life high in achievement and excellence could be a less well lived human life than one lower along this same dimension. Each of the two dimensions is a factor that must be taken into account in assessing the choiceworthiness of a human life.

There are at least two considerations that count in favor of understanding well lived human lives in terms of choiceworthy human lives so construed. The first consideration appeals to the deliberative standpoint from which we make decisions about how to lead our lives. Typically, people decide to pursue some project or enter into some relationship because they believe that doing so would be valuable or worthwhile. And they do not believe, at least not typically, that the value of the project or the relationship is strictly a function of how much it contributes to their own well-being. These beliefs, moreover, are perfectly rational. Indeed, the value of some projects and relationships cannot be realized if one engages in them out of a concern for one's own well-being. People's rational concerns extend beyond themselves. Now assume, as is plausible, that

the value of a human life is, in part, determined by how well it succeeds in the rational aims it has set for itself. Then the value of a human life will not simply be a function of its level of well-being. People rationally care that their lives go well, but this is not the same as caring that their lives are high in well-being (Scanlon 1998; Raz 2004).

The second consideration that counts in favor of construing objectively valuable human lives in terms of choiceworthy human lives appeals to cases. Consider a life of moral perfection—the life of a moral saint. Such a life may be choiceworthy, but it is possible that it is not because it involves too much self-denial (Wolf 1982). Similarly, lives of great nonmoral achievement could fall short because they lack other elements necessary for a well lived life. Consider the life of a great novelist, but one that contains much hardship and suffering. Perhaps this life is a good life, but not as good as it could be since it contains not enough that is personally rewarding or fulfilling. The point is well expressed in the language of choiceworthiness. The life of a moral saint, or a novelist who achieves greatness, may be admirable, but it could fail to be a choiceworthy human life.

Even if the perfectionist state ought to promote choiceworthy human lives, there are significant limits to what it may do in this capacity. A choiceworthy human life consists of successful actions. And these actions trivially must be done by the actors themselves. The state can help to secure the conditions that make well lived human lives possible and it can help to sustain a social environment in which valuable pursuits are favored. But beyond this there is not much it can do. Even the best run perfectionist state cannot guarantee that its subjects will lead good lives.

2. State Neutrality

People disagree as to what counts as a valuable human life. This disagreement is often deep and not easily resolved. How should the state respond to it? In recent political theory, one answer has attracted a fair amount of support. The state should aim to be neutral among rival views of the good. The envisioned neutrality is circumscribed in various ways, for no one really thinks that the state can avoid altogether taking any stand on questions of value. At a minimum, the state must enforce justice and so it cannot aim to treat just and unjust conceptions of the good evenhandedly.

There are different characterizations of the principle of state neutrality. I shall start with a statement of the principle that has become influential. Call it the liberal principle of state neutrality.

> LSN: It is impermissible for the state to intend to favor or promote any permissible ideal of a good human life over any other permissible ideal of a good human life, or to give greater assistance to those who pursue it (on the grounds that it is a better or more valuable ideal of a good human life).

A permissible ideal of a good human life, for the purposes of this principle, is an ideal of a good human life that is consistent with the requirements of justice for a modern democratic society, where the requirements of justice are not themselves founded on or tied to any particular ideal of a good human life (Rawls 1993). The details and implications of this principle need not detain us here. Quite clearly, LSN is a principle that the political perfectionist will reject. It straightforwardly conflicts with the perfectionist claim that the state should promote, deliberately and actively, valuable human lives over less valuable human lives.

LSN can be advanced as a bedrock principle of political morality, one that does not rest on or require any deeper rationale. Typically, however, the principle is not presented in these terms. It is defended on the grounds that it is an appropriate response to the fact of reasonable pluralism in modern democratic societies. Roughly, the idea behind reasonable pluralism is that people *not unreasonably* affirm different conceptions of a good or choiceworthy human life. Naturally, reasonable pluralism contrasts with unreasonable pluralism. Two senses of unreasonableness need to be distinguished. One is moral. A conception of the good that directs a person to treat others unjustly is not a reasonable conception of the good. An example would be a patriarchal conception of the good that directs men to treat women as their property. Unreasonable conceptions of the good in this sense are not permissible conceptions under LSN. The second sense of unreasonableness is epistemic. As one writer explains, "an understanding of value is fully reasonable just in case its adherents are stably disposed to affirm it as they acquire new information and subject it to critical reflection" (Cohen 1993: 281–2). An example of an unreasonable conception of the good in this sense would be one oriented around the achievement of a pointless goal, such as counting blades of grass in an open field. Such a conception of the good, I assume, would be appealing to a person only if he had false views about the significance of this activity, views that could not withstand critical reflection as he acquired more relevant information about the activity.

The grass counter, while epistemically unreasonable, could be fully reasonable in the moral sense. He could exhibit an exemplary commitment to justice. For this reason, the proponent of LSN must hold that the state should *not* aim to disadvantage his conception of the good. This opens up space for the political perfectionist to reject LSN, but accept the importance of responding to the fact of reasonable pluralism. The grass counter does not have a reasonable conception of the good, even if it is true that he is disposed to treat others reasonably. So, when the state favors reasonable conceptions of the good over unreasonable ones, it does not fail to take into account the fact of reasonable pluralism. Suppose now that the proponent of state neutrality proposes a new statement of the principle, one that is designed to take this point on board. Call it the restricted neutrality principle (Wall 2010).

> RNP: If two or more ideals of a good human life are eligible for those who live in a particular political society, and if these ideals have adherents in that political society, and if these ideals cannot be ranked by reason as better or worse than one another, then the state, to the extent that it aims to promote the good in this political society, should be neutral between these ideals in its support of them.

I shall not discuss whether RNP is a sound principle of political morality. My point in mentioning it here is that it is a principle that both responds to the fact of reasonable pluralism and is available to the political perfectionist. RNP does not rule out state action that promotes reasonable over unreasonable conceptions of the good. A proponent of this principle is free to accept the perfectionist claim that "certain conceptions of the good are worthless and demeaning, and that political action may and should be taken to eradicate or at least curtail them" (Raz 1986: 133).

RNP presents its own puzzles, however. It assumes that reason can fail to rank some conceptions of the good over others, but it is not immediately clear what this means. Consider the statement quoted above—"an understanding of value is fully reasonable

just in case its adherents are stably disposed to affirm it as they acquire new information and subject it to critical reflection." This statement itself requires explanation. What accounts for the possibility that rival understandings of value can be reflectively stable in this way?

One possibility is that there exists a plurality of ways of leading a fully good human life. The truth of value pluralism underwrites reasonable pluralism. This response is one the political perfectionist can accept, at least insofar as he believes that value pluralism is true. But the proponent of state neutrality typically has a different idea in mind. Reasonable disagreement is not to be explained by the truth of value pluralism. The expectation of reasonable disagreement is the more fundamental idea (Larmore 1994). But how then should the expectation of reasonable disagreement itself be understood? Suppose A and B reasonably disagree about some matter that bears on the human good. Each is stably disposed to affirm his view on the matter as he acquires new information and subjects it to critical scrutiny. It is possible that neither person has the better case. The arguments and evidence for the matter in dispute might not favor either view. Call this a case of *faultless disagreement*. Now suppose that A and B come to recognize that their disagreement is faultless. What should they do? It would seem that both of them should suspend judgment on the matter. But perhaps not. Perhaps A and B are rationally permitted "to take the sectarian route of affirming [his or her] own view, that is, believing it as a matter of faith" (Cohen 1993: 282).

Taking this idea on board, one could construe RNP as applying to rival ideals of the human good that are the subject of faultless disagreement. This construal would explain the sense in which reason fails to rank the ideals as better or worse than one another. This is an interesting proposal. The possibility of a genuine disagreement implies that one or more of the disputants are in error (Wright 1992), but the available evidence might not establish who is mistaken. However, it has been widely thought that state neutrality should extend to ideals of the good that are unreasonably held. The atheist must think the theist is in error; but, if he is a typical atheist, then he also will think that the balance of argument and evidence favors his side of the case. Still, he may insist that the state should be neutral between atheistic and religious views of the good. Thus, he will be inclined to reject a principle like RNP that limits state neutrality to ideals that have an equal claim to rational acceptance.

A principle of state neutrality in between LSN and RPN might capture what the neutralist is after. Consider a principle that appeals to the conscientious employment of our rational capacities. On this principle, the state should be neutral in its support of ideals of the good that are both (i) consistent with justice and (ii) affirmed (by at least some of its citizens) in good faith after they have engaged in careful reflection and well-informed deliberation on the merits of the ideals. Such a principle would not extend state neutrality to ideals for which nothing could be said, such as the grass-counting ideal, but it would not severely restrict state neutrality to ideals for which reason could not rank as better or worse.

The trick here is to explain why we should honor the conscientious, as opposed to the sound or reliable, employment of rational capacities. Defenders of state neutrality often argue that in justifying state policies we should prescind from controversial ideals of the good. Doing so, they argue, is necessary if we are to treat citizens with respect as rational agents. On this view, we treat a person disrespectfully if we subject him to state power that cannot be justified to him on his current conception of the good. He will view the exercise of state power as hostile to his interests; and it will be fully reasonable

for him to do so given his background commitments. A principle of respect for persons thus lies behind the demand that we respect people's conscientious convictions about the human good (Larmore 1990).

The political perfectionist can respond to this argument by pointing out that rational agents are not simply stuck with their current views about the good life. They can revise their views in response to evidence and argument. To respect them as rational agents is just to respect this capacity. When the state undertakes sound perfectionist policies, it might be true that its action is not justifiable to all citizens given their current convictions about the good life. It does not follow, however, that its action is not justifiable to them *tout court*. For if it is rationally possible for citizens to revise their convictions in a way that makes it possible for them to rationally accept the policies in question, then it can be said that the policies do not disrespect them as rational agents. Much here, of course, turns on the notion of rational possibility. We might know of a person that he will never abandon his convictions. Should we then take them as a fixed point for purposes of justifying political arrangements to him? Doing so respects him in one sense. It takes his perspective on matters seriously. But it disrespects him in another sense, for it treats him as an agent who cannot abandon his perspective in response to evidence and argument. These two contrasting notions of respect for rational agency go some distance toward explaining the divide between perfectionists and their critics over the issue of state neutrality.

3. Autonomy as a Perfection

In rejecting state neutrality, perfectionists seek to enlist the state in promoting the human good in their society. This aim does not sit well with a robust commitment to human freedom and individual autonomy, or so it may seem. After all, many people are not particularly good at managing their lives. Their prospects for leading good lives may increase as others take over important decisions in their lives. If autonomy is important, then it appears that there are significant limits to state perfectionism.

The appearances could be misleading. There might be distinctively perfectionist reasons to protect and promote the autonomy of persons. To be sure, autonomy is an elusive concept. We need to identify the sense of autonomy that is being considered before we can assess the extent to which perfectionism supports it. Following Feinberg, it will be useful to distinguish three senses of autonomy—autonomy as a condition, an ideal, and a right (Feinberg 1986). The latter two notions are prior to the first. Autonomy as a condition refers to the capacities and opportunities that a person must have if he is to realize the ideal of autonomy or exercise the right it designates. We understand the condition of autonomous agency by understanding what it would mean to exercise that agency.

Understood as an ideal, autonomy refers to a life of partial self-creation (Raz 1986). It is the ideal of a person charting his own course through life, fashioning his character by self-consciously choosing projects and taking on commitments from a wide range of eligible alternatives, and making something out of his life according to his own understanding of what is valuable and worth doing (Wall 1998). So characterized, this ideal might have perfectionist value. As Mill said of individuality, it may be an essential element of a well lived human life. That is a strong claim; and I shall not attempt a defense of it here. A weaker claim holds that autonomy has intrinsic value, even if it is not an essential element of a well lived human life. If either the strong or the weak claim is

granted, then there are perfectionist reasons for the state (and others) to let people make their own decisions about important aspects of their lives.

These considerations might not assuage the worry under consideration. Even if autonomy were an intrinsically valuable ideal, and even if its realization were an essential element of a good life, it would not follow that it trumps all other considerations. In promoting the human good, the perfectionist state might need to balance autonomy against other perfectionist values. The resulting balance, or so it may be worried, will fail to give autonomy its due. This worry will intensify if one affirms an additional perfectionist thesis about the value of autonomy. This thesis holds that autonomy has conditional value only. It is valuable when exercised in the pursuit of the good, but not the bad (Raz 1986). It follows from this thesis that when the state takes action to prevent people from engaging in worthless pursuits it need not thwart the kind of autonomous agency that has value and is worth protecting.

Some will suspect that the problem here is that autonomy is not well described as an element of a well lived human life. It refers instead to the prerogative a person has to make important decisions about his life, whether he is disposed to do so well or poorly. This brings us to the sense of autonomy understood as a right rather than as an ideal. Feinberg explains:

> The life that a person threatens by his own rashness is after all his life; it *belongs* to him and to no one else. For that reason alone, he must be the one to decide—for better or worse—what is to be done with it in that private realm where the interests of others are not directly involved.
>
> (Feinberg 1986: 59)

If autonomy is a right, then a person's autonomous engagement with worthless pursuits must not be interfered with. Doing so would infringe his right to self-rule.

The perfectionist state, it may now be said, fails to respect the autonomy of persons by failing to take account of this rights-based dimension of the value. This charge invites immediate qualification, however. For no one seriously maintains that the right to make decisions about one's own life extends to decisions that cause direct harm to others. As the remarks above make plain, the right to do as one pleases ends where the interests of others are directly involved. Thus to understand autonomy as a right, and consequently to understand the objection that the perfectionist state fails to respect it, we must consider the complex relationship between autonomy so understood and the harm principle. (Here I assume that one can harm a person without setting back his autonomy. Later I consider the possibility that the right to autonomy is limited only by others' equal right to it.)

4. Moral Environmentalism

Let us assume then that autonomy is a right of persons, and not merely an element of a well lived life; and let us assume further that it speaks against coercive and manipulative state interference, even when this interference correctly targets options of no value (providing these options do not involve direct harm to others). On these assumptions, should state perfectionism be rejected? Not obviously; for the perfectionist can argue that a due regard for autonomy merely restricts the means for pursuing perfectionist political objectives. It does not rule them out as such. Non-coercive and non-

manipulative state perfectionist measures remain untouched (Sher 1997). In reply, critics of perfectionism often insist that coercion lies behind all state action. "[E]ven if governmental activity is not overtly coercive, still in the last resort the government's ability to undertake any activity at all rests on its coercive power" (Waldron 1989: 1138–9).

Rather than retreat to (purportedly) non-coercive political measures, the perfectionist may stand firm and offer a more direct response to the objection we are now considering. The response traces out and extends the limits to autonomy understood as a right. Respect for autonomy does not cover autonomous actions that cause direct harm to others. This much is accepted on all sides. Controversy ensues once one starts to define the notion of "harm" for the purposes of applying the harm principle. Here there is a strong tendency for writers to either stretch the notion of harm to cover cases which warrant interference, but are not plausibly harms as commonly understood, or to reformulate the harm principle in terms of wrongful actions. The tendency is understandable, since there are compelling reasons to think that the limits to autonomy extend beyond limits that rule out actions that harm others. For example, as Mill contended, we might have an obligation to "perform certain acts of individual beneficence, such as saving a fellow creature's life." Failing to save the life of another need not harm anyone unless we say, as Mill himself did, that the failure to meet our positive obligations to others counts as harming them (Mill 1859).

Over and above positive obligations to specific persons, it is also possible that we have positive obligations to society as a whole. To mention another example from Mill, each of us may have an obligation "to bear his fair share in the common defense, or in any other joint work necessary to the interest of the society of which he enjoys the protection" (Mill 1859: 11). If accepted, these claims support an important conclusion. To understand what it means to respect the autonomy of others, we must identify the domain in which they are properly left free to act as they please. This domain is determined, in part, by the moral obligations—some negative, some positive—to which each of us is subject. It does not follow that every moral obligation should be enforced by the state. But however these obligations are specified, respect for the autonomy of people does not extend to those actions that contravene their obligations to others in their society.

To this conclusion, the perfectionist can add the claim that each citizen has a general moral duty to do his fair share in creating and sustaining a valuable moral environment in the society in which he lives (George 1995). When the state enables or facilitates efforts to discharge this duty, its action does not infringe the autonomy of anyone, for its actions are actions grounded in citizens' moral obligations to one another. That duty is a duty to do one's part in creating and sustaining the kind of public culture which best enables people to lead valuable lives. When the state effectively supports valuable pursuits and discourages base ones, it facilitates the efforts of citizens to discharge this general moral duty. When the state criminalizes the production and sale of hard drugs, for example, it employs coercive measures in an effort to help its citizens meet their duty to sustain the type of environment that will best enable its members to lead good lives. These measures are accurately characterized as paternalistic (indirectly paternalistic, to be precise), but whether they are ruled out by the harm principle must be argued for, not simply assumed. For the harm principle, as we have just seen, gets its content from a moral theory that specifies the duties of citizens to one another. The critic of paternalism can object that citizens have no general moral duty to promote a valuable environment, but here the argument will be one over the extent of their moral duties rather than an argument over the importance of respecting the autonomy of others.

You may suspect that this argument has moved too quickly. Grant that the harm principle must be informed by a moral theory and grant that state enforcement of the obligations specified by this theory need not infringe autonomy. But the critic of paternalism can object that the moral theory that informs the harm principle must itself give autonomy pride of place. He may insist that the guiding thought should be that autonomy can be restricted only for the sake of securing the equal right to autonomy of others. On this thought, harms to others that are not also infringements of their right to autonomy are not harms to others for the purposes of applying the harm principle, even if they constitute direct violations of moral duties to them.

This is a consistent position. It provides a principled basis for rejecting state paternalism of the sort I have been discussing and it does so in a way that is compatible with autonomy-based state perfectionism. But is it a position the perfectionist should endorse? To think about this matter, it will be helpful to imagine two possible moral environments for a given society. Environment 1 is autonomy-friendly in the sense that its members are free to pursue repugnant as well as valuable options, so long as the pursuit of the repugnant options does not detract from the autonomy of others. Environment 2, by contrast, is a better environment than Environment 1 in the sense that its members are more likely to lead valuable autonomous lives, but is less autonomy-friendly in that it coercively restricts the pursuit of some worthless options permitted in Environment 1. We now ask, which environment is better from the standpoint of perfectionist political morality?

The answer to this question will set the target for perfectionist state action for the society under consideration. The answer, however, is not straightforward, since the question is ambiguous. The question might concern what environment is best in terms of facilitating the successful pursuit of objectively valuable human lives. If this is the question, then it is very plausible to hold that Environment 2, at least in some scenarios, will be the better target. This is very plausible, since, as I have already noted, autonomy is but one element of a well lived human life and, on a perfectionist view of its value, it is only conditionally valuable. But the question posed might be a different one. It might be about which environment does the best job of respecting the autonomous agency of its members. If this is the question, then Environment 1 is the better option, since the state action required to secure Environment 2 will, by stipulation, be less autonomy-friendly.

This second version of the question is the one to focus on here, for we are presently considering the objection that the perfectionist state fails to show due respect for the right to autonomy of its members. Nevertheless, how the perfectionist responds to the first version of the question bears on how he should respond to the second. For the perfectionist can say that just as valuable autonomous agency, not autonomous agency as such, is what contributes to a well lived life, so too it is valuable autonomous agency, not autonomous agency as such, that commands respect. Respect for the right to autonomy, on this view, is derived from respect for persons understood as rational agents. The state fails to show respect for its members when its action interferes with their rational efforts to lead valuable lives. But when the state closes off repugnant options, then it need not interfere with these efforts and it may, in fact, help some of its members come to recognize the merits of better options. By so doing, it would express respect for their rational agency by refusing to view them as irredeemably tied to the pursuit of valueless options.

The discussion has brought us right back to the principle of respect for persons broached earlier. I have been suggesting that the perfectionist reading of this

principle—the reading that brings with it the rejection of LSN—may also foreclose the strong appeal to respect for autonomy needed to block paternalistic measures. This is only a suggestion. Doubtless more needs to be said on the issue. But these brief remarks have been intended only to make it clearer why an embrace of perfectionist politics poses a challenge to anyone who seeks a principled basis for rejecting state paternalism.

5. Politics and Imperfection

Like much contemporary political theory, this chapter has focused on general claims and principles. It has characterized a perfectionist approach to politics in terms of what it is morally fitting for the state to do; namely that it is legitimate for the state to promote, deliberately and actively, valuable human lives over less valuable human lives. And this chapter has drawn out some of the implications of this commitment by highlighting the perfectionist rejection of certain principles and arguments for limiting the scope of legitimate state power. But it is one thing to discuss what is appropriate political action in the abstract, quite another to offer concrete recommendations about what should be done in this or that circumstance. A serious worry about perfectionism in politics is that it demands too much from state actors. In the real world, it is unrealistic to expect political leaders to possess either the virtue or the competence needed to promote the good effectively.

It is a good question just how serious this worry must be for it to speak against perfectionism. For despite its name, perfectionism does not have to be an especially demanding view. It can be embraced by those with a healthy distrust of politics. And it certainly does not require its proponents to deny that much of modern politics centers on self-interested bargaining and compromise over competing interests. But perhaps this much can be said. In directing the state to promote the good, the perfectionist must think that this injunction is not generally self-defeating. He must think that there are realistic circumstances in which modern states can and will effectively promote the human good if they aim to do so. A perfectionism that always counseled its own rejection would not be one worthy of the name.

Related Topics

Pluralism, Natural Law and Rights Theory, Autonomy

References

Cohen, J. (1993) "Moral Pluralism and Political Consensus," in D. Copp, J. Hampton and J. Roemer (eds.), *The Idea of Democracy*, Cambridge: Cambridge University Press.
Darwall, S. (2002) *Welfare and Rational Care*, Princeton, NJ: Princeton University Press.
Feinberg, J. (1986) *Harm to Self*, New York: Oxford University Press.
George, R. (1995) *Making Men Moral*, Oxford: Oxford University Press.
Hurka, T. (1993) *Perfectionism*, Oxford: Oxford University Press.
Larmore, C. (1990) "Political Liberalism," *Political Theory* 18: 339–60.
—— (1994) "Pluralism and Reasonable Disagreement," *Social Philosophy and Policy* 11: 61–79.
Mill, J. S. (1859) *On Liberty*, Indianapolis: Hackett University Press.
Rawls, J. (1993) *Political Liberalism*, New York: Columbia University Press.
Raz, J. (1986) *The Morality of Freedom*, Oxford: Clarendon Press.
—— (2004) "The Role of Well-being," *Philosophical Perspectives* 18: 269–94.
Scanlon, T. M. (1998) *What We Owe to Each Other*, Cambridge, MA: Harvard University Press.

Sher, G. (1997) *Beyond Neutrality: Perfectionism in Politics*, Cambridge: Cambridge University Press.
Waldron, J. (1989) "Autonomy and Perfectionism in Raz's *Morality of Freedom*," *University of Southern California Law Review* 62: 1098–1152.
Wall, S. (1998) *Liberalism, Perfectionism and Restraint*, Cambridge: Cambridge University Press.
—— (2010) "Neutralism for Perfectionists: The Case of Restricted State Neutrality," *Ethics* 120: 232–56.
Wolf, S. (1982) "Moral Saints," *Journal of Philosophy* 79: 419–39.
Wright, C. (1992) *Truth and Objectivity*, Cambridge, MA: Harvard University Press.

Further Reading

A number of philosophers in the western tradition including Plato, Aristotle, Aquinas, Marx, (arguably) Mill, and T. H. Green were perfectionists about politics. An early contemporary defense of perfectionism is V. Haksar, *Equality, Liberty and Perfectionism* (Oxford, 1979). Raz's *The Morality of Freedom* is a major contemporary statement and defense of a liberal form of perfectionism. Hurka's *Perfectionism* and Sher's *Beyond Neutrality* both offer defenses of perfectionism that have become influential. Joseph Chan's "Legitimacy, Unanimity, and Perfectionism" (*Philosophy and Public Affairs*, 2000) presents a defense of a moderate form of perfectionism that is compatible with political contractualism. A collection of contemporary writings both in favor and critical of perfectionist politics is S. Wall and G. Klosko's *Perfectionism and Neutrality: Essays in Liberal Theory* (Rowman & Littlefield, 2003). This collection contains selections that defend state neutrality and reject perfectionism from, among others, R. Dworkin, J. Rawls, B. Ackerman, and G. Gaus. D. Rasmussen and D. Den Uyl's *Norms of Liberty* (Penn State, 2005) defends perfectionist ethics while rejecting perfectionist politics and J. Quong's *Liberalism without Perfection* (Oxford, 2011) presents a critique of contemporary arguments for perfectionism.

31
PLURALISM

George Crowder

The word "pluralism" has many meanings in political theory, but this chapter will focus on only one of these—namely, "value pluralism" or "moral pluralism," an idea whose explicit formulation is owed principally to Isaiah Berlin. (Berlin's most substantial formulations of value pluralism are found in Berlin 1990, 2000, and 2002, but see also the references in "Further reading" below.)

The core idea is that human values are irreducibly plural, frequently in conflict, and—crucially—sometimes "incommensurable" with one another. When values are incommensurable they are radically distinct, sharing no common measure and subject to no ranking or trade-off relation that applies in all cases. For example, if liberty and equality are incommensurable, they are intrinsically different considerations. Conflicts among liberty and equality then cannot be resolved by any single rule that measures them against one another in terms of a common denominator (e.g. utility), or that arranges them in an absolute hierarchy such that liberty always overrides equality or vice versa, or that trades them off according, for example, to market demand. Rather, conflicts among such incommensurables seem to leave us with hard choices—hard both in the sense that any decision will result in some degree of absolute loss in terms of the value foregone, and in the sense that there can be no one convenient rule that will tell us what to do every time.

Exactly how hard these choices are, and consequently how we ought to respond to them, is disputed. Varying interpretations of incommensurability give rise to rival accounts of the nature of choice under pluralism, and of the ethical and political directions that such choices should take.

Roughly speaking there are three main schools of opinion. First, some pluralists interpret incommensurability very strongly to mean that choices among plural values must be non-rational. This may be called the "agonistic" school, and its supporters have usually tended to emphasize the role of democratic politics in resolving pluralist conflicts. Second, other pluralists, allowing a more relaxed interpretation of incommensurability, stress the possibility of reasoned choice among incommensurable values within specific contexts. These "contextualist" thinkers have linked pluralism with a variety of political positions ranging from conservatism through communitarianism and multiculturalism to liberalism. Finally, an alternative "conceptual" approach proposes that reflection on the concept of value pluralism itself yields principles that suggest political implications, usually liberal.

1. Berlin and Value Pluralism

Intimations of the idea of value pluralism can be traced back through Western thought to the ancient Greeks, whose sense of tragedy in works such as Sophocles' *Antigone*,

and perhaps also whose polytheism, hints strongly at the notion of the ultimate incommensurability of human values. However, it is only in the twentieth century that the concept has been formulated explicitly. Max Weber is sometimes cited as an influential source, especially through his lecture, "Politics as a Vocation" (Weber 1948). But the thinker that contemporary pluralists cite as their inspiration more often than any other is Isaiah Berlin.

Berlin's pluralism is itself the product of many possible influences—he refers to Machiavelli, Vico and Herder, for example—but is best seen as emerging from his inquiry into the intellectual origins of twentieth-century totalitarianism. For Berlin, those origins lie most deeply in what he calls moral monism—that is, the deep-seated belief that every moral question must have a single correct answer and that the answer can be reduced to a single comprehensible formula that applies in all cases. From this arises the expectation that we can identify a uniquely correct way of life for all human beings, which in turn can be realized through a uniquely correct form of politics. These are the central pillars, in Berlin's view, of the political utopianism, pursued by both left-wing and right-wing extremists, that marred the history of the twentieth century. The ideal of the perfected society has been used to justify the sacrifice of all lesser values, together with the people whose values they are, and those sacrifices have been further intensified by the inevitable disappointments of dictators and social engineers frustrated by the stubborn reality of the "crooked timber of humanity" (Berlin 1990: 19).

The reality of human values, Berlin argues, is not unity but deep plurality and conflict: "The world that we encounter in ordinary experience is one in which we are faced with choices between ends equally ultimate, and claims equally absolute, the realisation of some of which must inevitably involve the sacrifice of others" (Berlin 2002: 214). The realm of human morality is one of multiple, clashing values requiring difficult choices. This is a matter not merely of appearances but of the deep structure of our ethical experience.

Beyond this broad picture Berlin provides few details, and he does not develop the idea of pluralism or its implications systematically. But gathering together what he does say, and adding to this the work of other writers, we can think of the basic concept of value pluralism as having four main elements. (Accounts of value pluralism other than Berlin's can be found in Chang 1997; Crowder 2002, 2004; D'Agostino 2003, 2004; Galston 2002, 2005; Gray 1995a, 1995b, 2000; Kekes 1993; Nussbaum 1992, 2001; Raz 1986; Stocker 1990.)

First, pluralists give some account of the "human values" that are said to be plural, although there is little consensus among pluralists over what this account should be. Opinion differs, for example, as to whether the plural values are universal or relative. While most pluralists, at any rate in the Anglo-American literature inspired by Berlin, accept that there are at least some universal and objective values, there is further divergence over the content of these values. Berlin himself sees the plural values as including universals that are "quasi-empirical," or ends "that dominate life and thought over a very large portion (even if not the whole) of recorded history" (Berlin 2002: 45). But he says very little about exactly which values fall into this category, only casually referring to liberty, equality, justice and courage in this connection (Berlin 2002: 172; 1990: 12; Jahanbegloo 1992: 37). An alternative model is provided by those pluralists who are closer to the natural law tradition, such as Martha Nussbaum with her ten-point list of "the central human capabilities" (Nussbaum 2000: 78–80).

A second element of pluralism is, of course, plurality. Values, at whatever level and with whatever content, are irreducibly plural rather than commensurable in terms of

a common measure or trade-off relation or conceivable as components of a single harmony or hierarchy. Even here there is dispute, however, with different pluralists giving varying answers to the question, "plurality of what?" Many pluralists speak of a plurality of ways of life or cultures or conceptions of the good. But the core notion of pluralism actually refers to a plurality of individual values or goods. The idea of incommensurable values has different implications from the idea of incommensurable cultures or systems of values, and pluralists have reason to reject the latter idea. I shall return to this point later.

Third, pluralists tend to emphasize the potential for conflict among the plural values. A standard pluralist theme is that to choose one value is often to forgo another, or at least to call for a compromise where one value is partially traded off in order that another be partially realized. Such conflicts might be conceptual in nature, where one combination of values necessarily excludes another—for example, a whole life based on the benefits of celibacy must rule out the rewards of sexual adventurism—or they might be more circumstantial, where for example we have to choose between different activities because of time constraints. At its most general the pluralist insight is that no single individual life, social order or conception of the good can accommodate all genuine values equally. Choices must be made, and every choice brings losses as well as gains.

The fourth element of pluralism, the incommensurability of values, is the most distinctive, and also the most controversial. Monists could accept everything said so far. They could agree that there are at least some universal values, and that these are plural and often conflicting. But they could respond that such conflicts can (at least in principle) be resolved by a single rule or formula that either commensurates or ranks the conflicting values in all cases. That is what pluralists most distinctively deny, and their denial appeals to the idea that values are incommensurable. Each value stakes a distinct claim such that there is no basis for commensuration or for a single absolute ranking or single scheme for trade-offs.

Here again there is disagreement, since value incommensurability has been variously interpreted. On what may be called the weak view, values are incommensurable simply when they cannot be measured against one another according to a common denominator such as utility. But this is a relatively weak interpretation because although it rules out quantified (or cardinal) ranking, it is consistent with qualitative (or ordinal) ranking. Someone could plausibly say that *War and Peace* is a better novel than *The Da Vinci Code* even though its superiority cannot be measured in commensurable units.

By contrast, a strong interpretation would insist that incommensurability always rules out not only cardinal but also ordinal ranking as well. On this view liberty and equality are such distinct considerations that we have no decisive reason to rank either ahead of the other in any circumstances. Of course each is itself a reason for action, but those reasons speak with such different voices that there can be no overriding or decisive rationale for preferring one to the other. When we have to choose among such values in cases of conflict, our choice is ultimately non-rational—"without criteria, grounds, or principles" (Gray 1995a: 61).

Between the weak and strong interpretations there is also what may be labeled a moderate view of incommensurability. On this view, incommensurability is not just about measurement, it bears on ordinal ranking too. However, the moderate position sees the strong view as too strong. Ranking for decisive reason may be impossible in the abstract, or in accordance with a single formula that applies in every case, but that is compatible with allowing the possibility of decisively reasoned ranking in a particular context.

Liberty does not always override equality, but there may be good reason to attach more weight to liberty (or vice versa) in a particular case.

These different understandings of value incommensurability have led to different accounts of the ethical and political implications of value pluralism. Setting aside the weak interpretation, I focus next on the political readings derived from the strong and moderate views.

2. The Agonistic Approach

The strong or "agonistic" interpretation of pluralism, according to which choice among incommensurables is ultimately non-rational, has been given a variety of political readings, but perhaps the dominant tendency has been towards an emphasis on democracy.

One version of this position is that of Chantal Mouffe (2000, 2005), who associates agonistic pluralism with a "radical" understanding of democracy. Following Carl Schmitt (1996), Mouffe argues that the deep plurality of value can be negotiated only by non-rational decisions or arbitrary acts of will. Different people's decisions will conflict, and conflict passionately, giving rise to a natural "antagonism." But while Schmitt believes that antagonism can be controlled only by coercive dictatorship—hence his support for Nazism—Mouffe thinks that antagonism should be moderated into "agonism," in which people oppose one another in a respectful way. Enemies become opponents. This is the realm of "the political," differing from liberal politics by replacing the regulating role of reason with that of "passionate commitment" and "hegemony," or dominant power (Mouffe 2000: 97, 99).

"Agonistic democracy" is also commended by William Connolly (2002, 2005), although in a somewhat different form and backed by a different justification. For Connolly, politics cannot be regulated by any overarching conception of human well-being, since this is contested by plural and incommensurable accounts of "identity." Politics is thus a realm of identity rather than reason. Further, since each identity makes its own unique claims to be heard, that suggests a link between agonism and democracy. But democracy is an ambiguous arena in which the expression of an identity risks a denigrating response from those who are different, leading to mutual aggression. What we need in order to control this tendency is an "agonistic ethic of care" in which people declare and defend their identities without "sanctifying" them or demonizing others (Connolly 2002: xv, 14). Such an ethic can find a basis in pluralism, since the multiplicity of value is such that no single way of life (or identity) can contain it—"life overflows the boundaries of identity" (Connolly 2002: 166). Thus I should respect others because their ways of life express genuine goods that I have forgone in my own life, but that are no less genuine for that.

A persistent difficulty with these agonistic positions is that it is hard to see how their political conclusions are adequately supported by their agonistic premises. If choice among incommensurables is really just a matter of passionate commitment, then why should we privilege Mouffe's agonism over Schmitt's antagonism? In the case of Connolly, the argument for agonistic care works only by looking beyond agonism to the kind of background notion of the human good—the value of the "life [that] overflows the boundaries of identity"—that Connolly began by rejecting. A purely agonistic chooser has, by definition, no reason to care about any identity more than another, so no reason to reject identities that are themselves exclusive and aggressive towards others. In short, on agonistic assumptions no political position is more justified than any other, and any is permissible.

Another critical question to ask of the agonists is whether we have to accept their key premise that choices among incommensurables must be non-rational or decisionist. This brings us to the next school of pluralist thought, which starts with a more moderate understanding of incommensurability.

3. The Contextual Approach

By contrast with the agonists, the "contextualist" school denies that incommensurability of values always precludes the possibility of practical reasoning. It is true that incommensurable values cannot be ranked for decisive reason in the abstract, or according to a single rule that applies in all cases, but that is consistent with ranking values for good reason within a particular context. This is not to say that every conflict of plural values has a rational solution; it is to say only that, given an appropriate understanding of what counts as rational in value judgment, reasoned decisions in this field are possible.

How exactly practical reasoning can be contextual is difficult to explain concisely. Part of the problem is that no simple set of rules can capture the variety and complexity of actual cases. That does not mean, however, that reasoning in particular cases of deep value conflict is impossible. The classic account is Aristotle's *phronesis*, but recent writers have argued something like the following. A particular context may involve, or generate, a "covering value" or set of background criteria according to which the goods in question (that is, the particular instances of these that concern us here) may be ranked or traded off for good reason (Chang 1997: 5). So, for example, in the context of a society-wide commitment to reform in the direction of greater social justice, there might be good reason to diminish the liberty of taxpayers in order to expand equality of opportunity in education, health care and so forth. The decision process would not be quantifiable, it would not be algorithmic, but it would still be rational in the sense that there would in that situation be a decisive reason to favor, on balance, one possible outcome over another. Variations on this kind of contextual account of choice among incommensurable values have been offered by several recent writers (e.g., Kekes 1993; Nussbaum 1992, 2001; Richardson 1997).

Another question is how we should conceive of the relevant "context." Different possibilities are canvassed in the literature, but the dominant idea along these lines is that choice among conflicting incommensurables can be determined or guided by local culture or tradition. Berlin, for example, refers to the guidance of "the course of conduct which least obstructs the general pattern of life in which we believe" (Berlin 2002: 47).

The political implications of this culturalist approach have in turn been variously interpreted. Berlin assumes that within a given political society it is the dominant national culture that should guide us, as long as this is tolerant of other national identities, including those of cultural minorities in its midst (Gardels 1991: 21).

This view would be criticized by many pluralists. More conservative pluralists would take a more assimilationist position, insisting on a thoroughgoing allegiance to a single culture "all the way down." Alasdair MacIntyre, for example, links pluralism with fidelity to a single comprehensive tradition as the only alternative to an irresolvable plurality of ethical commitments (MacIntyre 1985: ch. 17). Similarly, John Kekes argues that when it comes to choices among incommensurables "the grounds on which such judgments rest are the conceptions of a good life regarded as acceptable in the surrounding tradition" (Kekes 1993: 77).

For other contextualists, however, the proper response to pluralism moves away from cultural unity towards greater emphasis on cultural difference—hence the possibility of multiculturalist interpretations of pluralism. For example, Charles Taylor (1994) follows Berlin in stressing the value of positive cultural recognition to individual well-being, the moral authority of existing ways of life (especially in the face of conflicts among plural values), and the centrality of cultural identification at the national level. But he differs from Berlin in demanding not just toleration but also political recognition for minority nationalities, his own special concern being the self-determination claims of Quebec.

Bhikhu Parekh (2006) takes the multiculturalist line of argument further still. If cultures provide people with valuable guidance in the face of plural and conflicting incommensurable values, then that is true not only of national cultures, whether dominant or minority, but also of other forms and levels of culture, such as those of migrant or religious groups. On the same logic as that found in Berlin and Taylor, these groups, too, are deserving not only of ethical but also of political recognition, although not necessarily to the extent of according them the rights of self-determination that apply to nations. Taylor, in his more recent work, has also stressed the value of religious conceptions of the good, or "transcendence," in coping with value plurality (Taylor 2007).

However, the differences among these varying accounts raise a critical question concerning all culturalist approaches to pluralism. Supposing that we should look to culture or tradition to resolve conflicts among incommensurables, *which* of these should we look to, and why? Any particular culture or tradition can be seen as privileging one possible combination of values. But a pluralist will ask, why that combination rather than the alternatives? The pluralist perspective invites us to appreciate that there is a wide range of values or goods in the ethical experience of humanity. From that perspective, a reliance on existing culture or tradition looks arbitrarily narrow and selective.

One response could be to argue that the appeal to culture or tradition is simply a realistic recognition of our social and historical experience. No single society can embrace all possible values equally, so any society must be selective in this respect. Further, the particular selection to which a society is committed tends, in fact, to be the product of historical experience rather than the result of a "vindicatory" argument that has been justified rationally against all alternatives (Williams 2006: 189). But while this might be a compelling explanation of how we have, in fact, resolved value conflict at a social level, it is scarcely the end of the story for the social reformer who wants to question the status quo.

Here it might be thought that the complaint about selectivity is more of a problem for those conservative views that emphasize the claims of a single dominant culture than for the more accommodating multiculturalist kind of response. But even the latter raises problems from a pluralist point of view. Parekh, for example, defends the traditions of both indigenous and immigrant groups, but what happens when these conflict with each other or with the traditions of other groups? His answer is "intercultural dialogue," but dialogue needs rules and where are these to come from? In the end, the notion of resolving conflicts among incommensurable values by appeal to tradition seems merely to reproduce the problem at the level of multiple and conflicting traditions.

A different strategy might be to abandon the appeal to tradition, or at least to supplement it, by bringing in other kinds of contextual thinking. Joseph Raz, for example, accepts the incommensurability of values but makes a case for special weight to be accorded to the value of personal autonomy "as an ideal particularly suited to the condi-

tions of the industrial age and its aftermath with their fast changing technologies and free movement of labor" (Raz 1986: 369). Raz's argument suggests a pluralist case for liberalism, but the trouble with this is that non-liberal responses to modern industrialism seem to be just as successful at producing flourishing lives (depending on what one means by "flourishing"), as seen in the case of the East Asian "tiger" economies (Gray 1995b: 83). The wider the context to which one appeals, it seems, the less well placed one is to justify any particular political position within the pluralist spectrum of possibilities.

An alternative proposal, advanced by John Gray (2000), is to conclude that the search for a principled political position compatible with pluralism is doomed to failure, and to fall back on a kind of loose pragmatism. Thus Gray argues that in the end the appropriate pluralist politics is one of *modus vivendi*, according to which we should simply seek whatever settlement is necessary or effective in bringing peace among rival cultures, traditions or groups. Yet even this response is problematic since Gray's *modus vivendi* seems to uphold peace as a super-value.

A final set of difficulties with contextualist pluralism, especially of the kind that emphasizes cultural context, recalls the distinction made earlier between the plurality of values and the plurality of cultures. The culturalist response to pluralism, in effect, makes culture morally authoritative. Behind this is the assumption that not only values but also cultures are morally incommensurable—since if that is so then each can be judged only on its own terms. But pluralists have reason to deny this. As noted earlier, most pluralists accept that there are patterns of universal value that cross cultural boundaries within "the human horizon" (Berlin 1990: 11). This immediately means that whole cultures cannot be entirely incommensurable since they overlap on the universal values. To the extent that the values of different cultures are commensurable in this way, cultures can be critically compared and their performance ranked, at least within particular dimensions. For example, the freedom promoted by culture A might be superior to the freedom offered by culture B—that is, more complete, answering to more dimensions of liberty. So, on most pluralist views, existing cultural traditions cannot have the last word. It is the plurality and incommensurability of goods, not cultures, that is fundamental for the value pluralist.

4. The Conceptual Approach

The limitations of the agonistic and contextualist approaches will be especially worrying to pluralists sympathetic to liberalism. Basic liberal values such as human rights would be poorly defended if their advocates had to rely only on the agonists' ungrounded commitments or on the contextualists' varying accounts of the relevant context. So, liberal pluralists in particular have been interested in replacing or supplementing these with a third approach, one that seeks guidance in the concept of value pluralism itself.

Such a "conceptual" approach seems to be offered by Berlin when he observes that hard choices among conflicting incommensurables are unavoidable. "It is because this is their situation," he writes, "that men place such immense value upon the freedom to choose" (Berlin 2002: 214). The wider political implication appears to be that pluralism is best expressed by liberalism, in which the individual's freedom to choose is an especially salient value.

If this is Berlin's argument, it suffers from an obvious defect. The mere fact that we cannot avoid choices among plural values does not show that we should value either

the choices or the freedom to make them. Berlin himself notes that such choices may be painful and unwelcome. But can Berlin's conceptual argument be strengthened or reformulated and revived? It may be possible to argue as follows.

The truth of pluralism requires that we accept a basic principle of "respect for value plurality." If values are plural in the way that pluralists say they are, then the human good consists of many distinct dimensions, each of which makes its own unique ethical claim on us—one could imagine Nussbaum's human capabilities, for example, as supplying the content of these values. To acknowledge value pluralism is to respect each and every one of those claims as possessing a status fundamentally equal to that of other such claims. Further, to respect each claim is to hold, in principle, that it ought to be acted upon. Consequently, to respect value plurality is to accept that, prima facie, all such values have an equal claim to promotion in action.

Of course, no single person or society can promote every good in the human spectrum to the same extent and simultaneously. To emphasize individual liberty is to downplay allegiance to tradition, and vice versa. In particular cases choices have to be made, or at any rate a stress on one value will lead to less attention being paid to another. We must often choose against some important goods in order to advance others.

Nevertheless, we can still respect a good even when choosing against it. We can do this by insisting that we should, as far as possible, choose against a genuine good only when we have good reason to do so. To choose against a good simply by ignoring it, or by some arbitrary decision procedure, is to fail to take seriously the claim which that good has upon us, and so to fail to take seriously the full range of human values. The pluralist outlook implies a commitment to practical reasoning.

How exactly practical reasoning can proceed under pluralism is a notoriously hard question to answer, as admitted earlier. But in any case, liberal pluralists do not have to solve every conflict of values in order to make a case for liberalism. After all, liberal principles do not purport to answer all moral questions, but only to provide a political framework within which such questions can be debated and resolved freely and peacefully. Along these lines, liberal pluralists can argue that respect for value plurality gives us reasons to accept a distinctively liberal political framework. Within that framework incommensurable values will still collide and hard choices will still arise, but these conflicts and choices will be contained and managed by the framework.

Why might respect for plurality suggest a specifically liberal political framework? First, respect for value plurality suggests the desirability of a society which, roughly speaking, gives its members real opportunities to pursue more rather than fewer goods, or at least a greater share of the total spectrum of plural values. A society that deliberately or negligently narrows the range of goods open to its people—for reasons other than maintaining some necessary degree of coherence among the ends pursued—is unlikely to be one that adequately appreciates the force of the goods it denies. Liberal societies, with their emphasis on individual liberty, have a strong claim to satisfy this "diversity of goods" requirement. Of course, even liberal societies have costs in terms of goods excluded or discouraged. But they arguably impose fewer costs of this kind than non-liberal societies that limit the range of possibilities to those endorsed by authoritarian leaders or received traditions.

Second, to respect value plurality is to acknowledge that plural values can be legitimately combined in many different ways, suggesting a society that accommodates a range of different cultures or ways of life. Such a society will recognize a spectrum of "reasonable disagreement," to adapt a term from John Rawls (1993). Again, liberalism is a strong can-

didate for meeting this requirement, since it possesses an unrivalled capacity to accommodate multiple ways of life. Again, there are limits to the extent to which liberalism can be accommodating in this way, but liberals can argue that those limits should be accepted by pluralists. For pluralists as for liberals, not all ways of life are equally hospitable to the diversity of goods already mentioned, or to the individual autonomy I shall come to in a moment—or indeed to the core notion of respect for value plurality.

Third, respect for value plurality gives us a reason to place extra weight, especially for political purposes, on individual autonomy—that is, the capacity of individuals to control their lives in accordance with commitments arrived at through critical reflection on their own cultural, ethical and personal norms. Monist rules and received traditions cannot be decisive because they themselves embody value rankings that can be questioned. Those who are able to take a critically reflective view of the various values, principles and background conceptions of the goods in play around them are best placed to respond to the hard choices of pluralism in a way that does justice to the complexities involved. Assuming that a capacity for personal autonomy does not arise automatically but requires education, it follows that a society that genuinely respects value plurality will be one that promotes individual autonomy as a matter of public policy, principally through its system of education but perhaps also through a more deliberative approach to democracy. Again, liberals are well placed to argue that such a society will be a liberal society.

It should also be noted that this stress on personal autonomy is controversial even among liberal pluralists. For example, William Galston (2002, 2005) joins with many other contemporary liberals in regarding personal autonomy as too sectarian a foundation for liberalism, which must tolerate non-liberal groups that do not regard individual autonomy as a cherished value. But Galston's view can in turn be questioned by asking whether pluralist diversity should not be promoted within, as well as among, groups, and whether the ability of individuals to exit from their groups if they wish does not depend on personal autonomy after all (Crowder 2007).

It might seem an obvious objection that the case just outlined violates the cardinal pluralist insight that where values are incommensurable they cannot be subject to an absolute ranking. Do liberals not accord their favored values—toleration, rights and personal autonomy—just such an absolute priority? But neutralist and Rawlsian "political" liberals could reply that they prioritize liberal goods only within a political context rather than comprehensively. Perfectionist or comprehensive liberals, on the other hand, might point out that (as noted earlier), while single values might be incommensurable with one another, whole frameworks of value—including cultural and other conceptions of the good—are partially commensurable, since they overlap on generic universal values such as Nussbaum's central human capabilities. If so, the liberal good and its rivals can be critically compared and ranked at those points of overlap. In such an exercise the liberal view is likely to do well for the reasons already given. Once more, this is not to say that liberalism is cost-free in pluralist terms. Rather, the claim is that any society requires some such framework, and that the liberal proposal is likely to fare better than the alternatives.

5. Conclusion

There is a widespread although largely unstated assumption that the literature of value pluralism, although interesting, is somewhat tangential to the central concerns of contemporary political theory. In part this is because pluralism is so often taken to mean

that choice among incommensurables must be non-rational, which seems to undermine any prospect of reasoned agreement on the terms of desirable political reform (Dworkin 2001). Another influential argument labels Berlinian pluralism as too controversial a starting-point for modern political theory compared with the liberal neutrality championed by Rawls and others. While value pluralism represents a particular conception of the human good that many people would reject, Rawlsian "reasonable pluralism" simply captures the fact of reasonable disagreement about the good under modern conditions (Larmore 1996: ch. 7).

These attempts to dismiss or circumvent value pluralism are themselves highly contestable. As discussed above, pluralism does not necessarily rule out practical reasoning or the defense of political principles; only the agonists believe it does, and even they usually try to connect pluralism with political reform. As for the Rawlsian critique, pluralism is not a substantial formula for the good life of the kind bracketed by Rawls's neutrality. Further, there is reason to believe that Rawlsian reasonable disagreement itself depends, at least in part, on an acceptance of value pluralism (Crowder 2002: 165–71; Galston 2002: 46–7).

In fact, one has only to look at the list of thinkers mentioned in this chapter to see that value pluralism is now an abiding interest and concern for mainstream political theory. There is still much that is unclear and undeveloped in pluralist thinking, but it would be rash to conclude that it has no more to offer.

Related Topics

Liberalism, Conservatism, Contractualism and Political Liberalism, Perfectionism, Multiculturalism, Autonomy

References

Berlin, I. (1990) "The Pursuit of the Ideal," in *The Crooked Timber of Humanity: Chapters in the History of Ideas*, ed. H. Hardy, London: John Murray.

—— (2000) "My Intellectual Path," in *The Power of Ideas*, ed. H. Hardy, London: Chatto & Windus.

—— (2002) *Liberty: Incorporating* Four Essays on Liberty, ed. H. Hardy, Oxford: Oxford University Press.

Chang, R., ed. (1997) *Incommensurability, Incomparability, and Practical Reasoning*, Cambridge: Cambridge University Press.

Connolly, W. (2002) *Identity/Difference: Democratic Negotiations of Political Paradox*, expanded edition, Minneapolis: University of Minnesota Press.

—— (2005) *Pluralism*, Durham, NC: Duke University Press.

Crowder, G. (2002) *Liberalism and Value Pluralism*, London: Continuum.

—— (2004) *Isaiah Berlin: Liberty and Pluralism*, Cambridge: Polity.

—— (2007) "Two Concepts of Liberal Pluralism," *Political Theory* 35: 121–46.

D'Agostino, F. (2003) *Incommensurability and Commensuration: The Common Denominator*, London: Ashgate.

—— (2004) "Pluralism and Liberalism," in G. Gaus and C. Kukathas, eds., *Handbook of Political Theory*, London: Sage.

Dworkin, R. (2001) "Do Liberal Values Conflict?," in R. Dworkin, M. Lilla and R. Silvers, eds., *The Legacy of Isaiah Berlin*, New York: New York Review Books.

Galston, W. (2002) *Liberal Pluralism: The Implications of Value Pluralism for Political Theory and Practice*, Cambridge: Cambridge University Press.

—— (2005) *The Practice of Liberal Pluralism*, Cambridge: Cambridge University Press.

Gardels, N. (1991) "Two Concepts of Nationalism: an Interview with Isaiah Berlin," *New York Review of Books* 21 November: 19–23.

Gray, J. (1995a) *Berlin*, London: HarperCollins.

—— (1995b) *Enlightenment's Wake: Politics and Culture at the Close of the Modern Age*, London and New York: Routledge.
—— (2000) *Two Faces of Liberalism*, Cambridge: Polity.
Jahanbegloo, R. (1992) *Conversations with Isaiah Berlin*, London: Peter Halban.
Kekes, J. (1993) *The Morality of Pluralism*, Princeton, NJ: Princeton University Press.
Larmore, C. (1996) *The Morals of Modernity*, Cambridge: Cambridge University Press.
MacIntyre, A. (1985) *After Virtue: A Study in Moral Theory*, 2nd edn, London: Duckworth.
Mouffe, C. (2000) *On the Democratic Paradox*, London: Verso.
—— (2005) *On the Political*, London: Routledge.
Nussbaum, M. (1992) *Love's Knowledge: Essays on Philosophy and Literature*, Oxford: Oxford University Press.
—— (2000) *Women and Moral Development: The Capabilities Approach*, Cambridge: Cambridge University Press.
—— (2001) *The Fragility of Goodness: Luck and Ethics in Greek Tragedy and Philosophy*, 2nd edn., Cambridge: Cambridge University Press.
Parekh, B. (2006) *Rethinking Multiculturalism: Cultural Diversity and Political Theory*, 2nd edn., London: Palgrave.
Rawls, J. (1993) *Political Liberalism*, New York: Columbia University Press.
Raz, J. (1986) *The Morality of Freedom*, Oxford: Clarendon Press.
Richardson, H. S. (1997) *Practical Reasoning about Final Ends*, Cambridge: Cambridge University Press.
Schmitt, C. (1996) *The Concept of the Political* [1932], trans. G. Schwab, Chicago, IL: Chicago University Press.
Stocker, M. (1990) *Plural and Conflicting Values*, Oxford: Clarendon.
Taylor, C. (1994) "The Politics of Recognition," in A. Gutmann, ed., *Multiculturalism: Examining the Politics of Recognition*, Princeton, NJ: Princeton University Press.
—— (2007) *A Secular Age*, Cambridge, MA: Belknap.
Weber, M. (1948) "Politics as a Vocation," in H. Gerth and C. Wright Mills, eds., *From Max Weber*, London: Routledge & Kegan Paul.
Williams, B. (2006) *Philosophy as a Humanistic Discipline*, ed. A. W. Moore, Princeton, NJ: Princeton University Press.

Further Reading

Isaiah Berlin's comments on value pluralism are scattered throughout his work: see, e.g., his essay on Machiavelli in *Against the Current: Essays in the History of Ideas*, ed. H. Hardy (London: Hogarth, 1979), and his treatment of Vico and Herder in *Three Critics of the Enlightenment*, ed. H. Hardy (Princeton, NJ: Princeton University Press, 2000). Other significant sources for his views in this area include the interviews contained in "Isaiah Berlin in Conversation with Steven Lukes," *Salmagundi* 120: 52–134; and I. Berlin and B. Polanowska-Sygulska, *Unfinished Dialogue* (Amherst, NY: Prometheus, 2006). Further commentary on Berlin's pluralism can be found in A. Ryan, ed., *The Idea of Freedom: Essays in Honour of Isaiah Berlin* (Oxford: Oxford University Press, 1979); R. Dworkin, M. Lilla, and R. B. Silvers, eds., *The Legacy of Isaiah Berlin* (New York: New York Review Books, 2001); G. Gaus, *Contemporary Theories of Liberalism* (London: Sage, 2003), ch. 2; and G. Crowder and H. Hardy, eds., *The One and the Many: Reading Isaiah Berlin* (Amherst, NY: Prometheus, 2007). Other thinkers whose work has addressed or intersected with value pluralism include S. Hampshire, *Morality and Conflict* (Cambridge, MA: Harvard University Press, 1983), *Innocence and Experience* (Cambridge, MA: Harvard University Press, 1989), *Justice is Conflict* (Princeton, NJ: Princeton University Press, 2000); M. Oakeshott, *Rationalism in Politics and Other Essays* (London: Methuen, 1962), *On Human Conduct* (Oxford: Clarendon, 1975); M. Walzer, *Spheres of Justice: A Defence of Pluralism and Equality* (Oxford: Blackwell, 1983). For edited collections dealing with pluralism, see D. Archard, ed., *Philosophy and Pluralism* (Cambridge: Cambridge University Press, 1996); M. Baghramian and A. Ingram, eds., *Pluralism: The Philosophy and Politics of Diversity* (London: Routledge, 2000); and the "Symposium on Pluralism and Ethical Theory", *Ethics* 102 (1992): 707–834.

32
VIRTUE ETHICS AND POLITICAL PHILOSOPHY

Daniel C. Russell

Virtue ethics is a form of normative ethics that appraises action and guides choice by focusing on the virtues of character. Since the middle of the twentieth century, virtue ethics has seen tremendous growth as a modern approach to normative and applied ethics. This is no small accomplishment, since virtue ethicists have had much work to do just to claim space for character-based ethics alongside more familiar rival theories that focus more on rules, duties, and actions (see e.g. Frankena 1980 and Louden 1984 for representative challenges). The impetus for such a movement can be traced to the 1950s, when Elizabeth Anscombe (1958) and Philippa Foot (1958) each called for a renewed focus in moral philosophy on human nature, well-being, wisdom, and character, and in the following decades there grew a hope for an "ethic of virtue" as a distinctive form of normative ethics. Since then, virtue ethics has become widely recognized as a distinct alternative in ethical theory, demonstrating a unique and useful approach to action appraisal and guidance (on which see esp. Hursthouse 1995 and 1999). Today, some of the most exciting work in virtue ethics focuses on extending that approach to political philosophy as well.

It is no small thing to shift from ethics to politics, though, in part because the things being appraised—institutions instead of actions—are so different, but also because the state is by its nature coercive. Before we can ask what states ought to do, we must ask whether and how such a coercive monopoly could be legitimate in the first place, and always keep in mind that whatever the state does is done by force. These issues have no analogs in other types of ethical thought; we cannot move breezily from "what it would be virtuous for someone to do" to "what it would be just for the state to do." Furthermore, the shift to politics is potentially perilous for virtue ethics particularly, since the ancient perspective on ethics—especially Aristotle's—that virtue ethics commonly revives also pre-dates modern concerns over the very justification of political order. It is perhaps no accident, for instance, that Alasdair MacIntyre's (1981) self-consciously pre-modern view of the state, with its emphasis on the virtues as inextricable from their cultural and political situation, has not managed to become the political face of virtue ethics.

So the time seems right to ask what form or forms we might expect virtue ethics to take in political philosophy. I shall make a beginning in this chapter, by examining several major virtue-based approaches to political philosophy developed over the past two decades. As I do so, I draw attention to some broad structural differences between these

approaches and to how they fare in confronting fundamental issues peculiar to political philosophy. Hopefully this will be instructive for thinking about future directions for virtue-ethical political thought.

Before beginning, I should point out the sorts of approaches that fall outside our scope. Not every account of what the virtues are is a case of virtue ethics (see e.g. Driver 1996 and 2001), and so too not every political discussion of the virtues is virtue-ethical. For example, Hobbes defined the moral virtues as dispositions to obey the natural laws (see *Leviathan* XV, 79–80). In doing so, he illuminated the virtues in terms of prior moral and political principles, whereas virtue ethics hopes for illumination in the opposite direction. Likewise, in recent years many political philosophers have identified certain so-called "citizen virtues" that support political orders, such as liberalism (see Galston 1991) or republicanism (see Pettit 1999). Others have identified "institutional virtues" that consist in the internalization of the norms of a social or political institution by its members (Audi 2007). By contrast, virtue ethics understands "the right" in terms of "the good," namely acting from good character. The account of good character is prior to action appraisal in virtue ethics; if it were not, the central claim that the right thing to do is what it would be virtuous to do would be merely trivial (Watson 1990; see also McDowell 1997: 141; Hursthouse 1999: 30–1; Russell 2009: 65–7). Virtue ethics should, therefore, be distinguished from perfectionism as well, which understands the virtues as dispositions to do what is right, namely, to promote or value goods such as perfected human nature or achievement; such dispositions can also be thought of as forms of human perfection themselves (see Hurka 2001; see also Watson 1990; Wall 2007).

1. Direct Approaches: From Virtue to Politics

The most direct approach would be to assess laws and institutions in the very same terms in which virtue ethics assesses the actions and choices of individuals. We can think of this as virtue ethics "writ large," so to speak, what we might call "virtue politics." For example, one sort of virtue politics is familiar from Plato's *Republic*, where he defines justice in the state as structurally isomorphic to virtue in the soul: in each case, a certain proper hierarchy of the constituent "parts." In our own time, Michael Slote has made justice and virtue isomorphic in other ways. In the 1990s, Slote identified the justice of a state with the virtuousness of its citizens; more recently, he has identified the justice of particular laws and institutions with the virtuousness of citizens in creating them. Let's consider these in turn.

In "Virtue Ethics and Democratic Values," Slote identifies civic justice with the virtuousness of citizens and their relations to each other. Slote characterizes the fundamental citizen-virtue as "self-sufficiency," which is the opposite of dependency and parasitism and is a source of superabundant generosity. Slote then maintains that "a society is just to the extent its people exemplify all the forms of self-sufficiency" (1993: 16). (Slote 1998 retains this basic theoretical structure but with the Ur-virtue of "caring.")

It is important to note that this approach does not locate the injustice of repression, say, in repressive social institutions themselves. Rather, Slote argues that repressive rule by dominant classes is unjust because such classes are "socially parasitic" and thus not self-sufficient in his sense (Slote 1993: 17–18, 24–5). The injustice, then, lies in what the perpetrators themselves are like. Likewise, Slote argues that societies with repressed classes are unjust also because the members of those classes are then "intellectual and volitional parasites" (1993: 19).

However, surely the injustice of a repressive society lies in what is done by the repressors to the repressed—the acts by which the repressed are robbed of self-sufficiency—or in the institutions that make the repression possible, rather than in anyone's character per se. In fact, if the very presence of vices in citizens is what constitutes injustice, then the vices of the repressed would seem to constitute injustice in exactly the same way that the vices of the repressors do. Does that mean we could ameliorate the injustice by shipping out the repressed, and their vices with them? Slote's approach is therefore importantly different from, say, that recently developed by Lisa Tessman (2005), who identifies much of the injustice of repression with its depriving the repressed of opportunities to develop the virtues that are important for human flourishing.

Slote's approach also rests on some very important but unstated and undefended assumptions about the nature of political obligation. For instance, he argues that because it is less generous to give to others while reserving the right to stop doing so, a just society will enforce redistribution of property and indeed will tend strongly towards economic equality (Slote 1993: 27–30; see also 1998: 188). By making the virtuousness of citizens the determinant of state justice, Slote takes for granted that it is appropriate for the state to use coercion to ensure that the actions of its citizens count as virtuous (in this case, generous). A virtuous person would act generously; would a virtuous person also coerce others to act generously? Would that make him even more generous?

Slote does not make his presuppositions about coercion, obligation, and liberty explicit, and this may be no accident given Slote's more general turn towards an "ethic of care" (see Slote 1998, 2007, 2010), which focuses more on sentiment and intimacy than on considerations of obligations and liberties. But surely even those who rightly appreciate the importance of caring and intimacy in their lives must confront the moral questions raised by the very fact that states wield coercive force. Indeed, those who would build their lives around such values must surely have an interest in defending their liberty to do so (see Hampton 1993).

Slote takes a different approach in his 2001 book *Morals from Motives*, evaluating particular institutions in terms of the virtues and vices of those who create and administer the institutions. In doing so Slote makes the justice of institutions exactly parallel to the rightness of actions: in each case, he argues, the proper moral assessment depends entirely on the moral quality of the actual motives expressed by the relevant agents, individuals in the case of rightness and public officials and voters in the case of justice (see 2001: 99–100). Of course, he recognizes that despite the best intentions social institutions can have actual results that are deeply regrettable, as for instance when innocent persons are mistakenly and even systematically punished; and in general, even the most "public-spirited legislators" can pass laws on the basis of "faulty or totally misleading information … despite their best efforts" (2001: 106–7; see also 1998: 191–2). However, because the justice of laws and institutions lies entirely in the motives behind them, Slote maintains, they can still be just despite any such regrettable consequences.

To be sure, it is highly controversial even among virtue ethicists whether the rightness of action is entirely dependent on the virtuousness of the agent's motives (see Russell 2009: ch. 2 for discussion). Surely it will be at least as controversial whether justice is entirely a matter of motives, since the things being assessed—actions and institutions—seem importantly different in precisely this respect. Even if all that matters in the assessment of action is the quality of the motive, one might still maintain that the assessment of laws as just or unjust should take into account their actual effects on people's lives.

A variation on Slote's approach might be possible. Many virtue ethicists hold that there are some things that it may be virtuous to do (e.g. removing life support from a suffering patient) even though they are things no virtuous person would ever want to do; in such cases we might distinguish between an act virtuously done and a right act that merits praise and satisfaction (see Hursthouse 1999: chs. 2–3; 1995: 62–6). Likewise, we might distinguish a law virtuously passed from a law that a virtuous person would regard as praiseworthy. But even though this would preserve isomorphism between evaluations of actions and of laws, it is not clear why that isomorphism should be worth preserving in the first place—the fact that that way of thinking about justice would be "virtue-ethical" is no argument at all. Furthermore, this approach sidesteps fundamental political questions: again, the case of justice is different because state action involves the use of force; so, what justifies such action? Perhaps on the approach in question we could ask what sorts of purposes a virtuous person would think justify the state's use of coercive force. But that is to say that the approach offers a way of framing the question; it is not clear whether that way of framing it is useful or even appropriate.

So there are at least two problems with Slote's direct "virtue politics" approach, in both of its guises. One, it shows at most that isomorphism is possible between virtue ethics and virtue-based political philosophy, but not whether it is well advised. And two, it sidesteps fundamental questions about the very justifiability of political order. The task is to show not that virtue ethics might say something about politics, but that what virtue ethics might say is something that we need to hear.

2. Indirect Approaches: The State and Well-Being

Virtue ethics is not just about evaluating one-off actions or guiding one-off choices. Its real focus is on the kind of character and practical intelligence that are important for leading a good human life. In fact, the oldest tradition in virtue ethics, eudaimonism, understands the virtues as traits indispensable for well-being or eudaimonia (see Annas 1993; Hursthouse 1999: chs. 8–11). Although eudaimonia is sometimes characterized as a perfectionist good, such as "being a good specimen of one's kind" (see Sumner 1996; Haybron 2008: ch. 8), this is a mistake. Eudaimonia is, in fact, a kind of well-being or happiness, understood as a life that is rewarding for the person living it both as an individual and as a human (Russell 2012). Virtue ethicists might hope to illuminate political thought by approaching it from this same perspective: what is important for a distinctively human form of well-being.

Rosalind Hursthouse outlines the basic structure of a eudaimonist approach to political philosophy in her 1990–91 paper, "After Hume's Justice." For Hursthouse, this approach comes down to a pair of ideas: human flourishing or eudaimonia is prior to the account of social justice, and social justice is prior to the account of rights, such as property rights (1990–91: 235–6; see also Nussbaum 1990a: 205, 208). Correspondingly, the first stage of what we might call "political eudaimonism" works by developing a conception of human nature and flourishing, and understands a just society as one that realizes the social conditions necessary for human flourishing. The second stage moves from this notion of social justice to principles of justice defining civic institutions and rights. These are the claims, freedoms, and protections that persons must have in order to live well together. For instance, since private property gives one continuing title to a thing beyond mere possession during actual use, the political eudaimonist must determine whether human flourishing involves activities that

require such durable control over parts of the extra-personal world (see also Nussbaum 1990a: 232).

However, within this broad structure there is room for substantial disagreement over the state's role in realizing the social conditions for eudaimonia, and thus over the extent of the state's involvement in citizens' lives. There are three main versions of this approach for us to consider: (1) the state is a *shaper* of citizens' lives so as to promote their well-being or flourishing; (2) the state is a *guarantor of opportunities* for constructing a flourishing life; and (3) the state is a *protective framework* within which flourishing can be pursued.

The view of the state as closely shaping the lives of its citizens for the sake of their virtue and well-being is familiar from the political writings of Aristotle. According to Aristotle, the state ought to enable its citizens to be their best and to flourish (*Politics* VII.2, 1324a23–25, 1325a7–10; cp. *Rhetoric* I.5), so that any inquiry into the nature of good political organization must take its start from a conception of eudaimonia (*Politics* III.9, 1280a31–1280b12, 1280b29–1281a10; VII.1, 1323a14–21; see also Jacobs 2004: 152, 160). The state then promotes eudaimonia so understood by carefully engineering the conditions of citizens' lives through centralized state regulation, rather than relying on private efforts (*Nicomachean Ethics* X.9, 1179b20–1180a19; *Politics* VII.15, 1334b4–5 with Newman's addendum). Because virtue is indispensable for eudaimonia and best acquired young, the state regulates all facets of social life that bear on the development of the virtues in the young, from procreation (VII.16) to what children learn in school (*Politics* VIII.1–2; *Nicomachean Ethics* X.9, 1180a14–24) and eventually to their choice of occupation (*Politics* VIII.3–7; see Miller 1995: ch. 6 for discussion).

A broadly Aristotelian conception of the state has its modern advocates as well. For instance, Colin Farrelly and Lawrence Solum advocate a virtue-based theory of jurisprudence on which "the purpose of law is to enable humans to acquire, maintain, and exercise the human excellences or virtues" (Farrelly and Solum 2008: 2; see also Huigens 1995). Likewise, following Aristotle and Aquinas, Robert P. George has argued that an end of sound politics is to encourage citizens to lead valuable and morally upright lives, and that towards that end the state may justly use coercive force to protect citizens "from the corrupting influences of vices" (George 2008: 25). Moreover, George and others are part of a growing, so-called "natural law" movement in political theory that defends the authority of the state to enforce what human nature and virtue would obligate one to do (see Feser 2010 for a recent discussion).

However, even if virtuous persons would wish their fellow citizens to live virtuously or at least not viciously, the more basic question is again whether it would be virtuous for persons to force a conception of living well upon their fellow citizens. Moreover, this approach reveals a grave danger in reviving Aristotle's politics: in reaching back to the fourth century BC from our own century, we risk bypassing the seventeenth century, in which the very possibility of justifying a state monopoly in coercive force became a fundamental issue for modern political philosophy. That bell has been rung, for us if not for Aristotle, and it cannot be unrung.

Other neo-Aristotelian political theorists, however, have taken a very different and far more liberal approach to political eudaimonism. Martha Nussbaum, in developing her version of the so-called "capabilities approach," has argued that the function of the state is to prepare citizens to flourish, but unlike Aristotle she understands this function as the much weaker one of preparing citizens to adopt and pursue their own freely chosen conception of the good life (2006: 182). As Nussbaum puts the point, the role of

the state is to bring every citizen to a point at which he or she will be genuinely capable of choosing a life with human dignity, whatever life he or she actually chooses (1990a: 214; see also 1992: 221; 2000: 86–96; 2006: 71). "The government aims at capabilities," Nussbaum says, "and leaves the rest to citizens" (1990a: 214).

Nussbaum argues that although the state must have a somewhat detailed conception of what people need in order to flourish, such a conception still serves only to identify goods that would be generally recognized, across times and cultures, as important for good human lives (see esp. 1995; see also 1992: 224–5). Furthermore, the importance of pluralism is actually a central theme in this approach, by seeking to bring people to a point at which they are capable to choose a good life according to their own conception of it (see 2006: 78–81, 296–8). This is because human nature consists in the pursuit of a life of activity in accordance with practical reasoning (1988, 1992). So our needs and capabilities are not all on an equal footing; rather, some are "architectonic"—namely, our planning and structuring our lives by practical reasoning, and our doing so with each other—in the sense that they give all of our other capabilities a distinctly human shape (1990a: 226; 1988: 179). In that case, the possibilities for flourishing human lives will show endless diversity. What makes them all good human lives is that they are lives of activity shaped by practical wisdom, and when we think of it that way, this variety of ways of living is actually expressive of our shared human nature, rather than just evidence against it (cf. Hursthouse 1999: 219–20).

For this reason, although the conception of good human functioning at work in the capabilities approach is "thick," dealing with functioning across life as a whole (beyond, say, economic well-being), it is also "vague" in the sense of being an "'outline sketch' of the good life" that admits "of many concrete specifications" (Nussbaum 1990a: 217, 234–6). And although an even more detailed conception of flourishing is required as the state addresses the particular needs of particular groups of people, it is nonetheless possible to decentralize such state action, and entrust it to parties "on the ground" that are close enough to understand these more fine-grained needs (e.g. 1990a: 215–17, 236–8).

A driving theme in Nussbaum's approach has been the importance of improving people's lives by giving them opportunities to develop human capabilities that they may have been brought up to deny themselves. In this respect, much of Nussbaum's work has focused on the plight of women in under-developed countries (see esp. 2000; see also 1992), and in fact Nussbaum has envisioned her approach as ultimately calling for world-wide collective action (2005: 211–18; 2006: chs. 4–5; see also 2000: 101–5).

In an important break with Aristotle, Nussbaum describes the capabilities approach as a form of liberalism (2006: 6, 221) in virtue of divorcing its ethical ideals from divisive ideologies and seeking instead "overlapping consensus" among citizens (2006: 163; 2011), as well as attaching high value to individuality, freedom, and choice (2006: 217). She therefore considers her view a form of political liberalism, maintaining state neutrality between citizens' particular conceptions of the good life, as opposed to perfectionist liberalism (see 2008, 2011). Nussbaum contrasts the capabilities approach with other forms of liberalism as well. For instance, she distinguishes her approach from John Rawls' contractarianism insofar as the justification of citizens' entitlements on the capabilities approach is not procedural (see Rawls 1971: §§14, 20–25). Rather, the approach starts with the intuitive idea of human dignity as a political goal and then seeks institutions that will realize that goal as fully as possible (Nussbaum 2006: 37, 82, 281). Moreover, Nussbaum insists that the approach treats liberties as "abilities to do something" (2006:

290), and thus that liberty requires positive state action and not merely protection against interference. She therefore breaks sharply with classical liberalism, such as that of John Locke (e.g. Nussbaum 2006: 44–5), and has written in favor of socialism on occasion (e.g. 1990b: 205). Nussbaum has also said that, while human flourishing requires private ownership, what anyone may own is for the state to decide: "the very determination of who owns what property awaits a full scrutiny of need and surplus," the purpose of which is to create new "ownership rights" for those with less out of the surplus of those with more (2006: 19). The private property protected by the state is, therefore, whatever is left after state institutions have collected from each his due in supporting the flourishing of others both domestically and internationally (2006: 309–10, 316–17).

There is no denying that Nussbaum's neo-Aristotelian approach is far more liberal than Aristotle's own approach. By putting her focus on the centrality in human life of practical reasoning and living a life of one's own through one's own conception of the good (see esp. Nussbaum 2008, 2011), Nussbaum shows how important it is for the state to refrain from engineering citizens' lives, not in spite of the importance of living well but precisely because of it. Nonetheless, Nussbaum still does not raise any fundamental questions about the justification of the state. Her central political conviction can be summarized as follows: "The aim of political planning is the distribution to the city's individual people of the conditions in which a good human life can be chosen and lived" (1988: 145). But while Nussbaum tells us what she thinks the state ought to do, she does not say why there should be a state at all, or whether the purpose she assigns to the state warrants the use of force that is distinctive of state action. This silence might stem from the ambiguity of her crucial notion of a "just society": in *Frontiers of Justice*, for instance, it is often difficult to tell whether "what a just society would do" is what virtuous persons would do or what virtuous persons would coerce others to join them in doing. Unfortunately, for Nussbaum as for Aristotle, "society" and "state" (the Greek "*polis*" can mean either) tend to run together. Instead of neo-Aristotelian politics, we might have wished for neo-Ciceronian politics; since Cicero, like Locke after him, appreciated the fundamental importance of carefully limiting the state's jurisdiction (see esp. *On Duties* III.72).

Last, Nussbaum draws a lot of attention to the fact that her conception of good human functioning is broad-based and that it allows citizens considerable latitude in the lives they lead. Even so, the question of who should decide how political institutions are to be arranged is more basic than questions of how meritorious this or that arrangement might be. What Nussbaum might have to say about that question, however, remains unclear.

This brings us finally to Douglas Rasmussen and Douglas Den Uyl (2005), who have developed a virtue-based and eudaimonist political philosophy that places strict and carefully defined limits on the role of the state in the lives of citizens. On their view, the state is for the sake of the flourishing of citizens, but only by providing a peaceful framework within which citizens can pursue the good life according to their different conceptions of it. This framework consists mainly of institutions that protect citizens from interference and provide mechanisms for resolving disputes. The rights to be guaranteed, they argue, are aimed at ensuring that citizens' various pursuits of flourishing do not come into conflict (see esp. 2005: 266).

Rasmussen and Den Uyl distinguish sharply between the account of human flourishing as a kind of *ethical* theorizing, and the account of institutions and rights as a kind of *political* theorizing. Such an approach, they argue, "separates politics from ethics as far as possible without lapsing into either relativism, nihilism, or historicism" (2005: 37, see also 34–7, 63). The separation is motivated instead by a certain division of normative

labor, so to speak. Reflection upon and pursuit of the good is something for *people* to do, and the systematic study of the nature of the good is the domain of ethical philosophy (among others). The role of the *state*, by contrast, is not to determine the nature of the good or how to pursue it, but instead to provide a stable context in which such ethical reflection and pursuit can take place (cf. Nozick 1974: ch. 10). In short, then, the aim of eudaimonist political principles here is to protect a sphere of action within which good lives can be chosen and lived with minimal unwanted incursions or state interference.

Rasmussen and Den Uyl argue that in virtue of this division of normative labor, their approach offers a form of liberalism that avoids what has been an important problem for traditional liberalism (see e.g. 2005: 78). Not having distinguished ethical from political theorizing in form, they argue, liberalism has traditionally faced the dilemma of either making ethical theory robust—focused on rich human flourishing and not mere social cooperation—and allowing the state greater involvement in people's lives, or else restricting state involvement and making ethical theory correspondingly too thin. According to Rasmussen and Den Uyl, their approach avoids both horns of this dilemma: by separating ethical from political theorizing, the approach is liberal in its strong protection of individuals from state interference while also being robust in its views about what is ethically salient and important in human life (2005: ch. 3).

While they distinguish ethical from political theorizing in this way, Rasmussen and Den Uyl also argue that ethical theory should support political theory: if there is no such support "for a society's political/legal order, then its legitimacy is, to say the least, dubious" (2005: 84–5). This, they argue, is where a neo-Aristotelian ethic that focuses on flourishing and the virtues joins forces with political liberalism (see 2005: ch. 1). Such an ethic "conceives of human flourishing as the ultimate moral standard," and since "self-direction" or the pursuit of a life of one's own planning is "the central necessary constituent or ingredient of human flourishing" (2005: 87), that ethic "upholds a political/legal order that sees protection of individual liberty as its chief aim. Contrary to the conclusions of many contemporary liberals and conservatives alike, an ethics of human flourishing or self-perfection does not require a perfectionist politics" (2005: 85).

Even so, this approach shares with the others a certain view about the extent of the coercive power of the state: it should extend just as far as the theory's conception of virtue or human well-being requires. Put another way, if Rasmussen and Den Uyl call for a much less extensive state than Nussbaum (not to mention Aristotle) does, this is because they disagree over the point at which increased state action becomes inimical to flourishing. But for all that, they agree that the justification of the use of state coercion, as well as any resistance to it, must ultimately be framed in terms of what one thinks serves the well-being of those for whose sake the state acts. Such a political principle might offer citizens great latitude, but again, a more basic question is who should decide what principle of political association should frame political institutions.

3. Basic Approaches: Virtue and Political Obligation

More than once now we have encountered the question whether citizens might have moral claims against state coercion even when the state would coerce them for benevolent purposes. Do citizens in fact have such claims? Richard Kraut has recently argued that they do not: it might be good to enjoy some discretion over how one will live, but none can claim to be owed such discretion. There just is no sense to be made, Kraut

argues, of moral claims that would limit what the state may do for the sake of promoting flourishing (see Kraut 2007: 234–43, esp. 241).

However, other virtue ethicists disagree. In particular, Mark LeBar has recently developed a virtue-based, eudaimonist account of justification and obligation that places limits on how we may act towards each other that are indifferent to the purposes for which we might act towards each other. In "Virtue Ethics and Deontic Constraints" (2009), LeBar develops this approach by drawing on Stephen Darwall's account of how persons direct what he calls "second-personal reasons" to each other (Darwall 2006). By definition, to direct such a reason to someone (e.g. not to stand on your foot) is simultaneously to make a claim as one with the authority to make that claim, to acknowledge him as one capable of recognizing and honoring that authority of yours, and to honor his claim against you that you move him to act only by appealing to him as a rational agent (see LeBar 2009: 647). Accordingly, to direct reasons to another in this way is to acknowledge the other as having the very same authority to address you with second-personal reasons as well; part of addressing one another in this way, then, is to acknowledge the *reciprocity* of our authority to make claims with respect to one another.

Moreover, to adopt the "second-person standpoint" is to view oneself as a member of a community of rational agents who address each other with second-personal reasons. And it is here that LeBar's approach to obligation reveals itself as distinctly eudaimonist: the same conception of eudaimonia that gives persons a reason to develop the virtues also gives persons a reason to occupy the second-person standpoint, since doing so is indispensable for living a flourishing human life (see LeBar 2009: 652–4, 661–2; see also Darwall 2006: 278).

LeBar extends this general account to political obligation, and in particular how we make demands of one another through state institutions. In "Virtue and Politics" (2012), LeBar argues that since the creation of state institutions is an important form of interaction with each other, the same reasons we have to acquire the virtues are also reasons for us to authorize only those institutions by which we obligate one another within the second-person standpoint. Indeed, LeBar argues that adopting that standpoint is an indispensable part of virtuous character itself. To illustrate this point, LeBar asks us to imagine a society comprised of Slote, Nussbaum, and Rasmussen and Den Uyl, and notes that although such a society would be wholly comprised of reasonable, intelligent, and thoughtful persons, there would still be deep disagreement over fair terms of their association. In that case, whichever of them were to set those terms, he or she would do so unilaterally, not reciprocally; that is, he or she would assume an authority to set terms for others without recognizing any similar authority on their part. For LeBar, then, the crucial normative issue here is not the latitude one's theory allows others to enjoy, or even the nobility of one's purpose. Rather, given the importance of equality of authority, "the issue really is *who gets to decide* what that latitude is" (emphasis in original). By assuming for oneself the authority to make that decision for others, one abandons the second-person standpoint—and that, LeBar argues, no virtuous person would do. In that case, part of virtuous character as a whole is a deep commitment to restricting political obligation to what can originate in the second-person standpoint.

Notice, then, that whereas all of the other approaches we have considered assume that a certain conception of virtue and/or flourishing authorizes state coercion for the sake of virtue or flourishing so understood, LeBar's approach questions that very assumption. Whatever the merits of one's conception of virtue or flourishing, for LeBar the real question is whether pointing to that conception would suffice for addressing to others

a second-personal reason to submit to state coercion—the sort of reason that a virtuous person would give.

It remains to be seen how LeBar's virtue-based account of political obligation might be developed into a full-blown approach in political philosophy, so it is too early to say very much either way about the merits of such an approach. Nonetheless, that approach clearly begins with fundamental questions about what such goods as virtue or eudaimonia might have to do with the legitimacy of state action in the first place. I think that there is an important lesson in that.

4. Conclusion

Much constructive work has been undertaken in recent years to connect virtue ethics to political philosophy. So, what can we expect the future political face of virtue ethics to look like? The answer is still the one that Rosalind Hursthouse gave over a decade ago: "Who knows?" (1999: 6). However, today at least two things seem clear about the future of virtue-based political philosophy. One is that the developments in this area to date make it most doubtful that there will be any such thing as *the* political face of virtue ethics—and that, I think, is all to the good. The other point is about what progress here must look like, and I would say that the best way for virtue ethics to move forward is to take a big step back. Virtue ethics needs to begin not *in medias res* with questions about how to inject the virtues into political discourse. Virtue ethics must begin at the beginning, with the basic justificatory questions that are fundamental to anything we can recognize as a genuine contribution to modern political philosophy.

Note

Support for research on several topics discussed in this paper was provided by a visiting scholarship at the Social Philosophy and Policy Center in fall 2008, and final preparation of the paper was made possible by a visiting scholarship in summer 2011.

Related Topics

Aristotle's Social and Political Philosophy, Perfectionism, The Capability Approach (and Social Justice)

References

Annas, J. (1993) *The Morality of Happiness*, Oxford: Oxford University Press.

Anscombe, G. E. M. (1958) "Modern Moral Philosophy," *Philosophy* 33: 1–19.

Audi, R. (2007) *Moral Value and Human Diversity*, Oxford: Oxford University Press.

Darwall, S. (2006) *The Second-Person Standpoint*, Cambridge, MA: Harvard University Press.

Driver, J. (1996) "The Virtues and Human Nature," in R. Crisp (ed.), *How Should One Live?* Oxford: Oxford University Press.

—— (2001) *Uneasy Virtue*, Cambridge: Cambridge University Press.

Farrelly, C. and L. B. Solum (2008) "An Introduction to Aretaic Theories of Law," in C. Farrelly and L. B. Solum (eds.), *Virtue Jurisprudence*, New York: Palgrave Macmillan.

Feser, E. (2010) "Classical Natural Law Theory, Property Rights, and Taxation," *Social Philosophy and Policy* 27: 21–52.

Foot, P. (1958) "Moral Arguments," *Mind* 67: 502–13.

Frankena, W. (1980) *Thinking about Morality*, Ann Arbor: University of Michigan Press.

Galston, W. (1991) *Liberal Purposes: Goods, Virtues, and Diversity in the Liberal State*, Cambridge: Cambridge University Press.
George, R. P. (2008) "The Central Tradition—Its Value and Limits," in C. Farrelly and L. B. Solum (eds.), *Virtue Jurisprudence*, New York: Palgrave Macmillan.
Hampton, J. (1993) "Feminist Contractarianism," in L. Antony and C. Witt (eds.), *A Mind of One's Own*, Boulder, CO: Westview Press.
Haybron, D. (2008) *The Pursuit of Unhappiness*, Oxford: Oxford University Press.
Huigens, K. (1995) "Virtue and Inculpation," *Harvard Law Review* 108: 1423–80.
Hurka, T. (2001) *Virtue, Vice, and Value*, Oxford: Oxford University Press.
Hursthouse, R. (1990–91) "After Hume's Justice," *Proceedings of the Aristotelian Society* 91: 229–45.
—— (1995) "Applying Virtue Ethics," in R. Hursthouse, G. Lawrence, and W. Quinn (eds.), *Virtues and Reasons: Philippa Foot and Moral Theory*, Oxford: Oxford University Press.
—— (1999) *On Virtue Ethics*, Oxford: Oxford University Press.
Jacobs, J. (2004) *Aristotle's Virtues: Nature, Knowledge, and Human Good*, New York: Peter Lang.
Kraut, R. (2007) *What is Good and Why: The Ethics of Well-Being*, Cambridge, MA: Harvard University Press.
LeBar, M. (2009) "Virtue Ethics and Deontic Constraints," *Ethics* 119: 642–71.
—— (2012) "Virtue and Politics," in D. Russell (ed.), *The Cambridge Companion to Virtue Ethics*, Cambridge: Cambridge University Press.
Louden, R. (1984) "On Some Vices of Virtue Ethics," *American Philosophical Quarterly* 21: 227–36.
MacIntyre, A. (1981) *After Virtue*, South Bend, IN: Notre Dame University Press.
McDowell, J. (1997) "Virtue and Reason," in R. Crisp and M. Slote (eds.), *Virtue Ethics*, Oxford: Oxford University Press.
Miller, F. (1995) *Nature, Justice, and Rights in Aristotle's Politics*, Oxford: Oxford University Press.
Nozick, R. (1974) *Anarchy, State, and Utopia*, New York: Basic Books.
Nussbaum, M. (1988) "Nature, Function, and Capability," *Oxford Studies in Ancient Philosophy* supp. 1: 145–84.
—— (1990a) "Aristotelian Social Democracy," in R. B. Douglass, G. M. Mara, and H. S. Richardson (eds.), *Liberalism and the Good*, London: Routledge.
—— (1990b) *Love's Knowledge*, Oxford: Oxford University Press.
—— (1992) "Human Functioning and Social Justice: In Defense of Aristotelian Essentialism," *Political Theory* 20: 202–46.
—— (1995) "Aristotle on Human Nature and the Foundations of Ethics," in J. E. J. Altham and R. Harrison (eds.), *World, Mind, and Ethics: Essays on the Ethical Philosophy of Bernard Williams*, Cambridge: Cambridge University Press.
—— (2000) *Women and Human Development*, Cambridge: Cambridge University Press.
—— (2005) "Beyond the Social Contract: Capabilities and Global Justice," in G. Brock and H. Brighouse (eds.), *The Political Philosophy of Cosmopolitanism*, Cambridge: Cambridge University Press.
—— (2006) *Frontiers of Justice*, Cambridge, MA: Harvard University Press.
—— (2008) *Liberty of Conscience*, New York: Basic Books.
—— (2011) "Perfectionist Liberalism and Political Liberalism," *Philosophy and Public Affairs* 39: 3–45.
Pettit, P. (1999) *Republicanism: A Theory of Freedom and Government*, Oxford: Oxford University Press.
Rasmussen, D. and D. Den Uyl (2005) *Norms of Liberty*, University Park: Pennsylvania State University Press.
Rawls, J. (1971) *A Theory of Justice*, Cambridge, MA: Harvard University Press.
Russell, D. (2009) *Practical Intelligence and the Virtues*, Oxford: Oxford University Press.
—— (2012) "Happiness and the Virtues," in D. Russell (ed.), *The Cambridge Companion to Virtue Ethics*, Cambridge: Cambridge University Press.
Slote, M. (1993) "Virtue Ethics and Democratic Values," *Journal of Social Philosophy* 24: 5–37.
—— (1998) "The Justice of Caring," *Social Philosophy and Policy* 15: 171–95.
—— (2001) *Morals from Motives*, Oxford: Oxford University Press.
—— (2007) *The Ethics of Care and Empathy*, London: Routledge.
—— (2010) *Moral Sentimentalism*, Oxford: Oxford University Press.
Sumner, L. (1996) *Welfare, Happiness, and Ethics*, Oxford: Oxford University Press.
Tessman, L. (2005) *Burdened Virtues: Virtue Ethics for Liberatory Struggles*, Oxford: Oxford University Press.
Wall, S. (2007 rev.) "Perfectionism in Moral and Political Philosophy," *Stanford Encyclopedia of Philosophy*, http://plato.stanford.edu/entries/perfectionism-moral/.
Watson, G. (1990) "On the Primacy of Character," in O. Flanagan and A. O. Rorty (eds.), *Identity, Character, and Morality: Essays in Moral Psychology*, Cambridge, MA: MIT Press.

33

NATURAL LAW AND RIGHTS THEORY

David S. Oderberg

1. Introduction

The terms "natural law" and "natural rights" cover a variety of theories most of which are significantly different from each other and often incompatible. It is as well to state from the outset what this chapter will not be considering. We will not examine what is sometimes called the "voluntarist" tradition of natural law, reaching back to William of Ockham (c.1288–1348), according to which the laws of nature that ground morality emanate directly from an act or acts of will. For Ockham it was the will of God, who is perfectly free to mandate any system of moral laws He sees fit. For later voluntarists such as Hobbes (1588–1679) and Pufendorf (1632–94), the emphasis was on the human will as the originator of all moral laws beyond the minimal natural law enjoined by God for humans to live together in peaceful society.

Nor will we consider the enormously influential natural law theory of Locke (1632–1704), which has molded much of modern social and political liberalism as well as inspired the radically individualist natural law theory embodied in the work of Robert Nozick (1974) and, even more extreme in its libertarianism, of Murray Rothbard (1982). And we will not look at the so-called "new natural law theory" found in the work of John Finnis, Germain Grisez, and Joseph Boyle (see for instance Finnis 1980 and Finnis *et al.* 1987). Its proponents see it as breaking with the classical natural law tradition we will outline by abjuring any grounding in metaphysics, specifically a metaphysic of human nature. Instead, new natural law theorists seek a path to the good through the agent's reflection upon practical reason.

The remainder of this chapter, then, will set out some of the fundamental ideas and principles of classical natural law theory (with the omission of "classical" for the most part). We will begin with some brief historical remarks, but the outline of the theory will be largely ahistorical, embracing contemporary ideas. The tradition has its wellsprings in Ancient Greece and Rome. Plato proposed the famous analogy between health as the natural condition of the body and justice as the natural order of the state (*Republic* IV: 444d–e; Hamilton and Cairns 1961: 687). Aristotle spoke of a "universal law" or "law of nature" binding on all human beings, even those having no "association or covenant" with each other (*Rhetoric* 1.13, 1373b5; Ross 1924). Seneca spoke of a guiding hand behind the universe, springing from "the command of eternal law" (*On Providence* 1, my translation; Reinhardt 2007: 3). Quotations could be multiplied from Cicero, the Stoic philosophers, and others. The common idea

behind the various formulations and emphases is that not just the physical universe, but the affairs of men—governed as they are by morality—are regulated by an eternal and immutable law to which "right reason" must conform if human agents are to be held to act morally.

These germs of classical natural law received their full flowering in the medieval period, the apogee of which was the work of St Thomas Aquinas, the greatest of the medieval philosophers and theologians. Aquinas crystallized the strands of natural law thinking from previous writers into a systematic and all-but-exhaustive account of morality as a system of laws built into the fabric of the universe and of human affairs in particular. (See his magisterial *Summa Theologica*, II.I, and especially the *Treatise on Law* (Regan 2002).) Aquinas saw the natural law as part of the eternal divine law, written into the consciences of human beings and never capable, at least in its first principles, of being erased. Though the Thomistic development of natural law long held sway in the universities of Europe, this inextricable entwinement of the natural law with the law of God was seen by many scholars, at least from the seventeenth century onwards, and rightly or wrongly, as partly responsible for the bitterly divisive religious wars that split Europe. Hence the move to a more "modern" conception of natural law as divorced from religion, becoming the scholarly domain exclusively of philosophers (at least *qua* philosophers rather than theologians). But now let us move on to the conception of the human being as it appears in classical natural law.

2. Man

Natural law theory has two characteristics that set it apart from most if not all other normative moral theories. (For more on these, see Oderberg 2010; Feser 2009: ch. 5; Feser 2010.) First, it is *essentialist*. In other words, it posits the existence of an immutable human nature that, while its manifestation in different people involves variations across culture, genetics, upbringing, race, gender, and historical context, is nevertheless fixed in the basic requirements for its fulfillment. Hence one of the reasons the theory is a theory of *law* is that it applies without exception to all human beings. To say, then, that natural law is written into the hearts—or better the consciences—of men is more than a rhetorical flourish. Natural law theorists see the moral law as imprinted into human nature in a way analogous to that in which physical laws are imprinted into the fabric of the universe. The natural law against murder, for example, is no more subject to exception than the law of universal gravitation. (Putative exceptions such as killing in war, self-defense, and capital punishment, are in fact not exceptions at all: see Oderberg 2000b.)

Second, natural law theory is *teleological*. This term is misleadingly applied in contemporary parlance to any theory that merely proposes the achievement of some state of affairs as the purpose (*telos*) of morality. In this sense, utilitarianism is sometimes described as a teleological theory because it holds maximization of utility (or sometimes "the good") as the purpose of moral agency, the state of affairs moral behavior must bring about. Natural law theory is teleological in the far more substantial sense of ascribing *intrinsic purpose* to the powers and faculties of the human being, whether bodily, mental, or some combination of the two. Human beings have built into their nature a suite of ends the satisfaction of which fulfills all the various goal-directed parts (including powers, faculties, functions) of that nature. Just as, for instance, the eye has the *function* and *power* of sight or the stomach that of digestion, so the human mind has

the *function* and *power* of grasping truth and, through its diverse faculties, understanding the world around it. So, too, the human being as a whole has the *function* (or rather *a* function) of living in society with other human beings. Human nature for the natural law theorist, then, is neither a blank slate to be shaped wholly by experience, nor a fixed program determined in the past by the course of evolution and/or set by the human genotype. Human nature has a certain plasticity, but only within the rigid boundaries set by the built-in functions and purposes of the components that go to make up the complete person.

Correlative with ideas such as the functions, purposes, and ends of human nature and its parts is the concept of *fulfillment*. Again, in this context fulfillment does not mean self-determination without constraint, being whatever one wants to be, or pursuing happiness in whatever way one happens to construe happiness. Rather, fulfillment is the successful pursuit of those objectives set by nature for the human person. Hence there can be no coherent natural law concept of fulfillment without the equally important concept of *the good* as that which regulates human activity. Goodness, for the natural law theorist, is built into the structure of reality. This "natural goodness" (Foot 2001, though she is not a natural law theorist as such) is *ex hypothesi* not superimposed upon human affairs by any agreement between people, nor by the power of the state or any other political authority. It pre-exists any social compact or form of government, these latter being themselves regulated by it. Natural law theorists typically define the good as a complex of a finite number of basic ends that fulfill human nature. (For discussion, see Oderberg 2004; for a well-known statement by a new natural law theorist, see Finnis 1980.) These typically include such things as life, knowledge, work and play, friendship, and for most, religion. Since these goods are as much a constituent of reality as the goodness of, say, water for a plant or food for an animal, there is no question of any naturalistic fallacy or fact-value distinction. Natural law theory recognizes neither since the normativity essential to natural goodness is itself a fact about the world. Hence the world is literally not value-free.

Since human beings are by nature bound to pursue the good—that is, rationally bound, bound on pain of incoherence or a kind of metaphysical deviance—they must have a *right* to do so. If they did not, the natural law would give with one hand what it takes away with the other, which would make morality incoherent. (For more, see Oderberg 2000a: 53–63.) Just as human positive law confers rights on its subjects (a loose way of talking—see below), so too the natural law confers rights. The common law, for example, allows the formation of contracts for sale. But it also gives the seller a *right* to be paid on provision of the thing sold. What point or force would there be to a contract if the law gave no rights to its parties to enforce the terms? The analogy with natural law might not be perfect, but it is clear enough: since the natural law mandates (it does not just allow, it *requires*) that human beings pursue the goods that fulfill their natures, so it protects that pursuit by a kind of network of *powers* that enable the pursuit to be carried out. These powers are rights.

Note what this implies about classical as against contemporary natural rights theory. At least some rights, for the classical theorist, involve no discretion whatsoever. Every person has a right to life but no right, for example, deliberately to end it. Everyone has the right to be healthy, but no right intentionally to undermine their health. The general right to pursue the good is founded on a *duty* so to do, and so there is no correlative right to refrain from such pursuit. For the contemporary rights theorist—at least one who does not accept a foundation in some kind of natural law—the idea of rights

without discretions borders on incoherence. For the classical theorist, on the other hand, the concept of a non-discretionary right is imperative, since rights are protective powers: the choice of concept is because rights entail duties of non-interference, whereas duties typically do not—at least not without making an inference from one duty to another *via* rights.

Once more, "power" must be understood correctly. A right, for the natural law theorist, is a *moral* power, not a physical or political power. Every human being has natural rights even if neither positive law nor political authority recognizes them, and even if the human being concerned is not physically able to enforce respect for their rights or, for that matter, even to demand respect. It is this recognition of natural rights that grounds the equal dignity of human beings and bars any conception of the person as a creature of the state or of those other people with physical power over him. In this sense, then, no human law or political authority ever *gives* anyone rights or takes them away. They either recognize pre-existing rights (such as the rights to life and property) or specify detail in the exercise of rights that are only conferred in a generic way by the natural law (for example, what form a contract must take, what kind of justice must be meted out to what kind of criminal). These rights can be proclaimed in documents (such as a bill of rights or constitution) or enshrined in statute or common law, but putting them into concrete form does not bring them into existence *ex nihilo*.

Three prime characteristics of natural rights are as follows. (A brief but useful outline is in Cronin 1930: 662–5.) First, they are all *inviolable*. If a person has a right, say the right to life, then there is a correlative duty on everyone else not to interfere with the pursuit of the good protected by that right. Hence natural law rules out sacrificing the life of an innocent person no matter how much good it may achieve, or the state's passing a law banning marriage, or a judge's decision to enforce an agreement for the supply of slaves or child labor. Second, all rights have attached to them what is sometimes called the property of *coaction*, which is best thought of as a second-order right to use whatever means necessary (as qualified by the next feature), by the right-holder or another acting on their behalf, to defend their first-order right. So, for example, you have the right to use force to repel someone invading your house or assaulting your body. Not all coaction need involve violence, of course: much of it takes place through the civil legal system, via complaints to public authorities, or by such simple means as remonstration, protest, or even avoiding the potential violation.

Nevertheless, as must immediately be emphasized, coaction cannot involve any activity that violates the rights of another, which brings in the third all-important characteristic—*limitation*. All rights are limited by other rights, on pain of incoherence in the system of morality. For example, every person has a natural right to the fruit of their labor, but they are also bound by the natural law of charity, that is, of kindness to those in need. For this reason, shocking as it may be to the ears of libertarian natural law theorists and other individualists, a starving man has the right to take the surplus of someone else's food. This is not because the second person has no right to their food; of course they do. Nor does it mean that the natural law allows theft, which it could no more do than allow adultery or child abuse. Rather, the natural law of charity entails that property rights are limited by the rights of others to aid and assistance in need. (Question: can a person dying of kidney failure get her friend to frog-march a third person to the nearest hospital to have that person's kidney removed for transplant? If not, why not?)

3. Economy

The concept of limitation of rights brings us to the idea of human beings in society. The term "economy" (*oikonomía*) was used by the Greeks, especially Plato and Aristotle, to mean the management of a family or household. It also was used to mean administration generally, including public regulation, finance, and government. Today we restrict the term solely to matters of money, particularly on the national or international scale. But it used to mean something more. Aristotle famously remarked that ethical theory is a kind of political science (*politikē*; *Nicomachean Ethics* 1.2, 1094b10; Ross 1925). By this he did not mean that there is no such thing as the theory of individual virtue or of the good for a person. Rather, he meant that a complete ethical theory must recognize that man is a social animal whose nature tends to the formation of social groupings at all levels from family and friends to political institutions in the modern, narrow sense.

Natural law theory takes the social nature of the human person very seriously by steering between the extremes of individualism on the one hand and totalitarianism on the other. The human being is not an ethical atom, free to determine herself according to her own ideas and desires. Neither is she an Hegelian cog in the machine of the state, with her existence and identity wholly subsumed by it. How is this abstract idea realized in practice? By the insistence upon a multi-layered society of social groupings along a sliding scale from the smallest and most intimate—the family—to the largest and most impersonal—the state. In between, as natural law theorists have always observed, there must be all sorts of associations of people serving the ends of their members and the wider good—for example: networks of friends; local communities; regional communities; trade and professional associations; corporations; academic and cultural organizations; religious bodies; special interest groups of all shapes and sizes (though "special interest groups" has come to have a somewhat sour smell in modern politics, confined as it is to lobby groups; not that lobbying one's politicians as such is contrary to natural law!); and so on. In many ways the family is usually considered the most important precisely because it is the most intimate—the group within which future members of society are most influenced and shaped, for good or for ill. Hence it is no surprise that many natural law theorists, including those of the "new" variety, put so much emphasis on resisting all attempts to distort or undermine family life, whether it be through opposition to any taxation that penalizes marriage, support for a "living wage" for single-earner households, or hostility to any legislative proposals to undermine the unique place of marriage between a man and a woman as the socially superior way of raising children.

A multi-layered society, then, serves a twofold purpose. First, it allows human beings to flourish according to their natures: people want and need special relationships of all kinds, they are poorer without friends or family, without associations to promote their professional or other rights, and the like. Second, all of these associations function as a buffer between the individual and the quasi-omnipotence of the state. Since political history around the world is more a story of governments taking away rights and freedoms rather than granting them, there is an ever-present danger that the more a government is able to intrude directly into the lives of its citizens, the more likely it is to ignore their natural rights and thwart their attempts to live according to the natural law. The many layers of a well-functioning society make this all the more difficult. As an example—though here many would argue that the associations are *too* good at protecting themselves from state intrusion—we may cite the legal and medical professions, which are probably the last remaining authentic descendants of the medieval guild system.

Through tight, effective organization and mutual aid, they have managed remarkably well to preserve the power, autonomy, and influence of their members, even in countries, such as those throughout Europe, that possess a state-funded healthcare system.

Just as rights are limited by each other at the individual level, so the rights of associations within a multi-layered society are mutually limiting. The power of a trade union, for example, must be limited by the rights of its members, the rights of employers, and the rights of society at large. Hence its purpose must be solely the promotion and protection of its members' welfare yet not at the expense of *the common good*. A strike that seriously compromised the provision of essential services (one cannot avoid vagueness, of course) or the use of violence against recalcitrant members would be ipso facto contrary to the natural law. The general point should be clear enough, namely that each association must perform the delicate balancing act of looking after its members while not undermining the common good. (For a brief but interesting outline of some of these ideas, see Maritain 1971.)

In the social philosophy of natural law, the common good is paramount. In some ways, it is easier to say what the common good is not than what it is. It is not maximal utility or maximal goodness; it is not a measurable quantity of anything; it is not a mere instrument by means of which each member of society is able to pursue their own path of self-determination. Nor is the common good, as Hobbes would have it, mere peace and stability. Perhaps the best way to think of the common good is as a kind of *plurality* or aggregate: it just is the aggregate of individual goods of all a society's members where those goods are, at least generically, common to each and every one of them. As Mark Murphy puts it (Murphy 2006: 64): "The common good aggregatively conceived is that state of affairs in which all of the members of a political community are fully flourishing." But this suggests that the common good is a *limit* concept as it is at least practically impossible to achieve such a state of affairs. Nevertheless, like the good of an individual itself, it is practically possible to pursue that state and to achieve it to a greater or lesser degree. Indeed, it is *obligatory* to do so, and any society that is not organized around the principle of striving for the common good—and achieving it to a decent enough degree (again, vagueness notwithstanding)—would be considered morally deficient to that extent.

Let us take some examples. (For illuminating discussion of these topics, see Feser 2010.) The natural right to property is, like all such rights, inviolable and coactive, but also *limited* by other rights (not just other property rights). Limitation, so the natural law account goes, is governed by the overarching principle that rights exist as moral powers and protections individuals have to enable them to fulfill their natures. Hence the natural law against theft is a law against doing what would frustrate that fulfillment. But it is not just theft by one person from another that the natural law forbids for this reason. Government expropriation of private property, at least without just cause (and usually fair compensation in the absence of national emergency), would be regarded as theft. Monopolies on the ownership of scarce or essential resources would be inherently opposed to natural law, at least if they involved (as they inevitably do in the private sector) such things as price gouging, market manipulation, or the denial of access to some people. There could be no place in a naturally just society for, say, a corporation that bought up all the water resources and sold them back at an exorbitant price (where "exorbitant" would have to be defined within the context of the relevant society and measured against alternative and cheaper means of supply). Moreover, the natural law regards all monopolies with suspicion as tending towards the frustration of the goods of

some people to the benefit of others. Competition is seen as exemplifying an important part of human fulfillment (not just—maybe not even—the movements of an "invisible hand"): human agents are more fulfilled according as they have greater control over the use and fruits of their labor. The limitation on private monopolies, then, is merely an example of the limitation of property rights for the sake of the common good. The obverse of this is that competition in some areas might work *against* the common good. For instance, if competition in rail services compromised safety due to miscommunication and lack of integration, this would be a ground either for government monopoly or even private monopoly heavily regulated and accountable to government. It is, therefore, simplistic to say that natural law theory favors a "small is beautiful" approach to economics, especially in highly technologically developed societies, though it is fair to say that it is greatly cautious of too much ownership of anything in too few hands.

Another limitation on property rights is that they may never be exercised for immoral purposes. The simple reason, again, is that such exercise would frustrate either the rightholder's or other people's fulfillment of their natures. In this area, natural law does not recognize compromise or moderate positions—the prohibitions are exceptionless. One may not use one's property to commit acts of violence against innocent people, or to enslave another, or to commit fraud. Whether and to what extent the state, via positive law, should forbid such activities, however, is another matter. The natural law is prior to positive law; it is broader than it in scope; and it has various means of sanction apart from state action, such as communal disapprobation, peer pressure, as well as the force of custom and tradition. Actions contrary to the natural law do not *ipso facto* require legislative sanction since in given social and cultural contexts a law might be unenforceable, too intrusive, likely to cause more harm than good. In such cases, the society needs to find another way to restore moral behavior. One may not use one's property rights to break a promise, for example, or to engage in perverse sexual practices in the privacy of one's home. The law, however, does not enforce promises as such, only the specific sub-class of contracts; and generally (less so than in the past) it keeps away from intruding on private sexual activity (though what counts as private is a large question).

As for economics in the more narrow and contemporary sense, natural law does not allow, in principle, free markets (even if regulated) in prostitution, pornography, slavery, child labor, or any other activity that is intrinsically immoral (rather than circumstantially immoral such as weapons trading, where there is no natural law prohibition as such on the sale or use of weapons but where circumstances might make it morally wrong). Again, the state may *tolerate* certain evils, and so not legislate against them, but the natural law requires that it try to eradicate them and certainly not promote them. (Compare prostitution to slavery. It is plausible to argue that the historically entrenched and ineradicable nature of the former makes it tolerable albeit immoral, whereas the latter has proven in fact easier to abolish through concerted social action. Only the indomitable free will of the human agent, with her tendencies toward evil as well as good, could make it possible for entrenched violations of the natural law to exist.)

On the question of taxation, the natural law does not and could not agree with the assertion that all taxation is a form of theft. The latter implies that the state is, by its nature, an imposition on the natural liberties of individuals, whereas morality holds it to be a natural organization mandated for the protection and promotion of the common

good. As such the state, whatever its form, has the right to levy taxes for that purpose and *only* for that purpose. Correlatively, citizens have the strict moral obligation to support the state in its achievement of that purpose, of which taxation is one manifestation. This is not to say that there could not exist a society without taxation; one such could, at least in theory, consist of highly civic-minded people who supported the state through voluntary contributions both of money and labor. Given the selfish biases in human nature this is an unlikely scenario but it is not contrary to the natural law. As Cronin (1930: 650) puts it, natural law sometimes prescribes an end while leaving it open what means can be used to achieve it (as long as those means are not themselves immoral). Aquinas points out that one function of positive law is to determine a more specific way in which some natural law principle is to be implemented, for instance the way in which a criminal is to be punished (*Summa Theologica* II.I, q.95 a.2; Regan 2002: 54).

That said, there is much room for debate concerning the forms and extent of taxation, as well as the size of the state itself. Here, a key natural law idea is the so-called "principle of subsidiarity," which holds that no social function should be carried out at a higher level than the one at which it can effectively be carried out. The state exists only to serve the interests of its citizens, and these involve the pursuit and fulfillment of their natural ends. Pursuit and fulfillment cannot easily be separated: the human being is an *agent*, not a passive receptacle for interests that need to be satisfied. Hence, except in those matters where an individual cannot fulfill her natural ends by her own free agency, she should be left at liberty to do so. Yet if this applies to individuals, it must also apply to groups of individuals acting together. If a family can support itself, educate its young, and decide prudently what is best for them and itself, it should be left free to do so. The same goes for a club, a village, a city, a trade union, and so on. Higher levels step in where groupings at lower levels cannot act adequately; this is why trade unions exist, namely to operate on a scale that individual workers could not for the purpose of securing fair wages and working conditions.

Clearly there is much that the modern state can do, with the plethora of resources and technology at its disposal. But this is always more than it *needs* to do to secure the common good. Hence the natural law forbids the state from taking on functions that violate the principle of subsidiarity. Highly controversial examples are health and education. Uncontroversial ones are national defense, which is the sole prerogative of the state, and the teaching of morality and/or religion, in which most natural law theorists would hold the state to have no business being directly involved (though it may do so indirectly, through the promotion and encouragement of right living). One thing seems fairly clear, namely that the explosion of the public sector in virtually all modern technocratic states, sometimes larger even than the private sector (which actually generates the wealth used by the public sector in the first place), is not only economically unjustifiable but itself a contravention of the natural law. In the words of Cronin (1939: 612):

> [I]t is the bounden duty of the State, an obligation binding in the strictest justice, to avoid a plethora of public officials, for the support of whom taxes have to be levied; and where, on account of special circumstances, a certain increase in the number of officials becomes temporarily necessary, it is the duty of government to reduce this number, so far as the existing rights allow, as soon as the special circumstances disappear.

4. State

We have seen some of the ideas and principles involved in the natural law theory of the state as guardian of the common good. We should now survey, all too briefly, some further points concerning government and the state in its relation to other states. Natural law does not of itself prescribe a particular form of government. The rule of a constitutional monarch has traditionally been held to be the superior institution, not, as critics would say, because of an accident of history or culture, but for two reasons of principle. First, government by a single ruler (the literal meaning of "monarchy") is more controlled and orderly, just as is the running of a corporation by a chief executive or of a family by a "head of the household" (anachronistic to modern ears though it might seem). Second, the principle of heredity was thought to ensure greater stewardship and care over a country, just as the owner of a house is more likely to take care of it if he knows he can pass it on to his children. (For arguments in support of monarchy, see Aquinas, *On Kingship*; Phelan and Eschmann 1949.)

That said, the natural law maintains the possibility of various kinds of government so long as they protect and promote the common good, and this must involve steering between such extremes as tyranny or autocracy on one hand (monarchy gone wrong, as it were, say without checks and balances) and mob rule on the other (democracy gone wrong, say without protection of minorities). Within these poles, highly authoritarian forms of government will be allowed as well as representative ones. Opposed to the natural law, however, will be socialism and other forms of egalitarianism (not merely because they are never in practice egalitarian but because egalitarianism is, itself, contrary to human nature) as well as liberalism in all its varieties. What unites liberal forms of government is their insistence on marginalizing religion as well as public morality, adhering to the Millian "harm principle" (the state will allow people to do anything as long as it does not "harm" others), and generally taking an officially non-judgmental attitude to the behavior of its citizens outside the narrow confines of positive law (even though, in practice, the non-judgmentalism manifests itself as the approval or even imposition of a set of highly permissive moral attitudes). All of this is generally seen by natural law theory as corrosive of the common good and unfaithful to basic truths of human nature. (See Cahill 1932: ch. IX.)

The state acting as a corporate person on behalf of its citizens is, for natural law theory, as much beholden to objective morality as when it acts within its own confines for its citizens. Hence there is a long and rich tradition of what is sometimes called the "law of nations," or international law. Its most fruitful period begins in the early modern era, with such Catholic theologians and philosophers as Francisco de Vitoria (1480–1546), Francisco Suarez (1548–1617), and Gabriel Vasquez (1651–94). Many would say it reached its zenith with the Protestant legal theorist Hugo Grotius (1583–1645), who developed a lengthy and highly systematic theory of natural law and rights as applied to relations between states (Tuck 2005). It has always been difficult to determine just how classical Grotius really was. His famous "impious hypothesis" was that the natural law would be valid even if God did not exist. Many contemporaries and later writers, apart from being scandalized, thought this meant Grotius wanted to separate morality from religion and propose a truly secular natural law. This is not obviously correct since at least some prior Scholastic philosophers and theologians (even the great Saint Robert Bellarmine) had proposed just the same hypothesis (Crowe 1977: 223–8). The idea is not so much that atheism and natural law are

compatible, rather that the essences and purposes from which natural law derives are objective and implanted by God in His design for the world. God, on this view, works remotely through what the Scholastics called "proximate" causes in decreeing the natural law. Moreover, one need not believe in God to be able to recognize at least the basic principles of natural law. To this extent alone, then, it could be regarded as a secular morality. Nevertheless, the real metaphysical underpinning would be God, and serious reflection on the natural law should point human intelligence in that direction.

Whatever the truth of the matter, Grotius did in large part enunciate a classical theory of international law, focusing mainly (and unsurprisingly) on the law of war. The key natural law concept here is that of the *just war*, that is, a war fought both for the right reasons and in the right way. Hence there is a *jus ad bellum* (the moral principles that govern going into a war) and the *jus in bello* (those governing behavior during war). All too briefly, the first principle of the *jus ad bellum* is that war is always a last resort, after all peaceful means of dispute resolution (negotiation, diplomacy) have failed. If a state goes to war, it must have a just cause. This category includes such reasons as: self-defense against attack; a war of independence against enslavement by another nation; a pre-emptive strike against another state's preparations for attack; a war to obtain compensation for previously inflicted damage.

The *jus in bello* has as its primary principle the safeguarding of non-combatants in warfare. In previous ages it was much easier to make the combatant/non-combatant distinction (brightly colored and distinctive uniforms were worn partly for this reason), but in the age of guerrilla warfare, camouflage, terrorism, and so-called "total war" in a technocratic age, one might wonder just who is *not* a combatant of some kind in a country at war. Suffice it to say that principled distinctions can and are made by natural law theorists, for instance between a military codebreaker and a singer who entertains the troops. (See further, Oderberg 2000b: ch. 5; Plater 1915: Part II.)

This is a mere snapshot of the way in which natural law theory approaches the topic of warfare. As usual the devil is in the detail, and one could do worse than begin with Grotius's magisterial work. Needless to say, there is far more to inter-state relations than war. Natural law theory has much to say about international organization, the force of treaties, the use of diplomacy, and many other topics that are the staple of international law. Whether the theory countenances supranational authority is a vexed question. The major documents of the United Nations, such as the Charter itself and the Universal Declaration of Human Rights, bear the stamp of influence by natural law theory, but they were formed by many hands and many persuasions. The United Nations itself is avowedly secularist in its approach to international law, not just as a matter of practice but of principle. More importantly, its bias is globalist, as is that of most international lawyers and tribunals. Applying the principle of subsidiarity to international relations should lead us to the view that states are better placed to work out their differences on a bilateral or multilateral basis than by ceding sovereignty to unaccountable supranational structures. Matters of genuine global concern, such as the law of the sea, the stewardship of the environment, and the use of space, if not manageable at the bilateral or multilateral level, might benefit from the cession of some genuine power to global bodies. But these should, as far as subsidiarity goes, be the exception rather than the rule, even in our highly interconnected world.

Related Topics

Plato's Political Philosophy, Aristotle's Social and Political Philosophy, Aquinas, Medieval Political Thought, War, Rights

References

Cahill, E., S.J. (1932) *The Framework of a Christian State*, Dublin: M.H. Gill and Son, Ltd.
Cronin, M. (1930) *The Science of Ethics*, 2nd edn., vol. 1, Dublin: M.H. Gill and Son, Ltd.
—— (1939) *The Science of Ethics*, 2nd edn., vol. 2, Dublin: M.H. Gill and Son, Ltd.
Crowe, M.B. (1977) *The Changing Profile of the Natural Law*, The Hague: Martinus Nijhoff.
Feser, E. (2009) *Aquinas: A Beginner's Guide*, Oxford: Oneworld.
—— (2010) "Classical Natural Law Theory, Property Rights, and Taxation," *Social Philosophy and Policy* 27: 21–52.
Finnis, J.M. (1980) *Natural Law and Natural Rights*, Oxford: Clarendon Press.
——, Boyle, J., and Grisez, G. (1987) *Nuclear Deterrence, Morality and Realism*, Oxford: Clarendon Press.
Foot, P. (2001) *Natural Goodness*, Oxford: Oxford University Press.
Hamilton, E. and Cairns, H. (eds.) (1961) *The Collected Dialogues of Plato*, Princeton, NJ: Princeton University Press.
Maritain, J. (1971) *The Rights of Man and Natural Law*, New York: Gordian Press.
Murphy, M.C. (2006) *Natural Law in Jurisprudence and Politics*, Cambridge: Cambridge University Press.
Nozick, R. (1974) *Anarchy, State and Utopia*, New York: Basic Books.
Oderberg, D.S. (2000a) *Moral Theory*, Oxford: Blackwell.
—— (2000b) *Applied Ethics*, Oxford: Blackwell.
—— (2004) "The Structure and Content of the Good," in D.S. Oderberg and T. Chappell (eds.), *Human Values: New Essays on Ethics and Natural Law*, London: Palgrave Macmillan.
—— (2010) *The Metaphysical Foundations of Natural Law*, in H. Zaborowski (ed.), *Natural Moral Law in Contemporary Society*, Washington, DC: Catholic University of America Press.
Phelan, G.B. and Eschmann, I.T., trans. and eds. (1949) *De Regno: On Kingship to the King of Cyprus, by Thomas Aquinas*, Toronto: Pontifical Institute of Medieval Studies.
Plater, C., S.J. (ed.) (1915) *A Primer of Peace and War: The Principles of International Morality*, New York: P.J. Kenedy and Sons.
Regan, R.J. (ed.) (2002) *Aquinas on Law, Morality, and Politics*, Indianapolis: Hackett Publishing Company, Inc.
Reinhardt, T. (ed.) (2007) *Seneca: Dialogues and Essays*, Oxford: Oxford University Press.
Ross, W.D. (ed.) (1925) *The Works of Aristotle*, vol. 9, Oxford: Clarendon Press.
—— (1946) *The Works of Aristotle*, vol. 11, Oxford: Clarendon Press.
Rothbard, M.N. (1982) *The Ethics of Liberty*, Atlantic Highlands, NJ: Humanities Press.
Tuck, R. (ed.) (2005) *The Rights of War and Peace by Hugo Grotius*, Indianapolis: Liberty Fund.

Further Reading

For a useful historical survey of natural law theories, see F. Oakley, *Natural Law, Laws of Nature, Natural Rights* (New York: Continuum, 2005). H.P. Kainz, *Natural Law: An Introduction and Re-examination* (Chicago and La Salle, IL: Open Court, 2004) combines history with accessible discussion of contemporary issues including the differences between classical and new natural law. M.C. Murphy, *Natural Law and Practical Rationality* (Cambridge: Cambridge University Press, 2001) provides a rigorous and useful contemporary exposition of natural law theory that has both classical and new natural law aspects. Jacques Maritain, *Natural Law: Reflections on Theory and Practice* (South Bend, IN: St Augustine's Press, 2001) collects some of the more important writings by one of the most prominent natural law theorists of the twentieth century. J.W. Evans and L.R. Ward (eds) *The Social and Political Philosophy of Jacques Maritain* (London: Geoffrey Bles, 1956) is also useful in this regard. T.J. Higgins, S.J., *Man as Man: The Science and Art of Ethics* (Rockford, IL: Tan Books, 1958 [reprint]) gives a clear and accessible introduction both to basic theory and social/political

issues. On economics there is not a lot of material available, but A. O'Rahilly's *Money* (Cork: Cork University Press, 1941) is a detailed exploration of economics from a classical perspective. On war and international law generally, see J. Eppstein, *The Catholic Tradition of the Law of Nations* (London: Burns, Oates and Washbourne, 1935), which includes primary source material such as religious and secular documents, quotations from the tradition, as well as commentary.

Part IV

DISTRIBUTIVE JUSTICE

34
LUCK EGALITARIANISM

Zofia Stemplowska

As Andrew Williams once observed (2006a: 187), one attraction of political philosophy is that with relatively weak assumptions it can deliver radical conclusions. Luck egalitarianism is a paradigm example. According to luck egalitarianism, disadvantages that are due to the disadvantaged person's bad luck are inegalitarian (i.e. objectionably unequal), while disadvantages that are of one's own making are not. The view is composed of three elements. First, it assumes that some form of equality is a key requirement of justice. Call this the requirement of egalitarian justice. Second, it is anti-luck, in that it sees disadvantages that are due to luck as upsetting equality and in this sense objectionable. Third, it is pro-choice (or pro-"choicist" to avoid famous connotations), in that it sees disadvantages that are of one's own making (or, for short, due to choice) as consistent with equality, and in this sense acceptable. The three elements are distinct, but they interact with each other and are sometimes, unhelpfully, run together.

The most intuitive application of luck egalitarianism concerns situations of scarcity in which we must decide whether to assist the unlucky person or the person who brought the disadvantage upon herself. Imagine that both Elinor and Marianne need life-saving medical treatment but that resources allow only one of them to be saved. Elinor's condition is due to an inherited genetic trait. Marianne's is due to her reckless recreational driving. The two women, their recovery prospects, and the costs of treating them are similar in all other relevant respects. Whom should we treat? The standard luck egalitarian answer is that we must treat the person whose condition is due to luck rather than of her own making. True, often we will find it difficult to determine the source of one's disadvantage but, if appearances are not misleading, it seems that, in this case, Elinor is a victim of bad luck while Marianne's condition is of her own making. If so, egalitarian justice requires that Elinor rather than Marianne be helped.

More controversially, even if Elinor's life were not at stake and her genetic condition consisted merely in shortsightedness, but we still had to choose between treating her and saving Marianne's life, it seems that luck egalitarians would require us to treat Elinor ahead of Marianne. After all, Elinor's disadvantage is due to luck and Marianne's is of her own making. In Section 4 I will cast doubt on whether this widespread reading of luck egalitarianism is justified.

1. Luckism, Choicism and Egalitarianism

Consider, first, the anti-luck or anti-"luckism" element of luck egalitarianism. (I borrow the term "luckism" from Arneson (2004: 2) though I use it differently.) In rejecting luckism, luck egalitarians object to some being worse off than others due to bad luck.

Being a victim of bad luck is, by definition, "arbitrary from the moral point of view" (Rawls 1971: 312)—luck's victims are not chosen because they ought to be affected; luck strikes with no purpose. It is this arbitrariness of luck that gives us a reason to neutralize its disruptive impact on whatever we take to be a desired distribution. Of course, neutralizing luck might not always be possible but, sometimes, it can be done by spreading its burdens more equally; by compensating its victims from the resources of those who have had good luck; by preventing bad luck from arising; or by preventing situations in which some bear a greater ex ante chance of being its victims than others.

Given luck's arbitrariness, many luck egalitarians adopt a symmetrical view about good luck and bad luck, objecting also to some being *better off* than others due to good luck. In practice, though, the symmetrical and asymmetrical versions rarely diverge since, often, one person's advantage due to luck (i.e. good luck for her) is someone else's relative disadvantage due to luck (i.e. bad luck for her). But whether symmetrical or not, luck egalitarianism's objection is not to the presence of luck in human life as such but to its ability to move people away from equality. Importantly, then, luck egalitarians are not egalitarian because they object to luck, rather their egalitarianism is an independent conviction of theirs. Indeed, neutralizing luck, as Susan Hurley once observed, would not lead to equality without a prior commitment to seeing equality as the baseline against which to measure luck (2003: 146–80).

In addition to their commitments to equality and anti-luckism, luck egalitarians are also pro-choicist. According to the last, equality does not require the elimination of disadvantages that are of people's own making (call them chosen or avoidable). The alternative names for luck egalitarianism—choice-sensitive egalitarianism or responsibility-sensitive egalitarianism—capture this feature. In accepting choicism, luck egalitarianism seems to appeal to the principle "you broke it, you fix it," though, as I will explain in Section 4, the luck egalitarian stance can be considerably more nuanced than this principle suggests.

The luck egalitarian rejection of luckism and the acceptance of choicism are logically distinct features of the view: a theory could object to luck-generated disadvantages on grounds of arbitrariness (i.e. be anti-luckist) and, unlike luck egalitarianism, also object, on different grounds, to chosen disadvantages (i.e. be anti-choicist). But the two features go hand in hand since, once combined, they single out luck-generated differences, as opposed to chosen ones, as the only markers of unacceptable inequality.

Importantly, while neither the rejection of luckism nor the acceptance of choicism suffice to commit one to equality—one can still insist, say, that some ought to have more than others—nonetheless they do help to define the content of what ought to be equalized. In other words they help to define the content of what has come to be known as the currency of egalitarian justice.

The luck egalitarian currency is determined by the application of two filters through which disadvantages must pass to qualify as genuinely equality upsetting. These are the luck vs. choice filter and what we might call the relevance filter (though sometimes, mistakenly, the currency is equated simply with the second filter). To see the relevance filter in operation consider two distinct luck egalitarian views that use different filters. According to the first view, equality consists in equal hedonistic pleasure—differences that do not affect hedonistic pleasure are classified as irrelevant as far as equality is concerned. Thus an unemployed person in a wheelchair is seen as no worse off than someone with a job who can walk if they both enjoy the same amount of hedonistic pleasure. According to the second view, equality consists in equal objective well-being—dif-

ferences in hedonistic pleasure that do not affect objective well-being are irrelevant. Both views use different relevance filters but both, being luck egalitarian, apply the luck/choice filter. That is, disadvantages that pass the relevance filter must also pass the luck or choice test: if they are due to luck then they are classified as genuinely equality upsetting, if they are of the person's own making then they are not. A helpful way of thinking of how these filters work together is to think of the relevance filter as defining the baseline of equality from which the choice/luck filter measures departures, ignoring those that are due to choice.

Famous luck egalitarian currencies of egalitarian justice include Arneson's equality of opportunity for welfare, Cohen's equality of opportunity for welfare and resources (also known as equality of access to advantage), Dworkin's equality of ambition-sensitive endowments, Rawls's equal expectations of primary goods. (See Arneson 1989; Cohen 2011 [1989]; Dworkin 2000 [1981]; Rawls 1971. Whether Rawls can be read as subscribing, in some form, to luck egalitarianism is a subject of controversy; see Freeman 2006; Kymlicka 2002; Quong 2010; Scheffler 2003.)

In sum, luck egalitarianism is an attempt to offer a vision of equality. In this vision disadvantages due to luck are equality upsetting while those of people's own making are not. But the vision splinters into various positions depending on whether welfare, resources, primary goods, or another dimension across which people's standing can be compared, is considered most important for equality. All this raises a lot of questions. Perhaps the most obvious one is what would it mean for a disadvantage to be due to luck.

2. Luck and Responsibility

a. Brute Luck

The luck to which rank-and-file luck egalitarians object is brute luck. Inheriting a genetic condition, being struck by a loose wing-mirror falling off a bus, being born in a poor country, or in a brutal dictatorship, or in a liberal-egalitarian society are all examples of (bad and good) brute luck to the person to whom they happen. But generalizing from the examples and defining brute luck is tricky. One intuitive approach sees brute luck as something with an impact on one's life but which the person did not control. All the above examples seem, at first glance at least, to fit the bill. But to equate brute luck with the absence of control is too narrow. Being struck by a loose mirror falling off a bus is usually related to one's prior decision to go near the bus; even the development of a genetic disorder will often relate to one's prior decision not to commit suicide, which gives the disorder time to develop. One solution might be to try to see the occurrences as partly due to brute luck and partly a matter of control. Another might be to stipulate that outcomes of choices with no acceptable alternative (e.g. committing suicide to avoid a genetic disorder) or extremely unlikely outcomes (e.g. being struck by the bus mirror) count as brute luck. Either way the absence of any control over an outcome does not seem to be a necessary condition for the outcome to be seen as a matter of brute luck.

There is a further difficulty. Some believe, for example, that even if all our choices are entirely determined by factors we did not control (our genes, family background, people we encounter, trains we miss), reading *Sense and Sensibility* on holiday can be seen, in a meaningful sense, as a choice. On the above definition of brute luck as the absence of control, this could make reading the novel simultaneously a matter of luck

and choice. Such a conceptualization would then not allow luck egalitarians to sort between disadvantages that are due to brute luck as opposed to choice. So the luck egalitarian framework gives us a reason to define brute luck in what Hurley (2001) has referred to as a "thin" way—as simply the inverse correlate of choice (or responsibility or whatever we think makes disadvantages of the agent's own making). Whatever that is (and whether any living person can enjoy it), brute luck is the opposite of it. And, indeed, iconic formulations of luck egalitarianism—that object when some are worse off than others due to no fault or choice of their own (Temkin 1993: 13; 2011: 62; cf. Cohen 2011 [1989]: 29)—assign luck precisely this derivative meaning by not even mentioning luck in the definition.

b. Responsibility

This thing with reference to which luck is to be defined as its opposite, is usually referred to as "choice" or "responsibility." Responsibility here does not refer to what Scanlon has called substantive responsibility (1998: 248). To be substantively responsible for a disadvantage means that it is duly one's to bear. Knowing who is substantively responsible for a disadvantage is important but if luck were contrasted with substantive responsibility then we would not be able to ask whether disadvantages that someone is substantively responsible for are hers to bear while those due to luck are not, since it would be so by definition. Plus we would still need to know what makes it the case that someone ends up substantively responsible for the disadvantage. Similarly, we could try to identify a cut between what is a matter of luck and what people deserve, as some luck egalitarians do, but, again, we will then need to discover what makes one deserving of a given outcome (Temkin 2011; Knight 2011).

The type of responsibility with which brute luck is to be contrasted tells us only whether a given action or outcome is genuinely of the agent's own making (no matter if it is also hers to bear). Attributing actions and outcomes to agents in this sense will, most plausibly, require two elements: the involvement of the agent's will (her agent responsibility) and also circumstances that do not restrict too much the scope for the exercise of the will.

Consider the agent's will. What must it be like for us to see what she does as of her own making? Usual suspects include freely willed choice; acting in a way one would if one enjoyed the ability to act otherwise (even if one lacks it); acting from a reason-responsive mechanism (e.g. one's actions are sensitive to reasons); acting from preferences one endorses; acting in the absence of addictions and phobias; or, simply, acting while conscious. These various conditions are meant to establish someone's agent responsibility (or her attributive responsibility; Scanlon 1998: 248).

But it also matters that the agent's actions do not merely reflect the absence of any acceptable alternatives to her adopted course of action. So brute luck is the opposite of what we might call avoidable responsibility, which is agent responsibility in the presence of an acceptable alternative. That said, for simplicity, "avoidable responsibility" can be referred to as "choice" and I will continue talking of the luck vs. choice distinction.

c. Option Luck

There are also cases of disadvantage that seem to arise out of a mixture of brute luck and choice. This is the domain of option luck. Dworkin, who has coined the term along with

that of brute luck, defined option luck as "a matter of how deliberate and calculated gambles turn out—whether someone gains or loses through accepting an isolated risk he or she should have anticipated and might have declined" (2000 [1981]: 73). The definition is ambiguous when it comes to gambles that were not deliberate and calculated but should have been anticipated and could have been declined, such as those we regularly make when, out of habit, we use old electric appliances or open the door to perfect strangers without considering the risks.

There are other definitional difficulties but Dworkin's brilliant insight is that insurance can transform brute into option luck. Having where one lives consumed by fire may be an instance of bad brute luck, but if affordable insurance was available then not having the money to rebuild it is a matter of option luck. So, if affordable insurance is available, brute luck can rule our lives much less than it would at first appear.

But are disadvantages due to option luck any more acceptable than those due to brute luck? Dworkin answers in the affirmative (for criticism see Lippert-Rasmussen (2001) and Vallentyne (2002)). His view is that, unlike brute luck, option luck should lie where it fell. Option luck creates three categories of people: those who gambled, so to speak, and won, those who gambled and lost, and those who did not gamble. Those who did not gamble end up worse off than those who gambled and won but, according to Dworkin, they have no complaint because

> people should pay the price of the life they have decided to lead, measured in what others give up in order that they can do so … But the price of a safer life [no gambling], measured in this way, is precisely foregoing any chance of the gains whose prospect induces others to gamble.
>
> (Dworkin 2000 [1981]: 74)

By extension, those who gambled and lost—either because their investments did not work out or, in the case of insurance option luck, the cover they spent money on wasn't needed—cannot complain since they had the option not to gamble. And, finally, those who gambled and won are entitled to their trophy since no one can object. Note, however, that there remains a question whether the gambles should have existed to begin with.

3. Insulting, Harsh and Exploitative?

Luck egalitarianism has many followers and many just as committed critics. The most important criticisms come from fellow egalitarians who believe that luck egalitarianism ought to be rejected in favor of social (or democratic) egalitarianism. Social egalitarians criticize both the relevance filters used by luck egalitarians and their use of the luck/choice filter. Before reviewing the objections, it is worth noting the two defining features of social egalitarianism. First, it sets its relevance filter to register only inequalities in the social and political standing of people as citizens (also referred to as relational inequalities). On this view, there may be no troubling inequality between a lecturer living on £25,000 a year and a manager living on £250,000 a year if they are both citizens of equal standing in the society with neither being socially excluded or dominated. This position is, in principle, open to luck egalitarians but no prominent luck egalitarian holds it (including, in addition to those already mentioned, Otsuka (2006), Roemer (1993), Rakowski (1991), and others who might qualify as luck egalitarians: Van Parijs

(1995), Steiner (1997)). Second, social egalitarians explicitly reject the luck/choice filter (though, as I will argue, the rejection is not absolute). They emphasize that it is inegalitarian if someone cannot function as an equal citizen in society, no matter whether their situation is due to luck or choice (Anderson 1999; Satz 2010; Scheffler 2003; Miller 1999: 230–44). So Elinor and Marianne from the opening example would be just as entitled at the bar of egalitarian justice to the life-saving operation.

a. Insulting?

Social egalitarians attempt to justify their narrow relevance filter by pointing out that it is insulting to consider some differences as disadvantages that ought to be eliminated. Seeing deafness or ugliness, for example, as a disadvantage, even if the deaf and the ugly can function as equal citizens, shows lack of respect and offensive pity. If deafness or ugliness is a disadvantage it is only to the extent that the deaf are not treated as equals by their co-citizens. Anderson (1999) has offered a memorable parody of a zealous official of a luck egalitarian utopia who informs the ugly through a letter that their ugliness is sufficient to qualify them for the equality restoring compensatory grant.

There are two issues here. One issue is whether to see deafness, ugliness, involuntary creepiness or lack of marketable talent as markers of disadvantage. The other is how to respond to it if the answer to the first question is positive. Regarding the first issue there is less disagreement than might at first appear. Social egalitarians, too, would accept that deafness that leads to social exclusion is problematic for egalitarians. It is the fact of social exclusion that is problematic but, to deal with the problem, it helps to identify one of its sources. Similarly, luck egalitarians can note that what they find problematic is when the deaf or the ugly or the creepy or those who faultlessly developed no marketable talents get systematically worse jobs and enjoy worse opportunities in life. The problem is the lack of opportunities. And it is unclear why it would be insulting to point out to someone who earns little that she is disadvantaged in many ways vis-à-vis those who earn a lot. She already knows this if she has ever attempted to buy a house in their neighborhood.

Regarding the second issue of the appropriate response to disadvantage, Anderson (1999) has argued that the proper way of dealing with the difficulties faced by the disabled, or the ugly, is through changing social attitudes, not through targeted, direct compensation. Luck egalitarians can respond that there is nothing in their doctrine that would make them opposed to this method of neutralizing bad luck. But Anderson is right to note that they might favor direct compensation should the social attitudes remain unchangeable. Such direct compensation could undermine the self-respect of its recipients (Wolff 1998) if it is offered only to those branded as untalented or ugly. This gives luck egalitarians a reason not to target direct compensation too closely.

b. Harsh?

What of the social egalitarian critique of the use by luck egalitarians of the luck/choice filter? Using the filter, they point out, is too harsh and also exploitative in that it can exacerbate the vulnerability of the already disadvantaged. In virtue of these two problems it is also inegalitarian.

The harshness objection is best introduced through an example. Recall the reckless driver, Marianne. But assume that, in this scenario, Elinor has not inherited the life

threatening genetic condition and so we could rescue Marianne without sacrificing Elinor's rescue (this example is adapted from Fleurbaey 1995). Nonetheless, if Marianne's accident is of her own making, luck egalitarianism seems to lack the conceptual resources to require that she be assisted as a matter of egalitarian justice. The available resources should be distributed equally between her and Elinor and everyone else in society, rather than given simply to Marianne who has already enjoyed the opportunity to lead a long and healthy life. Assume, until the next section, that this must indeed be the luck egalitarian view. The social egalitarian claim is that seeing Marianne as not entitled to assistance is too harsh and inegalitarian.

Luck egalitarians, such as Cohen (2011 [2006]), have responded by drawing a distinction between egalitarianism concerned with fairness and one concerned with fraternity. Not assisting Marianne, they concede, violates the requirements of egalitarian fraternity—that there be no social gulf between people—but preserves the egalitarianism of fair shares. If Marianne has already been given as good opportunities as others to do well, then to assist her when she has squandered them is to make others subsidize her choices. Assisting her might well be required as a matter of fraternity or as a matter of humanitarian concern to prevent suffering—and, if so, Marianne retains a claim to assistance—but it is not a claim of egalitarian justice when the latter is meant to preserve fair shares.

This response raises the question of the status of the luck egalitarian requirements and prohibitions. Does luck egalitarianism identify what all-things-considered distributive justice requires? If so, luck egalitarianism is compatible with recognizing that there can be reasons of charity to assist those whose disadvantage is of their own making, but it rules out assistance as a requirement of justice. Does it identify, instead, what egalitarian justice requires—or egalitarian justice concerned with fair shares in particular—but sees egalitarian justice as only one component of all-things-considered justice that must be sometimes curtailed by concern with freedom, privacy, solidarity, efficiency? For example, it could tell us that the moral value of an extra unit of welfare for a person is "greater, the larger the degree to which the person deserves this gain [on account of her prudent choices]" (Arneson 1999: 239–40). Or does it merely put forward a consideration that is supposed to apply only in cases where we need to break ties between those who are equally in need of assistance, as in the first case of Elinor and Marianne? The dominant view is that luck egalitarianism is a component principle of justice (or, in Cohen's vocabulary, a component principle of rules of regulation (Cohen 2008)). As such, it needs to be balanced with other principles before we can know what people are due.

c. *Exploitative?*

In addition to being harsh, luck egalitarianism has also been accused of ignoring, if not justifying, the plight of some vulnerable people. Leading the charge, Anderson has pointed out that those—usually women—who perform the socially useful and morally worthy task of caring for dependants often end up poor. Their efforts are not only badly rewarded but also take up time and energy, preventing them from seeking more lucrative occupations. Nonetheless, for some, becoming carers has been a choice. So their disadvantage would be of no concern to luck egalitarians.

One, admittedly partial, answer is available to luck egalitarians. They can point out that if carers perform a duty of care that falls not just on them but also on others who shrink from their responsibilities, then the carers are owed compensation. That is, those who have shrunk from their responsibilities—even if only by entering a de facto game

of chicken with others, safe in the knowledge that those with the caring disposition will blink first and do whatever is necessary to take care of others—owe the carers fair compensation for their efforts. On this view, improving the position of the carers has nothing to do with the fact that they end up disadvantaged but this argument is available to luck egalitarians.

However, the problem of the disadvantaged carers simply reemerges if the carers are acting beyond the call of duty or if others do not share the responsibility for the dependants. In fact, the problem can be restated in a more general form (Eyal 2007; Lippert-Rasmussen 2011; Temkin 2011): how should we respond to disadvantage that is due to someone's choice to do the moral thing when others are not required to assist one with this task?

Taking a step back, notice that the objections canvassed above assume that leaving people to bear the costs of their choices could mean leaving them in dire straits. But this assumption needs to be examined.

4. The Egalitarian Baseline: Hard Board or Soft Cushion?

Consider once more the case of the reckless Marianne and assume that we have the resources to treat her. We might decide—having bitten the bullets offered by social egalitarians—that we have a non-egalitarian duty to treat her but that she should repay us later; she is not entitled to the treatment for free as a matter of egalitarian right. She has made her choice to be reckless and she should now pay the full price for it. But asking Marianne to pay the full price of her choice raises the question of what the right price should be (Olsaretti 2009; Stemplowska 2009, 2011). How do we know that the right price for recklessness is the market costs of her medical treatment? Why isn't the right price for recklessness instead a requirement to wear a badge that says "we are disappointed in you, Marianne!"?

One reason why it seems that luck egalitarians are committed to setting the price at the market cost of the medical treatment is that they are committed to providing people with equality of opportunity, not equality of outcome. Marianne's choices unsettle equality of outcome but luck egalitarians, being pro-choicist, find this acceptable. The luck egalitarian view is, however, less straightforward. The iconic statement of luck egalitarianism—people should not be worse off than others due to no choice or fault of their own—is compatible with two main readings.

a. Luck Egalitarianism Version 1

On the first reading, luck egalitarianism requires that departures from the egalitarian baseline due to luck be neutralized while departures due to choice be accepted. But this tells us nothing about how we ought to specify what counts as the egalitarian baseline. The baseline could include the equal provision for all of opportunities that are very generous and cushion the consequences of recklessness or imprudence. The general thought is that there are many ways of structuring the initial baseline of equal opportunities—one way of doing it is to set up generous guarantees of assistance, for everyone, in case they end up badly off for whatever reason.

This might be thought to be against the spirit of luck egalitarianism. After all, even if departures from the baseline are still possible due to choice—Marianne might crash her car on purpose—such departures would be rare. And the more the initial baseline cushions the impact of mistakes or foolishness, the less scope there is for the luck/choice

filter to operate. But, in response, notice that the iconic formulation of luck egalitarianism is agnostic on how prominent a role the luck/choice filter must have: we only know that should people become disadvantaged due to choice relative to the baseline—whatever it might be—their disadvantage does not upset equality.

Some might object that the prudent and the hard working would be forced to subsidize the reckless and the lazy so that the cushioning baseline cannot count as egalitarian. But the baseline makes the prudent and the hardworking no worse off than the lazy and the imprudent are: they all equally enjoy the same opportunities to be imprudent and lazy and yet end up as well off as everyone else.

On this first reading luck egalitarianism loses some of the distinctiveness that has been attributed to it. But it still succeeds, in the terms set out by Cohen, in incorporating into egalitarianism "the most powerful idea in the arsenal of the antiegalitarian Right: the idea of choice and responsibility" (Cohen 2011 [1989]: 32). It also shows that merely pointing out that a disadvantage is of one's own making tells us nothing about what to do with it in the absence of a plausible account of the opportunities that people ought to have enjoyed to begin with.

b. Luck Eglitarianism Version 2

The second reading of luck egalitarianism sees it as offering directions regarding the nature of the baseline. According to this reading, not only must departures from the baseline that are due to luck be neutralized while those that are due to choice do not count, but also the baseline itself should be constructed in such a way as to allow maximum room for the luck/choice filter to differentiate between people. In effect, the opportunities we offer people must be such that people who make different choices end up with different outcomes: the hardworking and the prudent can maximize their advantage vis-à-vis the imprudent and the lazy (they can have, for example, what they would enjoy had they not shared the world with the imprudent and the lazy).

For the most part luck egalitarianism has been interpreted in this second way. This reading is one that fits well with desert luck egalitarianism whose justification for leaving those whose disadvantages are of their own making without assistance is that the prudent deserve to reap the rewards of their prudence. This second reading can also try to appeal to the thought that setting up the baseline in such a way as to allow the hardworking and the prudent to maximize their rewards maximizes everyone's freedom. But it is not clear that freedom is maximized. True, the prudent and the hard working acquire the option to use the resources they accumulate to purchase other things than the bailing out of the lazy and the imprudent. But the lazy and the imprudent lose the option to act in a lazy or imprudent way without suffering bad consequences or, at least, without ending up worse off than others.

Must luck egalitarians adopt this second reading? It might appear that pro-choicism requires that the baseline does not cushion the consequences of different choices: different choices should lead to different outcomes. But notice that different choices could be allowed to result in different outcomes even if the baseline did cushion them. The imprudent and the lazy could simply be asked, for example, to pay $100 extra a month towards some scheme and that would be the extent of the difference.

My main point is that although luck egalitarianism has for the most part been interpreted according to the second reading, this is not the only coherent reading of the doctrine. The doctrine is not committed to maximizing the role of the choice/luck filter.

Furthermore, this second reading of luck egalitarianism—while assumed by their critics—is not explicitly endorsed by Arneson, Cohen or Temkin (with the latter pointing out that his and Cohen's early statements were ambiguous; Temkin 2011: 62–3, n. 24). Moreover, as the critics of luck egalitarianism have long pointed out, the implications of the second reading are counterintuitive. This version of luck egalitarianism welcomes, for example, setting up the (criminal) justice system so as to prioritize the claims of those who made no choices increasing the chance that they be victims of crime. Those who opened doors to strangers who turn out to be robbers, or who dressed "provocatively" when going out late at night and were sexually assaulted, would have to wait in line until the claims of the more prudent citizens facing similar problems were met.

c. *Which Egalitarianism?*

Even if the second reading is rejected, luck egalitarianism need not opt for the maximally cushioning baseline. Such a baseline is also implausible. It is implausible even if the baseline concerns only the relational good of equal status as citizens. Indeed, social egalitarians appear to concede this. Thus Anderson notes that people "do not owe one another the real freedom to function as beach bums" (1999: 321). Indeed, she holds that "[i]n the typical case of an able-bodied adult, for instance, access to a decent income would be conditioned on responsible performance of one's duties in one's job, assuming a job was available" (1999: 328). These ideas are echoed by Scheffler who, approvingly explaining Rawls's view, notes that "[p]eople are asked to accept responsibility for their ends, in Rawls's sense … because it is reasonable to expect people to make do with their fair shares" (Scheffler 2005: 27). In sum, the cushion of the baseline does not stretch to cover cases where reckless and wasteful decisions of some would eat too much, if compensated for, into the resources of others. This is sensible, in my view. Due to scarcity of resources, the choices one person makes can affect the amount of resources available for others. And because we are beings with separate lives in circumstances of scarcity, we must respect each other as holders of entitlements that offer us some protection against the potentially expensive choices of others.

What *any* egalitarian view needs, then, is a systematic way of designing the baseline, deciding which choices should not lead to disadvantage (Williams 2006b). One possible answer is that we must ask whether acting to avoid a disadvantage requires greater sacrifice than offering compensation should it arise. For example, if private insurance is prohibitively expensive and the interests people have in not needing to avoid opening their houses to strangers (lest it leads to robbery) or cycling on an old bike (lest it leads to costly accident rescue) are more important than the interest people have in retaining the resources that are needed to compensate for such disadvantages, then compensation should be offered as a matter of policy. But offering compensation to those whose disadvantage is the result of drink-driving or going bungee jumping uninsured would not be required if avoiding such activities jeopardizes less important interests than providing compensation for the resulting disadvantage would. With regard to the standard example of imprudent conduct—smoking—our answer would depend on whether we agree with the former British Secretary of State for Health, John Reid, who suggested that for many very poor people (he singled out lone mothers on council estates) smoking is an essential and rare escape from stress. My point here is only that this is exactly the right question to ask: Does acting to avoid a disadvantage require an unreasonable sacrifice given the various interests at stake?

Judging when one set of interests outweighs another is not going to be easy. Three main routes are open to us. We can have a democratic vote. We can appeal to some objective preference ranking. Or we can opt for an idealized voting situation, a version of which is Dworkin's hypothetical insurance model (2000). On Dworkin's model individuals are asked what they would be willing to sacrifice to avoid a certain disadvantage. The question is posed in a hypothetical setting, in which people are informed about the general risks and costs but not their own risks of suffering a disadvantage (be it low salary or poor health). These answers are then aggregated, but what matters here is that the model attempts to weigh interests on both sides: the interests sacrificed by needing to bail out the disadvantaged and the interests sacrificed by staying disadvantaged or acting in a way to avoid disadvantage.

Whether or not we accept Dworkin's model, we should see luck egalitarianism as a view that is versatile enough to incorporate a whole range of payoffs for different choices. The commitment to equality, anti-luckism and pro-choicism rules out few substantive takes on what should be the point of equality.

Acknowledgments

I am grateful to Matthew Clayton, Fred D'Agostino, Jerry Gaus, Ben Jackson and Jonathan Quong for comments. I am also grateful to Fred D'Agostino for brilliant editorial help.

Related Topics

Left Libertarianism, Desert, Needs and Distributive Justice, The Capability Approach (and Social Justice), Equality, Health

References

Anderson, E. S. (1999) "What is the Point of Equality?" *Ethics* 109: 287–337.

Arneson, R. J. (1989) "Equality and Equal Opportunity for Welfare," *Philosophical Studies* 56: 77–93.

—— (1999) "Egalitarianism and Responsibility," *The Journal of Ethics* 3: 225–47.

—— (2004) "Luck Egalitarianism Interpreted and Defended," *Philosophical Topics* X: 1–20.

Cohen, G. A. (2008) *Rescuing Justice and Equality*, Cambridge, MA: Harvard University Press.

—— (2011) *On the Currency of Egalitarian Justice, and Other Essays in Political Philosophy*, ed. M. Otsuka, Princeton, NJ: Princeton University Press.

Dworkin, R. (2000) *Sovereign Virtue*, Cambridge, MA: Harvard University Press.

Eyal, N. (2007) "Egalitarian Justice and Innocent Choice," *Journal of Ethics and Social Philosophy* 2: 1–18.

Fleurbaey, M. (1995) "Equal Opportunity or Equal Social Outcome?" *Economics and Philosophy* 11: 22–55.

Freeman, S. (2006) *Justice and the Social Contract*, Oxford: Oxford University Press.

Hurley, S. (2001) "Luck and Equality," *Proceedings of the Aristotelian Society* Supplementary Volume 75: 51–72.

—— (2003) *Justice, Luck, and Knowledge*, Cambridge, MA: Harvard University Press.

Knight, C. (2011) "Responsibility, Desert, and Justice," in C. Knight and Z. Stemplowska (eds.), *Responsibility and Distributive Justice*, Oxford: Oxford University Press.

Kymlicka, W. (2002) *Contemporary Political Philosophy*, 2nd edn, Oxford: Oxford University Press.

Lippert-Rasmussen, K. (2001) "Equality, Option Luck, and Responsibility," *Ethics* 111: 548–79.

—— (2011) "Luck Egalitarianism and Group Responsibility," in C. Knight and Z. Stemplowska (eds.), *Responsibility and Distributive Justice*, Oxford: Oxford University Press.

Miller, D. (1999) *Principles of Social Justice*, Cambridge, MA: Harvard University Press.

Olsaretti, S. (2009) "Responsibility and the Consequences of Choice," *Proceedings of the Aristotelian Society* 109: 165–88.

Otsuka, M. (2006) *Libertarianism without Inequality*, Oxford: Oxford University Press.
Quong, J. (2010) "Justice Beyond Equality," *Social Theory and Practice* 36: 315–40.
Rakowski, E. (1991) *Equal Justice*, Oxford: Oxford University Press.
Rawls, J. (1971) *A Theory of Justice*, Cambridge, MA: Harvard University Press.
Roemer, J. E. (1993) "A Pragmatic Theory of Responsibility for the Egalitarian Planner," *Philosophy and Public Affairs* 22: 146–66.
Satz, D. (2010) *Why Some Things Should Not Be For Sale*, Oxford: Oxford University Press.
Scanlon, T. M. (1998) *What We Owe to Each Other*, Cambridge, MA: Harvard University Press.
Scheffler, S. (2003) "What is Egalitarianism?" *Philosophy and Public Affairs* 31: 5–39.
—— (2005) "Choice, Circumstance, and the Value of Equality," *Politics, Philosophy, and Economics* 4: 5–28.
Steiner, H. (1997) "Choice and Circumstance," *Ratio* 10: 296–312.
Stemplowska, Z. (2009) "Making Justice Sensitive to Responsibility," *Political Studies* 57: 237–59.
—— (2011) "Responsibility and Respect," in C. Knight and Z. Stemplowska (eds.), *Responsibility and Distributive Justice*, Oxford: Oxford University Press.
Temkin, L. (1993) *Inequality*, Oxford: Oxford University Press.
—— (2011) "Justice, Equality, Fairness, Desert, Rights, Free Will, Responsibility, and Luck," in C. Knight and Z. Stemplowska (eds.), *Responsibility and Distributive Justice*, Oxford: Oxford University Press.
Vallentyne, P. (2002) "Brute Luck, Option Luck, and Equality of Initial Opportunities," *Ethics* 112: 529–57.
Van Parijs, P. (1995) *Real Freedom for All*, Oxford: Oxford University Press.
Williams, A. (2006a) in Morten E. J. Nielsen (ed.), *Political Questions: 5 Questions on Political Philosophy*, Automatic Press/VIP.
—— (2006b) "Liberty, Equality, and Property," in J. S. Dryzek, B. Honig, and A. Phillips (eds.), *The Oxford Handbook of Political Theory*, Oxford: Oxford University Press.
Wolff, J. (1998) "Fairness, Respect, and the Egalitarian Ethos," *Philosophy and Public Affairs* 27: 97–122.

35
THE DIFFERENCE PRINCIPLE

Rex Martin

1. Introduction

Since the publication of his Theory of Justice in 1971, John Rawls (1921–2002) has been the dominant theorist of justice in the English-speaking world and in Western Europe. In Rawls's view justice is, or should be, a virtue of society, specifically of the political and social and economic arrangements that constitute its "basic structure." Justice requires of these basic institutions that, as a set, they exhibit (i) the principle of equal basic rights and liberties and (ii) that of equality of opportunity; in addition they must work together in such a way as to encourage contributions that provide for the production of goods and services, which in turn are so distributed as to (iii) improve, or at least not lower, the level of income and wealth of all the various income groups involved, when considered from the perspective of a representative person's whole life, a normal life. The principle of mutual benefit, (iii), is itself adequately fulfilled only where principles (i) and (ii) are respected. In a just or well-ordered society, inequalities in economic or social positions and attendant inequalities in income and wealth can be allowed—indeed, should be allowed—subject to meeting these three conditions. Since the third of the conditions allows for differences in income and wealth Rawls refers to it as the "difference principle" (for this first formulation of his principles of justice see Rawls 1971: 60–61, 75/1999: 53, 65). The present chapter focuses on Rawls's difference principle.

Rawls typically maintains that the difference principle is closely attached to the principle of equality of opportunity. Indeed, a single main argument is advanced for these two principles, linking them together so as to constitute a single principle of justice. Rawls calls this composite principle his "second principle of justice" (where the idea of equal basic rights and liberties provides the content of his first principle of justice). Rawls's argument begins with the fact that people have different natural endowments and are born into and grow up in different social circumstances, for which they are not responsible. Factors such as natural endowment and initial social circumstance are not negligible; they powerfully affect a person's life prospects, advantageously for some and disadvantageously for others. Indeed, they are among the main sources of inequality between people.

Rawls's main argument for the second principle sets off from this point. He first develops the idea of an appropriate equality of opportunity—conceived as the taking of remedial steps, conscientiously, to *reduce* the initial differential in advantages that accrues to

individuals, arbitrarily, from differences in endowments and social positions. State-supported primary and secondary education (of good quality and at no cost to the individual student) would be an example of a step designed to deal with disadvantages due to initial social circumstances. And public health and other measures (sanitation, clean air and water, vaccinations, adequate prenatal and childhood nutrition) would be an example of a step designed to deal with disadvantages due to natural endowment or to one's pre-birth or early life health history. But measures for equality of opportunity cannot be expected to yield strict or absolute equality once one includes the family, as Rawls does, in the basic structure of society (see Rawls 1971: 74/1999: 64); or once one considers how difficult it would be (it would be impossible in fact) to radically reduce the differences between persons as regards their natural endowments. The leading motif of fair equality, then, becomes trying to make people less unequal at the point where they actually enter into adult life, as citizens and as workers, and to make sure that everyone there, so far as possible, has the basic capabilities required to be contributing members of society.

It is precisely where a fundamental equality in starting points is not fully and strictly achieved, or cannot be, that concern for reducing the inequality of *resultant outcomes* is in order. And since Rawls believes that an absolute equality of opportunity with respect to such starting points can never be achieved, he introduces a further idea—the difference principle—to complement equality of opportunity and to complete the line of argument as regards outcomes. (For Rawls's main argument, see Rawls 1971/1999: secs. 12 and 13.)

The difference principle supplements the idea of conscientiously *reducing* the initial differential in advantages that accrues to individuals. What it adds is the idea of a distribution of income and wealth, via the basic structure, which is to "everyone's advantage" (Rawls, "Justice as Fairness" [1958] in Freeman 1999: 48, 50; Rawls 1971: 60–61/1999: 53). The object of this two-step procedure is to reduce and, ideally, to minimize the gap between persons by taking account of both starting points and end results. "The intuitive idea," Rawls says, "is that the social order is not to establish and secure the more attractive prospects of those better off unless doing so is to the advantage of those less fortunate" (Rawls 1971: 75/1999: 65). Society should not penalize those who, through no fault of their own, are already disadvantaged by their initial social circumstances or by their luck in the draw of the natural endowment lottery. Society must be arranged so that everyone should benefit from the distribution of positions implicated in a particular scheme of differential advantage. This is the understanding of the difference principle—the principle that concerns the production and distribution of resultant benefits—that has been reached in the argument so far. (For the important point about not sacrificing anyone's life prospects see Rawls 1971: 177–8, 180/1999: 155, 157.)

In the next two sections I will try to offer a more complete description of the difference principle and of its justification. The justification will be developed, as it was by Rawls, in two stages. In the next section I will take up what Rawls calls the "informal" or "intuitive" justifying argument. After we've canvassed this particular argument further, we will look (in the section after that) at what it would take (in Rawls's view) for this argument to become a fully formal and conclusive one. Here we will look at Rawls's argument in what he calls the "original position" for the difference principle.

2. The Difference Principle: Its Description and Informal Justification

In the previous section, I relied on the earliest of Rawls's formulations. Here the point of the difference principle is that a distribution of income and wealth is to "everyone's

advantage." In subsequent discussion in *A Theory of Justice* Rawls amended this formulation. He notes that the phrase "everyone's advantage" suggests that the expectations of the various income groups are "close-knit"; they go up and down together, and "there's no loose-jointedness, so to speak, in the way expectations hang together." But Rawls immediately adds, it is "quite conceivable" that some groups—(for instance) those least well-off—"are not affected one way or the other by changes in the expectations of the best off although these changes [may] benefit others" while having no effect on the least well-off (Rawls 1971: 80–83/1999: 70–72.)

The failure of close-knitness, in fact or in hypothesis, forces a revision on Rawls's initial formulation of the difference principle. A very general revision would say that every group benefits, or at least none becomes worse off. Accordingly, Rawls's principle of "everyone's advantage" needs to be revised to accord with this more nuanced understanding of mutual benefit. Thus, as some people improve their situations, others should continue to improve, to become better-off, or at least none becomes worse off. "Everyone's advantage" or mutual improvement so understood is an ongoing process.

Immediately after his discussion of close-knitness Rawls advances the claim that the difference principle requires that the distribution of income and wealth be "to the greatest benefit of the least advantaged" (see, for example, Rawls 1971: 83, repeated 302–3/1999: 72, repeated 266–7). Probably this is the best known and the most cited version of the difference principle. This new version introduces a complication. Throughout his writings, Rawls renders the difference principle in two distinct ways (the one "mutual benefit," the other "the greatest benefit of the least advantaged"). In his informal argument—as we actually find it in Rawls 1971/1999: ch. 2 (see section 13 for the heart of the discussion)—Rawls presents his governing concern clearly enough (in his discussion of fair equality of opportunity and its relation to mutual benefit). He also lays out his ultimate aim equally clearly (as the "greatest benefit" of the least well-off group). The problem is that he never gives us the intermediate step (the argument itself) whereby we can get from one of these points to the other.

I will try to show, by developing a constructive argument, that Rawls's governing concerns and the goal he aims at are indeed connected (for he clearly regards them that way). Such an argument takes various elements in Rawls's overall discussion, lays these elements down in a certain order, and achieves thereby a definite end result. This idea of a constructive argument is modeled on what Rawls calls a "procedure of construction" (see Rawls 1993/1996: lecture 3 [and p. 93 for the quoted phrase]). A procedure of construction is more like a constructive proof in mathematics—say, a geometrical proof—in which we get a result by making certain moves (constructing certain figures, making certain assumptions) than it is like a classical syllogism in which certain propositions simply follow, by logical entailment, from ready-made premises.

The constructive account begins by assuming that everyone is equal in terms of income and wealth. This is not a (putative) fact but simply an assumption in Rawls's argument. The next element is mutual benefit. The underlying idea here is that where different jobs and positions are appropriately held by people with different talents and abilities, this distribution of positions, with accompanying differentials in income and costs of training and education (for these costs see Rawls 1971: 102/1999: 87; Rawls 1971: 151) should lead to increased production and efficiency (Rawls 2001: 61) which, in turn, benefits everyone. Where this is so, changes that benefit everyone are acceptable, even though they might make for inequality or even for the increase of inequality (see Rawls 1971: 62/1999: 54–5).

Consider Figure 1. Here LWOG is the less well-off group and MWOG is the more well-off one; the axes for each indicate income levels that rise in the direction of the arrows. The dotted line running at an angle of 45° is the equality line; it represents the points of strict and continuous equality between the two groups. Point O is the origin (a point of equality between LWOG and MWOG), *a* and *b* are two points in the available space below the equality line.

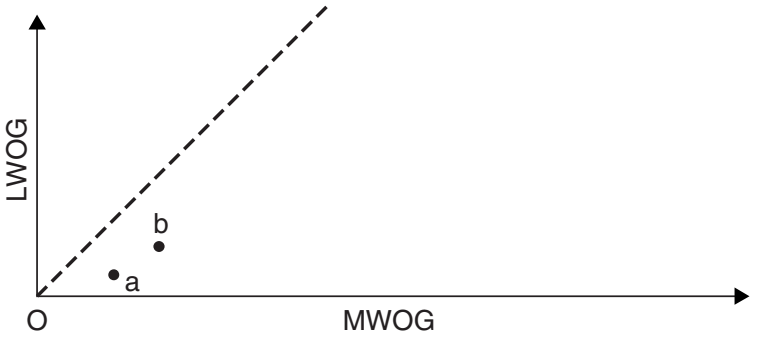


Figure 1

A move from O to *a* is mutually beneficial, although that move introduces inequality between the two groups; a further move from *a* to *b* is likewise mutually beneficial (although it increases the inequality). In both these moves (O to *a* and then from *a* to *b*) we assume that no coercive injustice (such as robbery) has occurred and that the moves are compatible with the two prior principles, i.e., with equal basic rights and liberties and with fair equality of opportunity. So we can conclude that moves from O to *a*, O to *b*, and *a* to *b* are all eligible ones so far as Rawls's theory of justice is concerned.

Let us next add, as an element in the argument, that mutual benefit (understood as making everybody—every income group—better off or at least none worse off) is *a continuing and ongoing process*. This point is worth emphasizing as it allows us to take the idea of mutual benefit to its conclusion.

Next consider Figure 2. This figure, similar to the first, but with the addition of a new curved line, is adapted from Rawls 1971: 76/1999: 66. Rawls calls this curved line a "contribution curve" (Rawls 1971: 76–7, 104/1999: 66–7, 89) or, as he sometimes puts it, a "production curve" (Rawls 2001: 61–2). The area constituted by this curve (the curve *OP*) and the space below it represent the possible positions people could occupy, given natural resources and productive capacity (and with the standard assumption added, that any position on or under this curve is compatible with the prior demands of justice—with the full satisfaction of equal basic rights and liberties and of fair equality of opportunity). Imagine then that one could draw a straight line from O through *a* or from O through *b*. Two such lines have been sketched in. Any position on either of these lines would represent mutual improvement. I have drawn them to suggest an overall trajectory or tendency, in each case, in keeping with the idea that mutual benefit is a continuing and ongoing process.

It should be noted that the lines running through *a* and *b*, respectively, intersect the contribution curve in a portion of the curve which has been darkened. The darkening

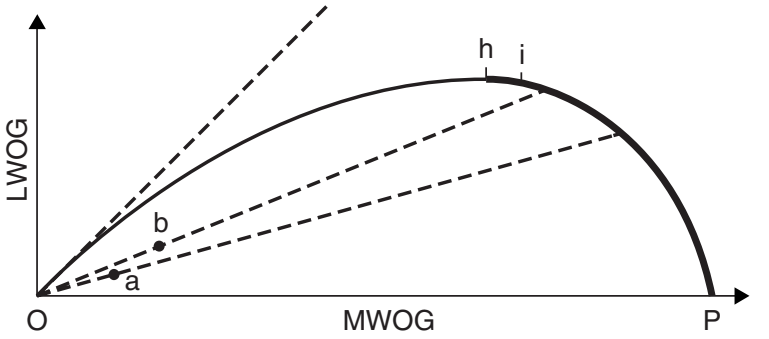


Figure 2

has been introduced to highlight a theme. For example, take points *h* and *i*. Any moves from *h* to *i* would result in LWOG's situation being hurt but MWOG's being improved; any moves from *i* to *h* would have the opposite result. These results for points *h* and *i* could be duplicated for points on *any* stretch of the darkened curve.

Generalizing, we could say that all points (from *h* to *P*) on this darkened curve represent points where any move on the curve, in either one direction or the other, would improve one individual's or one group's situation but only, at the same time, by making another's worse. The points on this curve, in the way I've just described them, are Pareto efficient (or Pareto optimal) points. For this well-known understanding of efficiency (or optimality) see Pareto 1971: 261–9 (especially 261–3, 267) and 451–2; Dyke 1981: 91–5 (especially 95 and n.2 there); Gauthier 1986:76–9 (especially 76); and Rawls 1971: 66–71/1999: 58–62. The idea of Pareto efficiency and of the efficiency frontier (the darkened curve) it defines is the next element in the constructive argument.

We can imagine a whole host of points (other than *a* and *b*) in the space below the curve, each of them representing mutually improving moves away from an initial state of strict equality (that is, away from point O). And straight lines of continuous and ongoing mutual benefit could be drawn from point O through each of these other points; each such line would intersect the efficiency frontier at a different point. We would at this stage have an indefinitely large number of points on the efficiency frontier, each one of them representing the trajectory of a continuing and ongoing process of mutual benefit. But so far as mutual benefit as a continuing and ongoing process is concerned, we have reached (at the intersection of these trajectories with the efficiency frontier) an end point. In short, moves *to* this curve would be mutually beneficial but moves *on* this darkened curve would not and could not be mutually beneficial.

The frontier is the conceptual stopping point for the idea of mutual benefit. But, if we look at the process by which they got there in the first place, every point on the frontier is an eligible point in an account of mutual benefit. Is there any reason to aim for one of these end points in particular? It is important to frame this question in the right way. We are engaged in a constructive *argument*, we are asking if there is a good reason to choose any one of these points as a basic goal point or an ultimate aim. We are not asking anybody to move, on the efficiency frontier, from where they now are to some other point; to do so would violate mutual benefit, the very idea we are working with (and have been presupposing throughout). The moves contemplated here are virtual moves,

moves from point O and in the eligible space below and on the contribution curve (in the limiting case, moves *to* but not *on* the efficiency frontier); they are simply moves in an argument. The question, then, is whether one of these end points should be selected as the best one, selected as the *one* to aim at. Since all are eligible end points in an account of mutual benefit, the choice cannot be made on grounds of mutual benefit.

What about *in*equality? Are we completely uninterested in equality after point O? This brings us to the next element in our constructive argument. Suppose then the following additional thought: just as we should conscientiously try to reduce inequalities in starting points (inequalities stemming from two main sources—initial social circumstance and natural endowment) we should also conscientiously strive to reduce inequalities in income and wealth as regards ending or result points. (I briefly suggested this thought earlier, in the first section.) So, among all the Pareto efficient points (among *all* the Pareto optimal points that serve as endpoints, stopping points, to trajectories of various continuing and ongoing processes of mutual benefit) we pick that point which is closest to the equality line. Point *h* is on the *darkened* curve; point *h* is the optimal point closest to the equality line. I think the best argument for choosing this particular point, as a goal or basic aim, is that it is the one point that allows full scope for mutual benefit at the least cost in terms of inequality. (And here I reply to points raised by Williams (1995: sect. VI).)

Further, I think this particular argument could serve as a bridge between the simple idea of "everybody's advantage" or of mutual benefit and the stronger notion of "the greatest benefit of the least advantaged." Two points should be made here. First, Rawls himself says, at several points, that his preferred point (his difference principle goal point) is the optimal point closest to the equality line, point *h* in Figure 2. (See Rawls "Reply to Alexander and Musgrave" (1974) in Freeman ed. 1999: 247; also Rawls "Social Unity and Primary Goods" (1982) in Freeman ed. 1999: 374 n.12; and Rawls 2001: 62, 68, 123.) Second, as Rawls also notes, point *h* in Figure 2 is the point at which the least advantaged, the least well-off, achieve their greatest benefit (in terms of income and wealth).

In the Figure 2 example, I have followed Rawls's lead in considering a two-group or two-person model case (though sometimes he uses a three-group model). Rawls asks whether the result reached for the two-group case (that is, to choose the optimal point which most reduces the inequality between the group that is below the median in income and the group that is above) can be generalized to an *n*-group case (where *n* designates three or more groups). "I do not know," he writes, "whether the focal point [i.e., the efficient point closest to equality] can be defined sufficiently clearly … when there are three or more relevant classes. The problem is that we need a measure of equality" (Rawls "Reply to Alexander and Musgrave" (1974) in Freeman ed. 1999: 247 and n.7). Though Rawls raises, he does not answer, this question.

I turn to Rawls's question now. I have been arguing for a particular way of formulating and justifying the difference principle. I have done so, first, by bringing Pareto efficiency into play (with Pareto efficiency understood as an efficiency frontier constituted by endpoints, stopping points, to trajectories of various continuing and ongoing processes of mutual benefit). And then, second, I have suggested bringing an egalitarian constraint on Pareto efficiency into consideration (where the egalitarian constraint specifies in effect that the inequality between the top group and the bottom group is to be minimized).

This formulation of the difference principle can be shown to be equivalent to Rawls's standard formulation of it as specifying the greatest benefit of the least advantaged, as the

goal or basic aim. More precisely, in the case where society can be classified into two economic classes, the Pareto efficient-egalitarian version of the difference principle can be shown to be exactly equivalent to Rawls's standard version of the principle. In the general case where society can be classified into n economic classes (where $n > 2$), the equivalence of the two versions can be shown assuming a mild regularity condition, called chain connection. (For a proof of this contention, see the Appendix by Prakash Shenoy in my book (Martin 1985: 197–201).) Chain connection means, roughly, that "if an advantage has the effect of raising the expectations of the lowest position, it raises the expectations of all positions in between" (Rawls 1971: 80/1999: 69). That is, so long as the "curve" of the least well-off class shows improvement, all the other classes do so as well. (For his main discussion of chain connection see Rawls 1971: 81–3/1999: 70–72.)

This analysis has a singular virtue; it explains a puzzling fact about Rawls's exposition of the second principle, a fact that holds throughout his writings. Rawls seems to shift repeatedly between talking about the difference principle as *improving* everybody's prospects and, alternatively, as *maximizing* the benefits of the least well-off group. But the puzzle dissolves once we see that both these versions belong to one and the same story, to one and the same basic argument. Each formulation specifies a single feature of the overall argument; one formulation emphasizes the idea that drives the difference principle—the notion of continual mutual benefit (and the underlying idea of society as cooperative)—the other emphasizes the optimal result, the goal or end point of the entire argument. That goal point can be stated in either of two distinctive ways: (i) as absolutely minimizing the difference (measured in terms of income or wealth) between the top-most and bottom-most group, consistent with the realization of everyone's continual betterment, or (ii) as achieving "the greatest benefit of the least advantaged," that is, the greatest benefit for the least well-off group. Here we have a distinction without a difference; the two formulations of the goals, (i) and (ii), say the same thing.

3. The Difference Principle: Its Justification in the Original Position

Rawls's contractarian method of justification (in what he calls the "original position") is very complex. The original position constitutes an arena for deliberation and decision about principles of justice; its various features are meant to frame and constrain the debate about such principles. "The idea of the original position is to set up a fair procedure so that any principles agreed to will be just" (Rawls 1971: 136/1999: 118). One feature that Rawls has continued to emphasize is that the "parties" to the contract are placed behind a thick veil of ignorance. Here they are constrained, in their subsequent reasoning, to ignore their own particular traits (traits that distinguish them from most or, at least, many other people), to be unaware of (or to ignore) their actual place in society, and so on. The point of the metaphor of the veil is to indicate that the parties should remove sources of bias and irrelevance from their deliberations. Other features are important as well. The parties understand that they are deciding about principles of justice (principles for distributing, by way of the basic structure, certain primary social goods—goods such as liberties, opportunities, income, and wealth—to individuals) and that they will have to live, for their entire lives, under the principles they have selected. Such principles, when looked at from a variety of perspectives, ought to be acceptable to persons in each of those perspectives—this Rawls calls the "unanimity requirement." (Rawls's main discussion of the original position is found in *A Theory of Justice*, chapter 3, and a very helpful summary of its main features is found in Rawls 1971: 146–7/1999: 126–7.)

Rawls envisions, in the *interpretation* presented here, two main roles for the original position: *screening* and *ranking*. In its first role the original position is to serve as a screening device for the candidate principles, that is, principles taken from a short list of main, historically available theories of justice—such as Plato's republic, various versions of utilitarianism, and so on. Here the features of the original position serve as a checklist against which the candidate principles are to be measured and assessed. Thus, for example, in Rawls's view, moral egoism (in any of its several forms) would be *screened* out, dropped from any further consideration, by features of the original position (see Rawls 1971: 132–6/1999: 114–17). But other principles, the various versions of utilitarianism, for example, might pass through the screening tests and remain in contention. Such principles can be formulated and argued for under the constraints of the original position. How would Rawls's second principle do under the screening test in the original position? A certain amount of role-playing is allowed there; individuals can assume the various standpoints on offer and then consider how things would play out in the deliberations of the parties. One could assume that one was in the least well-off economic class.

Where one took strict equality in starting points as a benchmark and remained behind the veil of ignorance, no one would prefer *disadvantageous* deviations, were they (by hypothesis) on the losing end, and hence they would veto such deviations (for example, disadvantageous deviations from strict equality). Only deviations advantageous to all would survive the screening process—that is, only such *advantageous* deviations could achieve the unanimity required of conclusive deliberations in the original position construct (see Rawls, "Basic Structure as Subject" (1978) in Rawls 1993/1996: 281–5). But it would be rational, in the eyes of each, to allow for mutually *beneficial* changes (as in the income example, in Rawls 2001: 62–3), where there were more benefits for each and all, or at least for some, and no losses. In short, the idea of mutual benefit, understood as just described, would—given the high degree of uncertainty, behind the veil, as regards starting points and outcomes for any particular individual—be endorsed in the original position.

Next, let us consider the argument for equality as a constraint on mutually beneficial outcomes. Where persons have an equal status (as parties to the deliberation) and each has equal claim to shares of primary goods, then the parties (as representing such persons in the original position) would prefer a mutually beneficial outcome that *reduced* the difference in income between the top-most and bottom-most group over one that *increased* that inequality. The idea is that, even after mutual benefit (understood as greater benefits for each and all, or at least for some, and no losses) is assured, one should continue to use egalitarianism as a standing constraint on such beneficial options, functioning often as a tie-breaker. Here, among available options, that efficient and mutually beneficial outcome which reduces inequality (or, ideally, minimizes it) is to be preferred. (The idea of using an egalitarian constraint on efficient and mutually beneficial outcomes is found in Rawls's writings at a number of points. See Rawls, "Distributive Justice: Some Addenda" (1968) in Freeman ed. 1999: 164–6; Rawls 1971: 69, 76–8 and sect. 17 entitled "The Tendency to Equality" (esp. pp. 104–5)/1999: 60, 65–7 and sect. 17 (esp. pp. 89–90); Rawls, "Some Reasons for the Maximin Criterion" (1974) in Freeman ed. 1999: 230–1; Rawls, "Kantian Conception of Equality" (1975) in Freeman ed. 1999: 262–6; Rawls, "Basic Structure as Subject" (1978) in Rawls 1993/1996: 281–5; Rawls 2001: sect. 39.)

I think Rawls's straightforward arguments for each of his two principles would fare well in the original position construct. These arguments could be formulated and would hold up, under the constraints identified there. They would pass the screening test of the original position.

Now, let us turn to the second main function of the original position construct, the *ranking* of the competing candidate principles that remain eligible after screening. (See Rawls 1971: 17/1999: 16, and also his "Basic Structure as Subject" (1978) in Rawls 1993/1996: 278, for the point about ranking.) The main ranking argument (called by Rawls the "maximin" argument) presupposes and builds on the arguments permitted and the assessments reached in the screening process. Maximin allows the parties to rank the characteristic or allowed-for outcomes generated by the main competing principles remaining under review—justice as fairness (Rawls's own theory), on the one hand, and its strongest competitor, the principle of maximizing average utility, on the other. Maximin recommends that a representative individual in a situation of deep uncertainty should select the *best* option from within the available set of *worst* outcomes generated by the competing principles (hence, the name "maximin"). Rawls's own gloss on this requirement is that one is to look at the "guaranteeable level" of each of the various worst outcomes on offer under the competing principles and to choose the best outcome among those (if any) that are "satisfactory" and to avoid absolutely all intolerable outcomes. (See Rawls 2001: 98; and also Rawls 1971: 154–5/1999: 134.)

Rawls's view is that utilitarian principles would allow the sacrifice or the serious weakening of some of the demands of justice as fairness, or would do so for some people at least. Here the argument focuses, in particular, on the loss of equal basic liberties of the sort enshrined in the *first* principle (see Rawls 1971: 156, 158/1999: 135, 137, and see Rawls 2001: 100 and also sect. 29). Rawls thinks the characteristic or allowed-for outcomes of the two principles would be satisfactory, but the utilitarian principle (at the point just specified) would yield an intolerable outcome. It is this fact that marks the primary ground, in Rawls's view, for *preferring* the principles of justice as fairness over their presumed closest competitor. Thus the maximin test provides a compelling reason (in the context of the original position) for ranking the two principles, *as a set*, above the principle of average utility.

Rawls came to see that this line of argument engendered no real case for the second principle per se, certainly not for the difference principle. Thus he later suggested that a two-stage original position argument, based on two distinct comparisons, was needed to supplement or replace the initial original position argument of chapter 3 in *Theory of Justice*. (See Rawls 1999: preface, at p. xiv.) Rawls developed this two-stage argument in *Justice as Fairness* (2001: part III; see also Rawls 1971/1999: sect. 49, where such an argument is first suggested). The first stage (the "first fundamental comparison") is precisely like the one we just considered in the previous paragraph, where the two principles "taken as a unit" were ranked above the principle of average utility, itself functioning as the "sole" alternative principle of justice (Rawls 2001: 96). But the second "fundamental comparison" is considerably different. Here Rawls compares his two principles to what he calls a *mixed* conception of justice. This mixed conception includes both the equal basic liberties and fair equality of opportunity, in each case exactly matching the language of Rawls's two principles. But in place of the difference principle, the mixed conception substitutes the principle of maximizing average utility with respect to such matters as economic and social positions, income and wealth. Finally, Rawls constrains the operation of average utility here by requiring a baseline, in what he calls "a social minimum," below which no person is to be allowed to fall. (See Rawls 2001: sect. 34.)

The mixed conception allows for a comparison of average utility directly with the difference principle. Even more important, the mixed conception (with its standing commitment both to equal basic liberties and fair equality of opportunity) is *impervious* to

the maximin argument that had sunk the principle of maximizing average utility in the "first fundamental comparison"; and the social minimum baseline makes it impossible for any person to fall (under average utility) to an *intolerable* economic level. Because it has these precise features, the mixed theory forces us to attend simply to the arguments that would make the difference principle per se (and not the difference principle as part of a package of principles) superior to the principle of maximizing average utility as a way of distributing offices and the income and wealth they generate. The crucial line Rawls now takes is to argue, first, that the difference principle cannot be equated with the idea of a social minimum (and here he responds to points made in Waldron 1986) and to argue, second, that the difference principle sets a *higher* standard of well-being for the least well-off group (say, the bottom 20 percent of wage earners) than does the social minimum.

But how it sets that higher standard is the heart of the matter. The difference principle is driven, not by any idea of a minimum or floor, but by the idea of *reciprocity*. In *Political Liberalism*, Rawls distinguishes "reciprocity" from "mutual advantage" (see Rawls 1993/1996: 16–18 (esp. n.18), 50, 54). Mutual benefit takes as its benchmark "each person's present or expected future situation as things are [now]"; reciprocity (which Rawls prefers) takes as its fundamental point of comparison that "everyone benefits judged with respect to an appropriate benchmark of equality defined with respect to that world" (Rawls 1993/1996: 17). The social minimum will prevent immiseration; but it will not give poorer citizens the sense that they really count, that they are a "part" of a truly reciprocal social order and withal "full members" of the public world. A continuing and conscientious deployment of measures for equal opportunity and the operation of an economy in compliance with the difference principle would. Thus it affords such citizens a higher standard of material well-being than would the social minimum and does so on principles of equal citizenship and of reciprocity, that is, of continual mutual benefit as constrained by egalitarianism (see Rawls 2001: sect. 36 and pp. 129–30 for the main argument here). Rawls alleges two other main considerations favoring the difference principle over average utility and its social minimum: publicity (Rawls 2001: sect. 35) and stability (Rawls 2001: sect. 37). A basic structure founded, at appropriate points, on the values enshrined in the difference principle—the values of reciprocity and equal citizenship—would prove stable.

The main use of the maximin argument in Rawls's hands is to rank his preferred set of principles of justice over utilitarian alternatives (in what he calls the First Fundamental Comparison). That argument is not used to defend the difference principle as such. Indeed, maximin does not figure in the Second Fundamental Comparison at all (Rawls 2001: 95–6). The idea that the difference principle is justified (let alone *specifically* justified) by maximin reasoning is, though widely held, mistaken. The standard line of interpreting Rawls's difference principle (at what I have called its goal or optimum outcome) has been that it was justified and rationalized by maximin reasoning. This maximin line of reasoning was attributed to Rawls by his initial expositors who often, then, severely criticized him for using it (see Arrow 1973 and Harsanyi 1975). More recent critics (such as Kymlicka 2002: 62–7, and Wolff 2006: 152–71) have continued to view maximin as the standard pattern of justification intended by Rawls for the difference principle. Maximin, however, adds nothing to the main justifying argument for the difference principle, as found in both *Theory of Justice* (1971/1999) and *Justice as Fairness* (2001), that I recounted in the earlier sections of this chapter.

Related Topics

Contractarianism, Utilitarianism and Consequentialism, Social Choice Theory

References

Arrow, K. J. (1973) "Some Ordinalist-Utilitarian Notes on Rawls' Theory of Justice," *Journal of Philosophy* 70: 245–63.

Dyke, Charles (1981) *Philosophy of Economics*, Englewood Cliffs, NJ: Prentice-Hall.

Freeman, Samuel, ed. (1999) *John Rawls: Collected Papers*, Cambridge, MA: Harvard University Press.

Gauthier, David (1986) *Morals By Agreement*, Oxford: Oxford University Press.

Harsanyi, J. C. (1975) "Can the Maximin Principle Serve as a Basis for Morality? A Critique of John Rawls's Theory," *American Political Science Review* 69: 594–606.

Kymlicka, Will (2002) *Contemporary Political Philosophy* (2nd edn), Oxford: Oxford University Press.

Martin, Rex (1985) *Rawls and Rights*, Lawrence: University Press of Kansas.

Pareto, Vilfredo (1971) *Manual of Political Economy*, translated by Ann Schwier, New York: Augustus Kelly Publishers. (The *Manual* was originally published, in Italian, in 1906; it was first translated into French in 1909; the first English-language version of 1971 was translated from the French edition of 1927.)

Rawls, John (1971) *A Theory of Justice* (1st edn), Cambridge, MA: Harvard University Press.

Rawls, John (1993/1996) *Political Liberalism*, New York: Columbia University Press.

Rawls, John (1999) *A Theory of Justice* (revised edition, with textual revisions and a new preface), Cambridge, MA: Harvard University Press.

Rawls, John (2001) *Justice as Fairness: A Restatement*, edited by Erin Kelly, Cambridge, MA: Harvard University Press.

Waldron, Jeremy (1986) "John Rawls and the Social Minimum," *Journal of Applied Philosophy* 3: 21–33.

Williams, Andrew (1995) "The Revisionist Difference Principle," *Canadian Journal of Philosophy* 25: 257–81.

Wolff, Jonathan (2006) *An Introduction to Political Philosophy* (revised edition), Oxford: Oxford University Press.

Further Reading

Brian Barry, *Theories of Justice* (Berkeley and Los Angeles: University of California Press, 1989) offers a profound criticism and reworking, in chapters 5 and 6, of Rawls's theory of distributive justice (chapter 6, in particular); the book also sharply contrasts mutual advantage with Barry's own preferred idea, that of impartiality. Matthew Clayton and Andrew Williams, eds., *The Ideal of Equality* (London: Macmillan; New York: St. Martins, 2000) contains thoughtful essays by a number of leading thinkers on equality and the grounds for deviating from equality; the idea of prioritarianism (which gives special weight to the welfare of the least well-off in determining the justice of such deviations) is developed in a number of essays (see especially the introduction by the editors, at pp. 5–8, and the discussion of Rawls by Derek Parfit on pp. 116–21). Gerald A. Cohen, *Rescuing Justice and Equality* (Cambridge, MA: Harvard University Press, 2008) offers a deep and bracingly written critique of Rawls, especially of Rawls's willingness to deviate from equality in outcomes and of his use of incentives, including differential incomes, to spur productivity by those well-off. Samuel Freeman, *Rawls* (Abingdon and New York: Routledge, 2007) provides historical background and a comprehensive survey of some of Rawls's main themes, with a useful chapter devoted to equal opportunity and the difference principle and another to the original position. Thomas Pogge, *John Rawls: His Life and Theory of Justice*, translated by Michelle Kosch (New York: Oxford University Press, 2007) is an updated and reworked version of the German original (Munich: Beck, 1994); it develops a very incisive and well-written critique of Rawls's domestic theory of justice.

36
LEFT LIBERTARIANISM

Hillel Steiner

1. Introduction

Uncontroversially, there are three standard measures of a normative theory's power. One is the extent to which its basic principles, when taken in conjunction with empirical information, can explain or justify widely affirmed normative intuitions and judgments. The second, as with any theory, is the exiguity of basic principles the theory contains—the fewer the better, so to speak—inasmuch as the more there are, the nearer the theory approaches being merely redescriptive rather than explanatory. And the third is coherence: if its set of norms yield contradictory judgments about the permissibility of a particular action, then the theory either is untrue or (what comes to the same thing) must be modified to be true.

Theories of distributive justice generate sets of moral rights. Left libertarianism claims to generate a coherent set of moral rights from just two basic principles:

> Left-libertarian theories of justice hold that agents are full self-owners and that natural resources are owned in some egalitarian manner. Unlike most versions of egalitarianism, left-libertarianism endorses full self-ownership, and thus places specific limits on what others may do to one's person without one's permission. Unlike the more familiar right-libertarianism (which also endorses full self-ownership), it holds that natural resources—resources which are not the results of anyone's choices and which are necessary for any form of activity—may be privately appropriated only with the permission of, or with a significant payment to, the members of society.
>
> (Vallentyne 2000: 1)

In fact, left libertarianism encompasses a number of different theories, some of which are far from recent (cf. Vallentyne and Steiner 2000a, b). What they all share in common is the view that sets of rights derived from successive exercises of those two basic entitlements—the right of self-ownership and the right to natural resources—are sufficient to embody the requirement of *responsibility-sensitivity* and, thereby, to account for many of our most significant pre-theoretical intuitions about the just interpersonal distribution of goods and services. Their being responsibility-sensitive associates them with the larger family of liberal theories that has come to be known as *luck egalitarianism* (cf. Vallentyne *et al.* 2005; Vallentyne 2002), though left libertarians would claim that other luck egalitarian theories—as well as right libertarian ones—are comparatively less systematic in satisfying that requirement: their normative theorems either partially

deviate from it or fail to yield determinate solutions in particular cases. Responsibility-sensitive theories are ones that capture that diverse set of common intuitions that all support entitling persons to—and encumbering them with—those things which are the consequences of their own respective choices, but not those things which are the consequences of their respective circumstances. A pressing issue, then, is how a responsibility-sensitive theory can assign rights to natural resources, since such resources are not the consequences of any person's choices. This issue will be addressed presently. Before doing so, we need to review several relevant general facts about the nature of rights.

2. Rights

Historically, arguments about what moral rights persons are vested with have relied upon one or the other of two rival conceptual models of rights (cf. Sumner 1987; Kramer et al. 1998). According to the Interest Theory, to be in possession of a claim-right is to be the person who has a significant interest in the performance of the duty correlatively entailed by that right. For the Will Theory, on the other hand, such a right is vested in that person who has the authority or powers to control that duty: that is, to demand (and, if necessary, secure enforcement of) performance of that duty or, alternatively, to waive it. Thus I have a Will Theory right against being killed by you if, and only if, my consent to your killing me is necessary and sufficient to waive your duty not to kill me. But if my consent cannot waive that duty—as is legally the case in jurisdictions that outlaw attempted suicide and/or voluntary euthanasia—then the right against my being killed either belongs to some third party (typically, the state) or it belongs to me but only as an Interest Theory right: a right whose correlative duty is vested in me, as well as others.

Libertarians, whether left or right, evidently deploy the Will Theory model of rights (Vallentyne 2007, however, explores a hybrid choice-prioritizing conception), and they regard as unjust any forcible interference with my decisions as to what may be done with what justly belongs to me, so long as my actions do not violate the rights of others. Thus, for example, my self-ownership vests me with the authority to enlist the voluntary aid of another in ending my life. The legal denial of such authority is seen as unjust and, at best, an instance of state paternalism: the state is acting as if my body (i.e. life) is its property rather than mine.

One of the strengths of the Will Theory is that rights that are modeled on that conception—as distinct spheres of personal discretion—are ones that can form *compossible* sets. That is, unlike Interest Theory rights which cannot be modeled so as to preclude interpersonal conflicts of the interests they ostensibly protect, Will Theory rights can constitute a set of rights which cannot conflict with one another—rights whose respective correlatively entailed duties are jointly performable (cf. Steiner 1994: 92; 1998: 290–3; Waldron 1989). The compossibility of all the rights in a set of rights is, indeed, the principal test of whether the theory of justice generating that set satisfies the aforementioned coherence condition.

Libertarian theories seek to satisfy the compossibility condition by construing all rights as reducible to property rights. Property rights partition the physical elements of what we may broadly call *action-space*—spatio-temporal locations and material objects—by severally assigning discrete portions of that space to different persons as constituting their respective domains. Thus, if two persons' chosen actions conflict—if the occurrence of one would preclude that of the other—which one of them is justly permissible

is to be determined *not* by reference to their comparative worth (about which those persons presumably disagree), but rather by reference to whose assigned domain contains the physical components of those jointly unperformable actions (cf. Nozick 1974: 238; Miller 1982: 275; Steiner 1994: 86–101; 1998: 262–74; 2006: 468–70). A well partitioned set of rights, by minimizing the possibility of such conflicts, greatly restricts the authority of the state inasmuch as it thereby minimizes the occasions for conflict-resolution by exercises of third-party (legislative or judicial) discretion, and can thus be seen as a necessary condition for *the rule of law*.

Relatedly, the duties entailed by our moral rights are typically understood by libertarians, and by liberals generally, as forming only a subset of all our moral duties, albeit one which enjoys primacy among them. It is these two features of moral rights that explain liberalism's traditionally avowed compatibility with a plurality of conceptions of the good life—its aforementioned neutrality on the comparative worth of conflicting actions—as well as its capacity to sustain the intelligibility of the common notion of *a right to do wrong* (cf. Otsuka 2003: 27–8). Thus, one of morality's primary rules or values may well be *charity*—a norm which encumbers me with duties to transfer some of my resources to persons considerably more in need of them than I am. Assuming that I am justly entitled to those resources—that I hold moral rights that others not interfere with my disposition of them—this entitlement does *not* entail that I do no wrong in refusing to act charitably and insist on withholding those resources from needier persons. All that is entailed by assigning primacy to moral rights, is that others would be committing a *worse* wrong by forcing me to make that transfer. In other words, morality's assigning such primacy to rights entails that the following three alternatives are listed in descending order of desirability: (a) my choosing to transfer my resources to the needy; (b) my withholding those resources; (c) my attempting to withhold those resources but being forced by others to transfer them. It is alternative (b) that represents having (i.e. exercising) a right to do wrong. The fact that my withholding is an exercise of my rights is insufficient morally to justify that act. All that it would suffice to justify are whatever actions might be necessary to prevent or redress my being forced to transfer (cf. Steiner 1996; 2006: 466–8).

3. Foundational Rights

Any account of rights that are responsibility-sensitive must distinguish between foundational rights and derivative ones. A right is foundational—non-derivative—if (i) it is one that can be exercised to create other rights, and (ii) it is not one that has been created by the exercise of another right. Foundational rights are, in that sense, natural rights: they are to be contrasted with derivative rights inasmuch as, although these too are moral rights, foundational rights are not the results of choices. All libertarians agree that the right of self-ownership is a foundational moral right. Some find it implicit in a foundational right against coercion, while others see it as implied by a foundational right to equal liberty.

Most right libertarians hold that self-ownership is our *only* foundational right. They maintain that it is through persons' various chosen exercises of this right—through Lockean *labor-mixing* or some equivalent process—that moral property rights in originally unowned natural resources (land, space, etc.) come to be established and thence to serve, along with self-ownership, as the derived rights from which all other moral rights are serially derived.

Left libertarians disagree. For them, rights to originally unowned natural resources

are not derivative ones but, rather, are as foundational as self-ownership. They contend that sets of rights derived solely from exercises of self-ownership rights and their serial derivatives both *compromise* responsibility-sensitivity and contain rights that are *incompossible*. These deficiencies are due to the fact that property distributions underwritten by those right libertarian theories are sensitive not only to persons' choices, but also to their comparative temporal locations. For such theories assign greater Hohfeldian powers and privileges to members of temporally prior generations than to their immediate successors. More specifically, in setting no limits on natural resource appropriation, they licence those earlier persons to acquire more unencumbered property rights in initially unowned natural resources than would, in Locke's familiar phrase, "leave enough and as good" for those others (cf. Otsuka 2003: 35–8). Since persons are not responsible for (the starting-point of) their own temporal locations, and since being vested with such proprietary control of natural resources endows those who have it with (at least) superior bargaining power, such unequal distributions as result from exercises of those unequal bargaining powers are far from being describable as responsibility-sensitive.

More strongly, many left libertarians have argued that any such theory is *strictly incoherent*. It is incoherent because, by thus empowering a subset of self-owning persons unilaterally to acquire unencumbered ownership of *all* natural resources, it implies that, in the absence of those owners' permission, later arrivals are encumbered with *unperformable duties of non-trespass*, the enforcement of which unavoidably violates those later arrivals' self-ownership rights. So the property rights of those resource owners are thereby *incompossible* with the later arrivals' self-ownership rights. That is, such theories can imply, of one and the same such enforcement action, that it is both *permissible* and *impermissible* (cf. Steiner 2009: 241). And in so doing, they fail to satisfy the aforementioned coherence requirement.

Accordingly, in a world where all natural resources have been appropriated by private individuals (or nations), left libertarians insist that the ownership of any natural resource is encumbered by a liability to compensate all others for their exclusion from the use of that resource. Initially unowned, natural resources are ones that everyone is at liberty to use. By establishing or retaining exclusive private ownership of such a resource, its owner coercively denies to others a liberty which they would otherwise have, since that ownership title entails that all other persons are subject to an enforceable duty to refrain from using it. The magnitude of that aforesaid compensatory liability is equal to the current market value of the natural resource appropriated. Broadly speaking, this is equivalent to the difference between that resource's gross market value and the market value of whatever improvements have been made to it by human labor, and each person is entitled to a portion of that difference (cf. Steiner 1994: 268–73; 2011: 121–3). It is that compensatory liability that "levels the playing field" between members of temporally overlapping generations and, in that sense, contributes to rendering the distribution of property rights among them responsibility-sensitive—in much the same sense that a necessary condition for the array of finishing positions in a running-race to be responsibility-sensitive is that the runners each started at the same time and place (cf. Steiner 1994: 224–8; 1997).

4. Scope of Libertarian Rights

For all libertarians, and for liberals generally, the primary subjects of rights are individual persons. Any group rights—such as those of corporations or nations—are derivative

ones, created and sustained by exercises of their individual members' rights. Hence the powers and liberties of such groups—and of persons authorized to act on their behalf—are restricted to those which have been authentically ceded to them by each of those members (cf. Locke 1690: esp. §§II.95, 118, 135; Otsuka 2003: chs. 5–6). Nations are no more empowered than their individual members to violate either the unceded foundational rights of their members or the rights derived from them. And the same is true with respect to the rights of non-members.

These restrictions have clear implications for significant aspects of international relations. All libertarians agree that neither nations nor individuals may initiate violence against groups or persons who are innocent: that is, who have violated no one's libertarian rights. They similarly agree that the state's moral authority extends neither to preventing emigration, nor to restricting immigration when at least some of their members are willing to employ and/or accommodate foreigners. And this same lack of regulative authority applies to international trade and investment. Moreover, both right and left libertarians support a right of *secession*. That is, landowners, unless they themselves have contracted otherwise, are empowered to withdraw their land from the state's jurisdiction (cf. Steiner 1994: 262–5; Otsuka 2003: 100).

However, where right and left libertarians are bound to differ is in the matter of obligatory wealth transfers between nations: "bound to differ," because of the left libertarian foundational right with regard to natural resources. National territorial jurisdictions are simply aggregations of various portions of appropriated land. The left libertarian requirement that landowners compensate *all* other persons for their exclusion from those sites is to be understood literally: that liability, like the duty not to initiate violence against innocent persons, is not one that somehow dissolves at the borders of one's own society and, hence, is not one that is owed merely to one's compatriots. Accordingly, the aggregate amount of compensation thereby owed can be conceived as forming a *global fund*, on which each person—wherever located—has a claim. In effect, left libertarianism attaches a price tag to nations' territorial acquisitions and retentions. And it seems reasonable to suppose that such a mechanism might well make a contribution to the peaceful resolution of territorial disputes between nations (cf. Steiner 1994: 265, 268–70; 2005).

International wealth transfers are not the only issue that appears to divide left from right libertarians, in regard to the scope of the rights they endorse. Since the Will Theory model of rights—which most libertarians embrace—vests rights in those persons who control the duties correlative to them, rights cannot be vested in those who are necessarily incapable of exercising such control. Obviously prominent among the latter are non-existent persons: that is, persons who no longer exist and (most) persons who do not yet exist. The dead have no rights and, consequently, any property that they own at the moment of their deaths counts as unowned immediately afterward. As such, it falls into the same normative category as initially unowned natural resources, and is thus subject to the same egalitarian distributive rule. In practice, what this would amount to is a 100 percent tax on decedents' estates (with, perhaps, designated heirs being given a right of first refusal on the components of those estates). The proceeds of that tax join natural resource owners' compensatory payments in the *global fund* (cf. Steiner 1994: 250–8, 273).

Correspondingly, members of most future generations—most persons who do not yet exist—are similarly incapable of holding Will Theory rights. Hence, for example, present persons do not owe rights-based duties of environmental conservation and/or other forms of saving to those future persons. This does not, of course, entail that present

persons are devoid of *any* conservation duties. For, first, they might have non-rights-based moral duties to conserve. And second, they might owe rights-based conservation duties to members of the next one or two generations, since those persons are likely to share some element of contemporaneity with present persons, and thus be able to demand redress from the latter for any failure to furnish them with (the value-equivalent of) a just share of natural resources (cf. Steiner 1994: 259–61; 2009: 241–2).

It does need to be noted, however, that some right libertarians concur in denying a right of bequest, and some left libertarians affirm the presence of a right to conservation in members of *all* future generations (cf. Steiner and Vallentyne 2009).

5. Free Markets and Exploitation

All libertarians favor free markets. They see such markets as unsurpassed not only in their widely acknowledged capacity to allocate goods and services efficiently, but also—and more significantly—in their freedom-enhancing capacity to allow individuals fully to exercise their property rights as they see fit. The primary role of the state in this regard is not to regulate economic activity beyond what is necessary to prevent or redress violations of those property rights, including the enforcement of contractual obligations. That said, however, libertarianism does have the conceptual capacity to recognize some free market transactions as exploitative and, hence, unjust. The fact that the parties entering into a free market exchange do so voluntarily is not regarded—by most left libertarians and at least some right libertarians—as being, in itself, a sufficient reason for holding that transaction's distributive outcome to be just. Why? A simple example will suffice.

Markets are auctions writ large. Suppose that, in an auction for Red's goods or services, Blue's offer of $10 is the winning bid. Further suppose that White, who would have outbid Blue with an offer of $15, had previously had his libertarian rights violated, say, by being barred from the auction or by being robbed. In other words, had that prior rights-violation not occurred, Red's auctioned goods or services would have sold for at least $15. The injustice of that violation thereby taints the $10 price Blue paid and renders his title to what he purchased similarly unjust. Evidently, such injustices, if unredressed, are likely to be strongly compounded over time by further exploitations, since Red's receiving only $10 instead of $15 means that she herself will have less bidding power, than she would justly have had, at any auction she would attend as a purchaser. Since the history of widespread and unredressed violations of libertarian rights—and of the exploitations they have fostered—is not a brief one, its ramifications can continue to exert a significant distributive impact even in societies where unredressed current rights-violations are no longer common. So although free markets *do* have the virtues attributed to them by libertarians, it is difficult to see how their distributive outcomes can be any more just than the sets of rights exercised in them. Left libertarians concur with right libertarians in doubting that state regulation of the economy is a plausible remedy for such injustice. And consequently, they look to wealth redistribution (of the sort previously described), as well as private law restitution, to accomplish at least some of that rectification (cf. Steiner 2010).

6. An Unresolved Issue

Left libertarianism, as was suggested above, is a family of theories: having much in common, they also appear to differ from one another in some respects. Most of these differ-

ences revolve around the question of what, precisely, the foundational right to natural resources is meant to equalize.

On one view of the matter, the relevant *equalisandum* is opportunities for welfare. Primary determinants of persons' opportunities are their health, talents, and the gifts they receive from others. The interpersonal distribution of these factors is, as Rawls claims, arbitrary from the moral point-of-view. And to the extent that this distribution is an unequal one, the requirement that property distributions be responsibility-sensitive entails that the aforementioned *global fund* should be differentially distributed to individuals in such a way as to offset those inequalities.

An alternative left libertarian view sees that foundational right's *equalisandum* as shares of the *global fund* itself. Recall that this fund's revenues are comprised of the value of decedents' estates and the value of natural resource ownerships. On this alternative view, natural resources are held to include the germ-line genetic information that parents use to procreate offspring and that varies in value from one set of parents to another. These variations are due to the unequal capacities of respective information-sets to produce genomes with genetic predispositions to develop into healthy and/or talented offspring. Parents, in procreating, are *appropriating* that germ-line genetic information and, like appropriators of other natural resources, are thereby liable for a payment to the *global fund* that is equal to the value of what they appropriate. Accordingly, each person's right to an equal share of that fund entails, *ceteris paribus*, that there is a net transfer of wealth from those parents who produce more valuable genomes to their counterparts who produce less valuable ones, thereby furnishing the latter with the means to offset their children's genetic disadvantage (cf. Steiner 1994: 237–49, 273–80; 1999; 2008; 2011: 112–15).

It is an open question as to whether the opportunities view of what is to be equalized is more or less responsibility-sensitive than its resources counterpart and, hence, whether those two views are extensionally equivalent to one another.

7. Conclusion

It was noted at the beginning that one measure of a normative theory's power is the extent to which its basic principles, when taken in conjunction with empirical information, can explain or justify widely affirmed normative intuitions and judgments. In the case of distributive justice theories, our most salient intuitions and judgments tend to reflect demands of both equality and responsibility-sensitivity. For right libertarians, the demands of equality are fully satisfied by each person's being vested with a foundational right of self-ownership: no one can, morally, be born into slavery or other forms of servitude. Provided that everyone is so vested, any distribution of property deriving from exercises of that right, and of the rights derived from them, satisfies the demands of responsibility-sensitivity.

Left libertarians deny that such distributions are fully responsibility-sensitive and claim that, to satisfy that requirement as well as the demands of equality, a property distribution must be one derived from *both* a foundational right of self-ownership *and* one to natural resources that equalizes persons' opportunities or their shares of natural resource value. Among the pre-theoretical intuitions and judgments that left libertarianism can thereby explain and justify, two are that persons are entitled to the fruits of their labor, and that voluntary exchanges in free markets can be exploitative. Another is that all persons have, unconditionally, some initial material entitlement that constitutes their

spheres of personal (negative) freedom, spheres where others may not physically interfere with their actions. A fourth is that basic universally vested or *human* rights include some material entitlements. A fifth is that persons should not be disadvantaged by their genetic inheritance. And a sixth is that parents should be held responsible for the consequences of their own procreative and nurturing choices. Finally, in embracing a tax-base that consists solely of natural resource values and those of decedents' estates, left libertarianism avoids what Dworkin and others have called the "slavery of the talented" and thereby allows unencumbered occupational choice: brilliant brain-surgeons can abandon their lucrative jobs to become mediocre poets, if they wish to do so.

Related Topics

Locke, Liberalism, Environmentalism, Natural Law and Rights Theory, Luck Egalitarianism, Libertarianism, Intergenerational Distributive Justice, Human Rights and Cosmopolitanism, Global Justice and Politics, Justice and Borders, Equality, Freedom, Rights, Toleration, Paternalism, Moralism, and Markets

References

Kramer, M., N. Simmonds and H. Steiner (1998) *A Debate Over Rights: Philosophical Enquiries*, Oxford: Oxford University Press.

Locke, J. (1690) *Two Treatises of Government*, London: Awnsham Churchill.

Miller, F. (1982) "The Natural Right to Private Property," in Machan, T. (ed.) *The Libertarian Reader*, Totowa: Rowman & Littlefield.

Nozick, R. (1974) *Anarchy, State and Utopia*, Oxford and Cambridge, MA: Blackwell.

Otsuka, M. (2003) *Libertarianism without Inequality*, Oxford: Oxford University Press.

Steiner, H. (1994) *An Essay on Rights*, Oxford and Cambridge, MA: Blackwell.

—— (1996) "Duty-Free Zones," *Aristotelian Society Proceedings* xcvi: 231–44.

—— (1997) "Choice and Circumstance," *Ratio* X: 296–312.

—— (1998) "Working Rights," in Kramer, M., N. Simmonds and H. Steiner, *A Debate Over Rights: Philosophical Enquiries*, Oxford: Oxford University Press.

—— (1999) "Silver Spoons and Golden Genes: Talent Differentials and Distributive Justice," in Burley, J. (ed.) *The Genetic Revolution and Human Rights: The Oxford Amnesty Lectures 1998*, Oxford: Oxford University Press.

—— (2005) "Territorial Justice and Global Redistribution," in Brighouse, H. and G. Brock (eds.) *The Political Philosophy of Cosmopolitanism*, Cambridge: Cambridge University Press.

—— (2006) "Moral Rights," in Copp, D. (ed.) *Oxford Handbook of Ethical Theory*, Oxford: Oxford University Press.

—— (2008) "Universal Self-Ownership and the Fruits of One's Labour: A Reply to Curchin," *Journal of Political Philosophy* 16: 350–5.

—— (2009) "Responses," in de Wijze, S., M. Kramer and I. Carter (eds.) *Hillel Steiner and the Anatomy of Justice*, New York and London: Routledge, pp. 239–41.

—— (2010) "Exploitation Takes Time," in Vint, J., S. Metcalfe, H. Kurz, N. Salvadori and P. Samuelson (eds.) *Economic Thought and Economic Theory: Essays in Honour of Ian Steedman*, London and New York: Routledge.

—— (2011) "Sharing Mother Nature's Gifts: A Reply to Quong and Miller," *Journal of Political Philosophy* 19: 110–23.

—— and P. Vallentyne (2009) "Libertarian Theories of Intergenerational Justice," in Gosseries, A. and L. Meyer (eds.) *Justice Between Generations*, Oxford: Oxford University Press.

Sumner, L. (1987) *The Moral Foundations of Rights*, Oxford: Oxford University Press.

Vallentyne, P. (2000) "Introduction: Left-Libertarianism—A Primer," in Vallentyne, P. and H. Steiner, *Left-Libertarianism and Its Critics: The Contemporary Debate*, Houndmills: Palgrave.

—— (2002) "Equality, Brute Luck, and Initial Opportunities," *Ethics* 112: 529–57.

—— (2007) "Libertarianism and the State," *Social Philosophy & Policy* 24: 187–205.

—— and H. Steiner (eds.) (2000a) *The Origins of Left-Libertarianism: An Anthology of Historical Writings*, Houndmills: Palgrave.
—— and —— (2000b) *Left-Libertarianism and Its Critics: The Contemporary Debate*, Houndmills: Palgrave.
—— and M. Otsuka (2005) "Why Left-Libertarianism Is Not Incoherent, Indeterminate, or Irrelevant," *Philosophy & Public Affairs* 33: 201–15.
Waldron, J. (1989) "Rights in Conflict," *Ethics* 99: 503–19.

Further Reading

Fried, B. (2004) "Left-Libertariansim: A Review Essay," *Philosophy and Public Affairs* 32: 66–92.
—— (2005) "Left-Libertarianism, Once More: a Rejoinder to Vallentyne, Steiner and Otsuka," *Philosophy & Public Affairs* 33: 216–22.

37
LIBERTARIANISM

John Meadowcroft

> He is a libertarian, not because he begins with an abstract definition of liberty, but because he has actually enjoyed a way of living (and seen others enjoy it) which those who have enjoyed it are accustomed (on account of certain precise characteristics) to call a free way of living, and because he has found it to be good. The purpose of the inquiry is not to define a word, but to detect the secret of what we enjoy, to recognise what is hostile to it, and to discern where and how it may be enjoyed more fully.
>
> (Oakeshott 1949: 387)

In the above passage Michael Oakeshott set out the belief that libertarianism begins from the hedonistic enjoyment of "a free way of living." Liberty is deemed valuable because it is conducive to human happiness, not because it fulfills certain abstract, logical criteria. The philosophical and political project of libertarianism follows from this enjoyment of a free way of living, leading to the desire to better understand it, defend it where necessary and make it more widely available.

This chapter will elucidate this libertarian project with particular reference to its implications for attempts to use the state to achieve social or distributive justice. The chapter will define a free way of living as a framework within which individuals may pursue their own conception of the good life, so that different moral values may coexist within it. It will be shown that this framework can be understood as rights that place side-constraints on the actions people may take in pursuit of their ends. A free way of living is therefore shown to be a process via which multiple experiments in living take place simultaneously as people explore the viable alternative lifestyles and practices. This libertarian perspective will then be applied to the question of whether the state can be legitimately used to redistribute income and wealth in order to achieve or advance social justice. A short, final section will conclude.

Two important caveats should be noted at the outset. First, this chapter draws upon the work of a number of scholars who are commonly identified as classical liberals or conservatives rather than libertarians—Michael Oakeshott is one example. However, the work of these scholars informs the broad libertarian philosophical tradition and therefore reference to their work enables a full appreciation of libertarianism. Second, the chapter synthesizes the work of a number of scholars into a consistent libertarian position, although in reality there are important differences of opinion among the scholars cited herein. Robert Nozick's entitlement theory of justice discussed later in this chapter, for example, has been criticized by other libertarian scholars. However, it is judged that the presentation of a consistent libertarian position is of more value to

the reader than dissection of the differences and disagreements among libertarian, or libertarian-inclined, scholars.

1. Liberty, the Good Life and the Separateness of Persons

Chandran Kukathas (2000, 2003) has argued that the crucial distinction within liberal political theory lies between (1) the view that a liberal society describes a community in which all members subscribe to an agreed set of moral values, such as freedom of speech, religious tolerance and freedom of assembly, and (2) the view that a liberal society describes a set of principles within which different moral values may coexist, so that different groups may pursue lifestyles that reflect different moral values, whether strict Christian observance, libertinism or other norms. In this chapter it is argued that libertarianism, or a free way of living, corresponds with the second of these two views—libertarianism describes a framework within which different moral values may coexist.

The libertarian approach is methodological and ethical individualist. Methodological individualism implies that there is no such thing as "society" other than the sum of the individuals who compose it. It is only individuals who can act, choose or hold beliefs. In the words of the Nobel laureate James M. Buchanan (1989: 37–8):

> Only individuals choose; only individuals act. An understanding of any social interaction process must be based on an analysis of the choice behavior of persons who participate in the process. Results that are predicted or that may be observed in social interaction must be factored down into the separate choices made by individuals.

Methodological individualism, then, can be understood as a research strategy that requires that social phenomena should be understood in terms of the actions, values and beliefs of individuals. It is claimed that societies, governments or firms can only be understood in terms of the actions of the individuals who compose those entities.

Ethical individualism implies that individuals are morally separate entities and no one individual is of greater value than any other. In the words of Buchanan, "Each person, as a separately existent unit of consciousness, looks at, hears, feels, and tastes that which is confronted through a 'different window'" (Buchanan 1999: 11). Although we might not approve of the choices that other people make, it follows that only the individual can know his or her desires, tastes and preferences and therefore we cannot choose for others because we do not know their wishes.

The separateness of persons leads to the belief that there is no single good life that is best for all. In arguably the most important single work of libertarian political philosophy, Robert Nozick's *Anarchy, State, and Utopia*, Nozick listed thirty-six celebrities and historical figures, including the philosopher Ludwig Wittgenstein, the artist Pablo Picasso, the car-marker Henry Ford, the spiritual teacher Baba Ram Dass and the anarchist Peter Kropotkin, plus "you, and your parents," and then challenged the reader: "Is there really *one* kind of life which is best for each of these people?" (Nozick 1974: 310). Nozick argued that the diversity and complexity of humanity meant that there must be many different versions of the good life and no one could say what kind of life was the most valuable or enjoyable for others.

For this reason a free way of living does not describe specific practices or lifestyle choices that everyone is assumed or required to enjoy. Instead, a free way of living

implies a framework within which many different practices and lifestyles can coexist. As Nozick (1974: 310–11) put it:

> There is no reason to think that there is one community which will serve as ideal for all people and much reason to think that there is not … [The idea that there is] one best society for everyone to live in, seems to me to be an incredible one.

It should be stressed that this view of the separateness of persons does not imply that people are akin to atoms that exist in isolation from one another, untouched by other people and the external world. On the contrary, as Buchanan (1999: 11) made clear, it is important to recognize that individuals live "in four-dimensional space-time." People must be understood to live in particular places and at particular times and hence must be located in their social and cultural context. Individual preferences, tastes and desires will be formed in the context of people's interactions with one another in particular social and cultural settings. (See also Hayek 1948.)

This view of the separateness of persons does imply, however, that other people, whether politicians, bureaucrats or philosophers, cannot legitimately make trade-offs between the happiness of different individuals. It is not legitimate to judge that the happiness of one individual may be sacrificed for the happiness of others. Individuals may make personal trade-offs in order to maximize their own happiness, but the same maximizing logic cannot be applied to trade-offs at the societal level. As Nozick (1974: 32–3) described:

> [T]here is no social entity with a good that undergoes some sacrifice for its own good. There are only individual people, different individual people, with their own individual lives. Using one of these people for the benefit of others, uses him and benefits the others. Nothing more. What happens is that something is done to him for the sake of others … To use a person in this way does not sufficiently respect and take account of the fact that he is a separate a person, that his life is the only life he has.

Individuals are not, as Nozick (1974: 33) put it, "resources for others," but are distinct entities with their own ends. In this sense, libertarianism reflects the basic Kantian principle that individuals must be regarded as ends in themselves, not as means to other people's ends.

It follows, then, that human happiness cannot be "maximized" by trading-off the well-being of different individuals, or by sacrificing the utility of one person for the greater good of the collective. Rather, it is contended that human happiness requires that individuals pursue their own personal conceptions of the good life, even if others do not share their view of the value of their choices.

2. The Libertarian Framework of Rights

The notion that a free way of living describes a framework within which different moral values may coexist raises the question of what constitutes that framework. For libertarians, the framework should be understood to describe individual rights. As Nozick (1974: ix) wrote in the very first sentence of *Anarchy, State, and Utopia*, libertarians

believe that "Individuals have rights, and there are things no person or group may do to them (without violating their rights)." Within libertarian political philosophy there are, however, different accounts of what constitutes these rights and also of the source or origins of rights.

Libertarianism may be understood as a proprietarian political philosophy—meaning that all rights are understood to be essentially property rights. Individuals are assumed to have a right of self-ownership which amounts to a property right in themselves. As Jan Narveson (2001: 66) has written, "the libertarian thesis is really the thesis that *a right to our persons as our property is the sole fundamental right there is*." (See also Rothbard 1998: ch. 15.)

Accordingly, it is argued that people may not physically harm or enslave others because such actions constitute a violation of people's property right in their own persons. Furthermore, individuals are assumed to have a right of ownership of their justly acquired property that similarly prohibits other people taking their income and wealth. It is on this basis that libertarians argue that compulsory taxation—said to constitute taking people's property without their consent—violates people's rights.

Indeed, Nozick argued that compulsory taxation was akin to slavery as it involved requiring people to work for ends other than their own:

> Taxation of earnings from labor is on a par with forced labor. Some persons find this claim obviously true: taking the earnings of n hours labor is like taking n hours from the person; it is like forcing the person to work n hours for another's purpose.
>
> (Nozick 1974: 169)

Kukathas (2003) has argued that the right of exit is the most important right of all that serves as the ultimate guarantee of individual freedom. As long as the right of exit is in place then individual liberty is protected:

> The right of exit is, in fact, nothing more or less than the right to repudiate authority. It arises out of what might be called the 'no-right' of any authority to coerce people into becoming or remaining members of a community or association.
>
> (Kukathas 2003: 97)

For Kukathas, it is essential that the individual has the right to exit from the membership of any group, association, political authority or legal jurisdiction.

Libertarian rights are usually understood to be side-constraints. Side-constraints do not impose ends on individuals, but limit what people may do in pursuit of their own ends: "In contrast to incorporating rights into the end state to be achieved, one might place them as side-constraints upon the action to be done" (Nozick 1974: 29). Accordingly, rights are said to act as side-constraints on individual and group action, prohibiting people from doing things that violate the rights of others.

Nozick did not provide an explicit source for rights and for this reason he was criticized for propounding "libertarianism without foundations"—libertarianism constructed upon a premise that is not supported, explained or justified with reasoned argument or empirical evidence (Nagel 1975). Other libertarians, however, have provided explicit accounts of the origins of rights, either based on the idea of natural rights or the theory that rights result from explicit political agreement among persons.

The concept of natural rights is usually traced back to the ideas of John Locke. According to Locke, each individual, by virtue of their existence, is said to possess natural rights to life, liberty, and property: "yet every Man has a *Property* in his own *Person*. This no Body has any Right to but himself. The *Labour* of his Body, and the *Work* of his Hands, we may say, are properly his" (Locke 1690: 287–8). As Eric Mack (2009: 4) has written, the idea of natural rights is not intended to suggest "that persons are protected by some strange metaphysical shell," but instead, "to assert that certain fundamental facts about each person provide *reasons* for others to be circumspect in their treatment of that person," so that they should "avoid treating that person as material which exists for their own uses and purposes."

An alternative account of the origins of rights within the libertarian tradition is the view that rights originate in voluntary agreements among people to respect one another's persons and property. Buchanan (1975: ch. 4) has argued that rights are created when people choose to act as if others have rights. In the state of nature, or anarchy, widespread agreement to respect the rights of others has not yet been reached and therefore rights cannot be said to exist. Agreement to respect the rights of others is part of the process via which people leave the state of nature and enter into political society. Rights can only be said to exist when and where there is general agreement to respect the rights of others and this agreement has been formalized into law.

The libertarian framework, then, describes rights that enable people to pursue their own conceptions of the good life, as long as they do not infringe the rights of others. Rights are perceived to impose negative constraints on individual action—actions that violate other people's rights are prohibited. But rights do not impose positive duties or obligations—rights do not require that people provide resources or goods for others.

3. The Good Life within the Libertarian Framework

It has been argued that libertarianism describes a framework of rights within which individuals pursue their own conceptions of the good life that are likely to reflect different moral values. The questions that then arise concern how these different, and perhaps conflicting, practices and lifestyles are to coexist and how people are to choose between alternative conceptions of the good life in this context.

One model of how people may pursue happiness within a libertarian framework was set out in the final chapter of Nozick's *Anarchy, State, and Utopia*. Here, Nozick asked the reader to imagine the best of all possible worlds—their utopia. Other people may choose to live in this utopia, or to live in an alternative best possible world of their own imagining. In order to be viable a utopia must be able to attract sufficient people. Hence, if one person's utopia requires other people to work for them as slaves, those other people are unlikely to join it and so it will not prove viable. Given that each individual has a right of exit, no person can be forced to live in a "utopia" against his or her wishes. If a community cannot attract sufficient members to be viable, then it will be necessary to reimagine it in a way that might prove more attractive to others. It is anticipated that different individuals and groups will simultaneously form, dissolve and reform a wide range of different communities.

Nozick, then, imagined utopia to be a framework within which different communities engage in experiments in living mediated by the fact that each individual has the right to exit from any community:

> Utopia is a framework of utopias, a place where people are at liberty to join together voluntarily to pursue and attempt to realize their own vision of the good life in the ideal community but where no one can impose his own utopian vision upon others.
>
> (Nozick 1974: 312)

In order for any conception of the good life to be realized the consent of others might be needed and this consent could be legitimately withheld. No individual may impose his or her vision of the good life on others. This means that it is at least conceivable that "if almost everyone wished to live in a communist community," the situation could arise where "there weren't any viable non-communist communities," so that a few individuals who did not wish to live in a communist association would have "no alternative but to conform" to life under communist principles. Such a situation would not, according to Nozick, violate the rights of an individual who had no feasible option but to live in a communist society even though communism was anathema to him: "the others do not force him to conform and his rights are not violated. He has no right that the others cooperate in making his nonconformity feasible" (Nozick 1974: 322). As long as the individual has the right of exit—even if in practice there is no more-desirable community to join—then a person's rights are not infringed if he or she has to live in a community that does not come close to their ideal.

Nozick's model of "a framework for utopia" was intended to be a thought-experiment about the nature of the good life, rather than a detailed blueprint of how a realistic libertarian future might be constructed. The thought-experiment was intended to illuminate the fact that there are many different conceptions of the good life and in order to be realized any conception will require the acquiescence of others which could be legitimately withheld. Nozick highlighted the fact that utopia is not an end-state, but a *process* via which simultaneous attempts are made to realize alternative ways of living that might prove enjoyable.

Nozick's account of the process via which different people may simultaneously pursue happiness in different ways has similarities with Buchanan's theoretical account of the process via which a social contract is agreed. Buchanan (1975: ch. 4) conceptualized a social contract as emerging from a series of bilateral agreements between individuals. Any social contract capable of commanding the unanimous agreement of an entire polity is unlikely to correspond with any one individual's ideal. Rather, it is anticipated that individuals must compromise in order to reach agreement. In the words of Buchanan and Tullock (1962: 315), people will be prepared to compromise in order to agree the content of a social contract, even though it limits their own choices and constrains their own behavior, because "each individual must recognize that, were he to be free to violate convention, others must be similarly free; and, as compared to this chaotic state of affairs, he will rationally choose to accept restrictions on his own behaviour."

In Buchanan's model of the social contract individuals can be expected to accept second- or third-best alternatives as preferable to the absence of political agreement that carries the threat of violent predation given the absence of widespread respect for rights. Similarly, in Nozick's account of the framework for utopia, people can be expected to accept second- or third-best alternatives as the viable communities that come closest to their utopian ideals.

A free way of living, then, describes a framework within which each individual may pursue his or her vision of the good life. Individuals may pursue practices or lifestyles that

impose severe restrictions on personal freedom, may make wholly permissive choices, or may choose somewhere between the two. The framework is intended to allow different moral values to coexist and, crucially, guarantee individual liberty by ensuring that each individual has the right to exit from any group or authority.

4. The Libertarian Critique of Social Justice

Libertarianism provides a critique of the belief that the state should be used to redistribute income and wealth in order to achieve or advance social justice. As methodological individualists, libertarians reject the notion that the justice or otherwise of a distribution of income and wealth can be evaluated in terms of the overall allocation of resources at a given moment in time. The idea of social justice understood in these terms would seem to imply that someone or some group has deliberately decided what income and wealth each individual, family or group should receive and that outcome can then be evaluated as either just or unjust. In the words of Nozick (1974: 149): "Hearing the term 'distribution', most people presume that some thing or mechanism uses some principle or criterion to give out a supply of things," and that "[i]nto this process of distributing shares some error may have crept." The concept of social justice would seem to demand that a deliberate but faulty allocation of resources should be corrected by government intervention.

The error of this analysis, according to libertarians, is that it does not reflect the actions of individual men and women in the creation of wealth and the exchange of goods and services that produce particular distributions of income and wealth. As Nozick (1974: 149) put it, "we are not in the position of children who have been given portions of pie by someone who now makes last minute adjustments to rectify careless cutting."

The resources that would be reallocated in the name of social justice are not manna from heaven that could be divided in any number of ways between any number of people (although if such resources really were manna from heaven then some deliberative process for fairly distributing them would surely be appropriate). Rather, the resources that exist in a market economy are holdings that particular individuals are entitled to possess as a result of their historical actions in the production and exchange of income and wealth. Hence, in the memorable words of Anthony de Jasay (2002: 186), social justice would seem to involve "slicing the cake nobody baked."

Accordingly, the notion of social justice is perceived to be an intellectual error—a misunderstanding of how the world works. Indeed, F. A. Hayek (1976: 62–3) argued that the demand for social justice was an example of anthropomorphism—the attribution of conscious intelligence to self-ordering, spontaneous processes that are, in fact, beyond conscious control.

Hayek (1976, 1988) argued that something akin to social justice might be applicable in the context of families, friendship networks or within firms. These are made-orders within which it might be appropriate to make deliberative judgments about the allocations of resources between individuals. But Hayek argued (1976: 67–70) that the notion of social justice was meaningless in a complex, modern economy because the distribution of income and wealth generated therein was not planned or intended by any individual or group. For Hayek, the concept of justice was inapplicable to the outcomes of a spontaneous process that was not the result of human design.

It is said to follow that a just distribution of income and wealth concerns the historical processes via which a particular allocation has come about, rather than the particular

pattern or end-state of distribution. Hence, Nozick argued that justice in the distribution of income and wealth could be exhaustively covered by the repeated application of three basic principles of justice in acquisition, justice in transfer, and rectification when the first two principles have been transgressed:

> [T]he holdings of a person are just if he is entitled to them by the principles of justice in acquisition and transfer, or by the principle of rectification of injustice (as specified by the first two principles). If each person's holdings are just, then the total set (distribution) of holdings is just.
>
> (Nozick 1974: 153)

These three principles set out Nozick's entitlement theory of justice. According to Nozick, any inequalities of income and wealth that happen to arise as a result of the legitimate acquisition and the legitimate transfer of resources are a matter of complete irrelevance when it comes to determining the justice or otherwise of a given distribution of income and wealth—the only relevant question concerns whether the holdings were legitimately acquired and/or legitimately transferred (Nozick 1974: 150–3).

Nozick did, however, add a proviso to his theory. This proviso was that a "process normally giving rise to a permanent bequeathable property right in a previously unowned thing will not do so if the position of others no longer at liberty to use that thing is thereby worsened" (Nozick 1974: 178). Nozick's proviso was intended to prevent someone appropriating the total supply of a necessity for life, such as all the drinkable water in the world. The proviso implies a narrow definition of "worsen," not including worsening others' position by leaving them less to appropriate, or appropriating materials in order to compete with other producers. The proviso not only limits what can be justly acquired, but also limits what can be legitimately transferred: "If the proviso excludes someone's appropriating all the drinkable water in the world, it also excludes his purchasing it all" (Nozick 1974: 179).

Nozick famously illustrated his entitlement theory of justice with the example of the wealth acquired by the basketball star Wilt Chamberlain in a fictional scenario. Taking as a starting point a distribution of resources that was considered just, which Nozick named D1, Nozick supposed that large numbers of people were willing to pay 25 cents directly to Chamberlain (in addition to the standard ticket price) in order to watch him play basketball. Supposing that one million people paid to watch him play during the course of a season, Chamberlain would acquire additional personal wealth of $250,000. As a result of the voluntary actions of a million basketball fans and Wilt Chamberlain, a new distribution of resources had emerged, which Nozick named D2. In this new distribution one individual—Wilt Chamberlain—has amassed a sizeable personal fortune (Nozick 1974: 160–4).

For Nozick, the new distribution of resources (D2) must be considered just because it meets the criteria of justice in acquisition and justice in transfer. The theoretical starting point of Nozick's example was a distribution of resources that was just (it was left to the reader to imagine what such a distribution might be) and then people voluntarily transferred their justly held resources to Chamberlain in return for a service (watching him play) that they considered more valuable than the 25 cents each paid. In neither case can an injustice be said to have occurred, and therefore Nozick argued that the new distribution that has emerged must logically be considered just:

> If D1 was a just distribution, and people voluntarily moved from it to D2, transferring parts of their shares they were given under D1 (what was it for if not to do something with?), isn't D2 also just? If the people were entitled to dispose of the resources to which they were entitled (under D1), didn't this also include their being entitled to give it to, or exchange it with, Wilt Chamberlain?
>
> (Nozick 1974: 161)

On the face of it, Nozick's Wilt Chamberlain example would seem to decisively show that justice concerns the historical processes via which distributions of income and wealth come about, rather than particular patterns or end-state distributions of resources.

However, as Nagel (1975: 147) has pointed out, the Wilt Chamberlain example appears to rely on the assumption that the entitlements to property that people have in D1 are absolute—that is, people are assumed to have complete possession of everything that they own, so that people may do whatever they wish with their property without being subject to, for example, land-use planning restrictions or compulsory taxation. But, according to Nagel, it is really the case that "absolute entitlement to property is not what would be allocated to people under a partially egalitarian distribution." In a polity committed to the pursuit of social justice, "[p]ossession would confer the kind of qualified entitlement that exists in a system under which taxes and other conditions are arranged to preserve certain features of the distribution" (Nagel 1975: 147; see also Cohen 1995).

Indeed, in probably the most important argument for social justice, John Rawls explicitly stated that "the rights of property" allocated within his theory of justice were not absolute, but were to be limited in accordance with the terms of the social contract agreed by all (Rawls 1999: 245). Hence, if the members of a polity agreed to tax income at 25 percent in order to help the least advantaged members of society, for example, then Chamberlain would assume ownership of only 75 percent of the money paid to him by the spectators, with the remainder being transferred to the least advantaged via the tax and benefit system. If holdings/entitlements/property rights are agreed by all to be less than absolute, then it is not clear that the taxation and redistribution of part of Chamberlain's income in accordance with such an agreement could be considered a rights-violation.

The question that arises, then, is on what basis Chamberlain can be assumed to have consented to the taxation and redistribution of (a portion of) his income and wealth. As Fried (2005) has pointed out, within the context of Nozick's framework for utopia described above it would seem that entitlements to income and wealth may be limited legitimately if a viable alternative community without such limits to property rights to which people might exit does not exist. If Wilt Chamberlain is placed in the position of the individual who must live in a communist community because there is no viable alternative, even though he has a nominal right of exit from that community, then it would seem to follow—on Nozick's own terms—that his income may be redistributed via compulsory taxation in accordance with the rules of the community without his rights being violated.

Within the context of contemporary liberal democracies, however, it is hard to see that the opportunities to exit from political authority satisfy the requirements of a libertarian right of exit. A principal characteristic of majoritarian democracy is the ability of numerical majorities to impose costs on minorities who have no opportunity for escape or redress. In the words of Buchanan and Tullock (1962: 89–90), "the

essence of the collective-choice process under majority voting rules is the fact that the minority of voters are forced to accede to actions which they cannot prevent and for which they cannot claim compensation for damages resulting." Once present tax liabilities have been imposed such liabilities cannot be removed by moving to another jurisdiction.

Furthermore, the frameworks of competing associations set out by Nozick and by Kukathas envisage different communities coexisting within single geographical locations. Kukathas (2003: 167–71), for example, defines a community as "a coming together of individuals" and notes that "few (if any) individuals are locked in a single community which leaves no room for other attachments … In this sense, (almost) all communities are partial communities." By contrast, contemporary democratic polities do not allow different communities to coexist in terms of *inter alia* tax liabilities and personal freedoms. The idea that a US citizen, for example, must accept every decision of the federal government driven by the electorate of 200 million people, or move to another nation-state, falls catastrophically short of the libertarian framework of competing communities set out in this chapter. When applied to contemporary liberal democracies, then, the libertarian argument that redistribution via compulsory taxation constitutes a rights-violation would seem to hold.

It might be objected, however, that Nozick's entitlement theory has limited applicability to contemporary capitalist societies where the justice of the initial acquisition of resources may be challenged. Indeed, Nozick only partially addresses this issue via his account of just acquisition derived from Locke's ideas. Here, Nozick (1974: 174–8) rehearses Locke's (1690) claim that property rights are created in a previously unowned object when a person mixes his or her labor with it, but leaves a series of unanswered questions when it comes to how this theory might be applied in practice. One plausible implication of Nozick's theory would seem to be that rectification should take place where it can be shown that land was unjustly taken, for example from Native Americans by European settlers, but Nozick does not explore such implications in detail.

Even if the main thrust of Nagel's objection was accepted, however, Nozick's argument might still retain much of its force. Unless income is taxed at 100 percent and all actions that may increase or decrease the value of property are prohibited, then even taking into account the operation of a standard social democratic tax and benefit system, Nozick's argument suggests that there will still be continuous deviations from any preferred pattern or ideal distribution of income and wealth and there is no obvious reason to believe that the new distributions that emerge are unjust.

As Mack (2002: 87) has described, the challenge to those who would advocate the maintenance of a particular end-state or pattern of income and wealth in the name of social justice is that:

> Unless people are effectively prohibited from transforming and exchanging their holdings … there will always be many individuals who will find unexpected ways to enhance the value of their holdings by unilateral alteration of these holdings or by mutually agreeable exchange.

Indeed, given that (as noted above) Nozick did not argue that property rights were absolute in all circumstances, but instead proposed a proviso preventing people monopolizing the absolute necessities of life, it seems clear that Nozick did not intend that his theory should rely upon a conception of absolute property rights.

Nozick does appear to have successfully shown that "no end-state principle or distributional patterned principle of justice can be continuously realized without continuous interference with people's lives" (Nozick 1974: 163). The maintenance of an ideal distribution of income and wealth is possible only via continuous intervention to "correct" new distributions of income and wealth that emerge as a result of the actions of individuals in making, consuming, conserving and exchanging things. For Nozick (1974: 164–6), such interference is an infringement of people's rights as side-constraints—rights that imply the ability to make choices about the disposal of one's resources in accordance with one's own preferences. A society in which people's rights are violated as a matter of routine, so that they are unable to pursue their own visions of the good life unmolested, is unlikely to be one in which happiness and well-being flourish.

5. Conclusion

This chapter has argued that libertarianism is founded upon the belief that (what has been termed) a free way of living is the most conducive to human happiness. This free way of living may be understood as a framework of rights within which people may pursue diverse conceptions of the good life, so that a variety of moral values may coexist. For libertarians, justice in the distribution of income and wealth concerns the historical processes via which resources have been acquired, rather than the existence of a particular desired pattern or end-state of distribution. The pursuit of social justice via the state is perceived to violate people's rights if their legitimately acquired income and wealth are confiscated via compulsory taxation. It is claimed that such rights-violations cannot be reconciled with happiness and well-being.

Related Topics

Locke, Liberalism, Contractualism and Political Liberalism, Natural Law and Rights Theory, The Difference Principle, Left Libertarianism, Freedom

References

Buchanan, J. M. (1975) *The Limits of Liberty*, Chicago, IL: University of Chicago Press.

—— (1989) "Rational Choice Models in the Social Sciences," in *Explorations in Constitutional Economics*, College Station, TX: Texas A&M University Press.

—— (1999) "Origins, Experience and Ideas: A Retrospective Assessment," in J. M. Buchanan and R. A. Musgrave, *Public Finance and Public Choice*, Cambridge, MA: MIT Press.

—— and G. Tullock (1962) *The Calculus of Consent*, Ann Arbor: University of Michigan Press.

Cohen, G. A. (1995) *Self-Ownership, Freedom and Equality*, Cambridge: Cambridge University Press.

Fried, B. (2005) "Begging the Question with Style: *Anarchy, State, and Utopia* at Thirty Years," *Social Philosophy and Policy* 22: 221–54.

Hayek, F. A. (1948) "Individualism: True and False," in *Individualism and Economic Order*, Chicago, IL: University of Chicago Press.

—— (1976) *The Mirage of Social Justice*, London: Routledge.

—— (1988) *The Fatal Conceit*, London: Routledge.

Jasay, A. de (2002) *Justice and its Surroundings*, Indianapolis: Liberty Fund.

Kukathas, C. (2000) "Two Concepts of Liberalism," in J. C. Espada, M. F. Plattner and A. Wolfson (eds.), *The Liberal Tradition in Focus: Problems and New Perspectives*, Oxford: Lexington Books.

—— (2003) *The Liberal Archipelago: A Theory of Diversity and Freedom*, Oxford: Oxford University Press.

Locke, J. (1690/1994) *Two Treatises of Government*, P. Laslett (ed.), Cambridge: Cambridge University Press.

Mack, E. (2002) "Self-ownership, Marxism and Egalitarianism: Part I: Challenges to Historical Entitlement," *Politics, Philosophy and Economics* 1: 75–108.
—— (2009) *John Locke*, New York: Continuum.
Nagel, T. (1975) "Libertarianism without Foundations," *Yale Law Review* 85: 136–49.
Narveson, J. (2001) *The Libertarian Idea*, Ontario: Broadview Press.
Nozick, R. (1974) *Anarchy, State, and Utopia*, New York: Basic Books.
Oakeshott, M. (1949/1991) "The Political Economy of Freedom," in *Rationalism in Politics*, Indianapolis: Liberty Fund.
Rawls, J. (1999) *A Theory of Justice, Revised Edition*, Cambridge, MA: Harvard University Press.
Rothbard, M. N. (1998) *The Ethics of Liberty*, New York: New York University Press.

Further Reading

The key works of libertarian political philosophy are: R. Nozick, *Anarchy, State, and Utopia* (New York: Basic Books, 1974); J. M. Buchanan, *The Limits of Liberty* (Chicago: University of Chicago Press, 1975); J. Narveson, *The Libertarian Idea* (Ontario: Broadview Press, 2001); M. N. Rothbard, *The Ethics of Liberty* (New York: New York University Press, 1998).

D. Boaz, *Libertarianism: A Primer* (New York: Free Press, 1997) provides a reliable basic introduction to libertarian ideas. J. Paul (ed.) *Reading Nozick* (London: Rowman and Littlefield, 1981) collects the most important scholarly responses to Nozick's *Anarchy, State, and Utopia*, while R. Bader and J. Meadowcroft (eds.) *The Cambridge Companion to Nozick's Anarchy, State, and Utopia* (Cambridge: Cambridge University Press, 2011) contains reflections on Nozick's work and D. Schmidtz (ed.) *Robert Nozick* (Cambridge: Cambridge University Press, 2002) engages with Nozick's whole philosophical project.

38
DESERT

David Schmidtz

People ought to get what they deserve. What we deserve can depend on effort, performance, or on excelling in competition, even when excellence is partly a function of our natural gifts. Or so most people believe. Philosophers sometimes say otherwise. Ever since Karl Marx complained about capitalist society extracting surplus value from workers, thereby failing to give workers what they deserve, classical liberals have worried that treating justice as a matter of what people deserve would license interference with liberty. Rawls likewise rejected patterns imposed by principles of desert, saying (1971: 104),

> no one deserves his place in the distribution of natural endowments, any more than one deserves one's initial starting place in society. The assertion that a man deserves the superior character that enables him to make the effort to cultivate his abilities is equally problematic; for his character depends in large part upon fortunate family and social circumstances for which he can claim no credit. The notion of desert seems not to apply to these cases.

Rakowski (1991: 112) calls this "uncontroversial," and so does Scheffler (1992: 307). Hayek (1960: 94) says, "A good mind or a fine voice, a beautiful face or a skilful hand, and a ready wit or an attractive personality are in large measure as independent of a person's efforts as the opportunities or experiences he has had." Gauthier (1986: 220) says, "We may agree with Rawls that no one deserves her natural capacities."

Rawls's view indeed seems compelling. Inevitably, our efforts are aided by positional advantages, natural gifts, and sheer luck, so how much can we deserve? If our very characters result from an interplay of such factors, how can we (capitalists and workers alike) deserve anything at all?

Does Rawls leave any room for desert? Rawls's intent might have been narrower: simply to eliminate a rival to his difference principle as a test of the justness of basic structure. Whatever Rawls intended, though, his critique of desert has no such surgical precision. We know Rawls intended his two principles to apply only to society's basic structure, but his critique of desert is not similarly constrained, and cannot be constrained merely by stipulating or intending that it be so constrained. If Rawls's attack on principles of desert is warranted, then the skepticism he is justifying is sweeping. (Rawls sometimes says his intent is only to argue against desert as a preinstitutional notion, but see Schmidtz, 2008.) In the aftermath, Scheffler notes (1992: 301), "none of the most prominent contemporary versions of philosophical liberalism assigns a significant role to desert at the level of fundamental principle." But see Freiman and Nichols (2011).

1. The "Big Bang" Theory

Feinberg (1970: 58) coins the term "desert base" to refer to what grounds desert claims. The idea is that well-formed desert claims are three-place relations of the form "P deserves X in virtue of feature F." Many will say we deserve reward for our efforts. Some say we deserve things in virtue of being human, or being in need. We need not, and perhaps could not, produce a complete catalog of possible desert-bases. Suffice it to say, the standard bases on which persons commonly are said to be deserving include character, effort, and achievement.

What are we doing, ordinarily, when we deem someone deserving—when we acknowledge character, effort, or achievement? Suggestion: to judge Bob deserving is to judge Bob worthy. It is to judge that Bob has features that make a given outcome Bob's just reward (Sher 1987: 195; Narveson 1995: 50–1). Less obviously, to acknowledge that there are things Bob can do to be deserving is to acknowledge that Bob is a *person*: able to choose and to be responsible for his choices. (See also Lamont (1994) and Vilhauer (2009).)

By contrast, the skeptics' theory, in its most sweeping form, depicts desert as follows. To deserve X, we must supply not only inputs ordinarily thought to ground a desert claim. We also must be deserving of everything about the world, including its history, that put us in a position to supply that input. But this never happens, because even character, talent, and other internal features that constitute us as persons are products of chains of events containing at least one arbitrary link. In short, every history traces back to the Big Bang. Therefore, nothing is deserved. As Rawls repeatedly stressed (1971: 74), "Even the willingness to make an effort, to try, and so to be deserving in the ordinary sense is itself dependent upon happy family and social circumstances." (See Brock's 1999 section on "How can we deserve anything since we don't deserve our asset bases?" But for an alternative view, see Narveson (1995: 67), Sher (1987: 24), and Zaitchik (1977: 373).)

The counterargument is as follows. Needless to say, we all have whatever we have partly in virtue of luck, and luck is not a desert maker. Every outcome is influenced by factors that are morally arbitrary. Everyone is lucky to some degree. Nevertheless, the more a person supplies in terms of effort or excellence, the less weight we put on the element of luck. Crucially, being lucky is not the same as being merely lucky. The bare fact of being lucky is not what precludes being deserving. Being *merely* lucky is what precludes being deserving, because to say we are merely lucky is to say we have not supplied inputs (the effort, the excellence) that ground desert claims.

To rebut a desert claim in a given case, we must show that inputs that *can* ground desert claims are missing in that case. On a nonvacuous conception of desert, there will be inputs that a person can supply, and therefore can *fail* to supply. A general point about the methods of philosophical argument is pivotal here: namely, showing that X falls on one side of an alleged division is interesting only if falling on the other side is a real possibility.

Some causal chains work their way through features internal to persons. Desert makers, if there are any, are relations between outcomes and internal features of persons. We need not (and normally do not) assume anything about what caused those features. If an aspiring but naive skeptic says, "What people do, what they are trying to do, how well they are doing it—these things do not matter," a more worldly skeptic replies, "Compared to what?"

Is it odd that we normally make no assumptions about a desert maker's causal history? What if we had been talking about features of nonpersons? Joel Feinberg (1970: 55) observes, "Art objects deserve admiration; problems deserve careful consideration; bills of legislation deserve to be passed." John Kleinig (1971) says the Grand Canyon deserves its reputation. Such remarks are offered as small digressions, but they point to something crucial. We *never* say the Grand Canyon deserves its reputation only if it in turn deserves the natural endowments on which its reputation is based. We *never* question artistic judgments by saying, "Even the greatest of paintings were caused to have the features we admire. Not one painting ever did anything to deserve being caused to have those features." Intuitively, obviously, it doesn't matter.

Skeptics assume it does matter in the case of persons, but the assumption has never been defended, and indeed seems groundless. As with nonpersons, when a person's internal features support desert claims, the support comes from appreciating what those features are, not from evidence that they are uncaused.

Here, then, is where matters stand. We have two options. First, we can say no one deserves anything, and that is what we will say *if* we assume we can deserve credit for working hard only if we in turn deserve credit for being "destined" to work hard. The second option is to say we deserve credit for working hard not because we deserve to have been destined to work hard but simply because we did, after all, work hard. The latter is our ordinary practice.

Neither option is compelling. We are not forced to believe in desert; neither are we forced to be skeptics. We decide, and bear the consequences of deciding one way or another. A world of sweeping skepticism, where no one ever genuinely gives people credit for what they do, would be unattractive, but there is no denying that skepticism is an option, and that some do choose to be skeptics. The task, for those who *want* an alternative to skepticism, is to make room within a philosophically respectable theory for the idea that there are things we can do to be deserving.

In summary, a genuine theory of desert tells us what to look for when investigating what particular people have done. A genuine theory will not say what "Big Bang" theories say: namely, we need not investigate actual histories of particular people, since we know a priori as a perfectly general feature of human nature that no one deserves anything.

2. Deserving a Chance

When we consider how much sheer luck we needed to get where we are today, we naturally wonder, "Do I deserve this?" If we translate the question as "What did I *do* to deserve this?" then the question will have an answer. Of course, I did not do anything before I was born, or at the moment of the Big Bang, but history did not end at either of those points. It makes sense to ask "What did I do to deserve this?" and to treat what I've actually done as at least potentially relevant.

Also eminently sensible would be to ask, "What *can* I do to deserve this?" We easily can imagine it being true in a given case that, as a matter of fact, there is nothing I can do, but that is not preordained. A theory that lets us ask and answer this question is a theory that lets the concept of desert be what it needs to be in human affairs: a message of hope that is at the same time life's greatest moral challenge. Such a theory acknowledges the existence of persons: beings who make choices and who are accountable for the choices they make.

But there is a puzzle here. James Rachels (1997: 176) says, "What people deserve always depends on what they have done in the past." David Miller (1976: 93) says, "desert judgments are justified on the basis of past and present facts about individuals, never on the basis of states of affairs to be created in the future." Joel Feinberg (1970: 48; emphasis added) says, "If a person is deserving of some sort of treatment, he must, necessarily, be so in virtue of some possessed characteristic or *prior* activity."

If we are not careful, we could interpret such statements in a way that would overlook an important, perhaps the most important, kind of desert-making relation. It is conventional that what we deserve depends on what we do, and that we deserve no credit for what we do until we do it. There is a further assumption one could add to this, though.

> Further Assumption: When we first receive (for example) our natural and positional advantages, if we have not *already* done something to deserve them, it is too late.

This further assumption is unwarranted. We acknowledged that being merely lucky precludes being deserving. This does not mean that being merely lucky at t_1 precludes being deserving at t_2. In particular, we do not deserve our natural gifts at the moment of our birth, but the moment of birth is the beginning of a story, not an end. In that story, what matters, if anything at all matters, is what we go on to do. Here is a claim that may at first seem counterintuitive:

> We sometimes deserve X on the basis of what we do after receiving X.

Upon receiving a surprisingly good job offer, a new employee vows to work hard to deserve it. No one ever thinks the vow is paradoxical. No one takes the employee aside and says, "Relax. There's nothing you can do. Only the past is relevant." But unless such everyday vows are misguided, we can deserve X on the basis of what we do after receiving X.

How can this be? Isn't it a brute fact that when we ask whether a person deserves X, we look backward, not forward? If we concede for argument's sake that we look back, we would still need to ask: back from where? Perhaps we look back from where we are, yet mistakenly assume we look back from where the recipient was at the moment of receiving X. If we look back, a year after hiring Jane, wondering whether she deserved the chance, what do we ask? We ask what she *did* with it. When we do that, we *are* looking back even while looking at what happened after she received X. From that perspective, we see we can be deserving of opportunities. We deserve them by not wasting them—by giving them their due, as it were. Therefore, even if we necessarily deserve no credit for what we do until after we do it, it does not follow that if we have not already done something to deserve an opportunity by the moment we receive it, then it is too late.

To some readers, this will seem perfectly obvious. Others will feel uncomfortable with this conclusion and will want to know why they feel uncomfortable. Part of the answer is that philosophers have been taught to focus on desert as a *compensatory* notion. The idea is, desert makers we supply before getting X put a moral scale out of balance, and our getting X rebalances the scale. To those who see desert as necessarily a compensatory notion, we deserve X only if X represents a restoring of moral balance. We deserve X only if we deserve it qua reward—only if our receiving X settles an account.

Outside of academic philosophy, though, in ordinary use, desert is more than a com-

pensatory notion. Sometimes it is a *promissory* notion. Sometimes our receiving X is what puts the moral scale out of balance, and our subsequently proving ourselves worthy of X is what restores it. X need not be compensation for already having supplied the requisite desert makers. Life is full of perfectly ordinary situations where it is the other way around. There are times when supplying desert makers is what settles the account.

In either case, two things happen, and the second settles the account. In compensatory cases, desert-making inputs are supplied first; responding with rewards settles the account. In promissory cases, opportunities are supplied first; responding with desert-making inputs settles the account. Thus, a new employee who vows, "I will do justice to this opportunity. I will show you I deserve it" is not saying future events will retroactively cause her receiving X to count as settling an account *now*. Instead, she is saying future events *will* settle the account. She claims not that she is getting what she already paid for but that she is getting what she *will* pay for. (For related discussion, see Feldman (1995), Fred Miller (2001), and Waldron (1995).)

So why does James Rachels (1997: 176) assert that, "What people deserve always depends on what they have done in the past"? Rachels (1997: 180) says, "the explanation of why past actions are the only bases of desert connects with the fact that if people were never responsible for their own conduct—if strict determinism were true—no one would ever deserve anything." Crucially, when he says, "past actions are the only bases of desert," Rachels is stressing "actions," not "past." What Rachels sees as the unacceptable alternative is not a *promissory* theory, but rather the kind of *compensatory* theory that says people deserve to be rewarded for having natural endowments. He is thinking of past actions versus past nonactions, and is not considering whether actions postdating X's receipt might be relevant. (Rachels himself confirmed that this was so in a 2001 personal communication.) That is why Rachels could see himself as explaining why "past actions are the only bases of desert" when he says, "if people were never responsible for their own conduct, … no one would ever deserve anything." Notice: this argument does not tie desert-bases to events predating X's receipt. The argument connects desert to action, but not particularly to past action.

Rachels (1997: 185) also says, "People do not deserve things on account of their willingness to work, but only on account of their actually having worked." There are reasons for saying this, and Rachels may be right when speaking of rewards. It appears analytic that rewards respond to past performance. However, rewards are not the only thing that can be deserved. We sometimes have reason to say, "she deserves a chance." We may say a young job candidate deserves a chance not because of work already done but because she is plainly a talented, well-meaning person who wants the job and who will throw herself into it if given the chance.

If we say a job candidate deserves a chance, and then, far from throwing herself into the job, she treats it with contempt, that would make us wrong. The promissory aspect of desert will have failed to materialize. She had a chance to balance the account and failed. If she treats the job with contempt, then she supplies neither the performance nor even the good faith effort that the hiring committee expected.

By contrast, if the candidate fails through no fault of her own, then we cannot hold it against her. By analogy, suppose we intend to give salt a chance to dissolve in water, but what we actually end up doing is giving salt a chance to dissolve in olive oil. If the salt fails to dissolve, we still insist the salt would have dissolved in water, given the chance.

The possibility of bad luck notwithstanding, the fact remains that we sort out appli-

cants for a reason. In normal cases, the point is not to reward someone for past conduct but to get someone who can do the job. That is why, by the time we reach t_2, the question is not what she did before the opportunity but what she did with it. The question at t_2 need not and often does not turn on what was already settled at t_1.

A note on examples. Realistic examples are complex, raising issues beyond those intended by the theorist who brings them up. In this case, real-world hiring committees must juggle several criteria, not all of them having to do with desert. Some points might be better illustrated by speaking of tenure and promotion committees, where decisions are more purely a matter of desert but where candidates have enough of a history so that it is harder to sort out backward versus forward looking grounds for judging whether a candidate is deserving. Candidates often see their case as purely backward looking, but tenure committees do not. Tenure committees want to know that a candidate will not become deadwood—that past performance was not spurred mainly by a prospect of tenure qua *reward*. They want to be able to look back years later and say the candidate deserved tenure qua *opportunity*.

3. Grounding Desert

I explained how in everyday life we grasp the concept of deserving a chance in virtue of what we will do with it. I would not appeal to common sense to justify our common sense understanding, though. To justify, we look elsewhere. For example, Margaret Holmgren (1986: 274) says justice "demands that each individual be secured the most fundamental benefits in life compatible with like benefits for all," but then she adds that "the opportunity to progress by our own efforts is a fundamental interest." Richard Miller (2002: 286) concurs: "Most people (including most of the worst off) want to use what resources they have actively, to get ahead on their own steam, and this reflects a proper valuing of human capacities."

Against this, the Rawlsian supposition that inequalities should be arranged to maximally benefit the least advantaged seemingly rules out the idea that people deserve more—and thus should get more—if and when and because their talents and efforts contribute more to society. However, Holmgren notes, people in Rawls's original position would know (because by hypothesis they are aware of perfectly general features of human psychology) that people not only want to be given stuff; they want to be successful, and they want their success to be deserved. Accordingly, even grossly risk-averse contractors, focusing only on the least advantaged economic class, would be nonetheless anxious to ensure that such people have opportunities to advance by their own effort. "Rather than focusing exclusively on the share of income or wealth they would receive, they would choose a principle of distribution which would ensure that they would each have this opportunity" (Holmgren 1986: 275).

Holmgren's claim seems incompatible with Rawls's difference principle *if* we interpret that principle as Nozick interprets it, as a ground-level prescription for redistribution. In that case, the idea that Jane deserves her salary threatens to override our mandate to lay claim to her salary on behalf of the least advantaged. However, Nozick's way is not the only way to interpret the difference principle. Suppose we interpret the principle in Holmgren's way (and surely this also is Rawls's settled view): not as a mandate for redistribution but rather as a way of evaluating basic structure. That is, we evaluate basic structure by asking whether it works to the benefit of the least advantaged. On the latter interpretation, we choose among rules like "try to give people what they deserve"

and "try to give the least advantaged everything" by asking which is best for the least advantaged in actual empirical practice.

Read in Holmgren's way, Rawls's difference principle, far from competing with principles of desert, can support the idea that people deserve a chance. The difference principle supports principles of desert if Holmgren is correct to say the least advantaged want and need the chance to prosper by their own merit. Likewise, the difference principle supports principles of desert if it is historically true that the least advantaged tend to flourish within, and only within, systems in which honest hard work is respected and rewarded.

Likewise, utilitarians and nonutilitarians alike can care about consequences. Feinberg (1970: 80) says, "The awarding of prizes directly promotes cultivation of the skills which constitute bases of competition." Rawls (1971: 6) says, "Other things equal, one conception of justice is preferable to another when its broader consequences are more desirable." So, both Feinberg and Rawls can correctly insist that utility is not a desert maker, while also recognizing that (a) things that are desert makers (effort, excellence) can as a matter of fact make people better off, and that (b) making people better off is morally significant. Rachels (1997: 190) adds:

> In a system that respects deserts, someone who treats others well may expect to be treated well in return, while someone who treats others badly cannot. If this aspect of moral life were eliminated, morality would have no reward and immorality would have no bad consequences, so there would be less reason for one to be concerned with it.

Finally, consider how desert claims might be grounded in a conception of human dignity. When wondering whether a person did justice to an opportunity, we typically do not look back to events occurring before the opportunity was received. There is something necessarily and laudably ahistorical about simply respecting what people bring to the table. We respect their work, period. We admire their character, period. We do not argue (or worse, stipulate as dogma) that people are products of nature/nurture and thus ineligible for moral credit. Sometimes, we simply give people credit for what they achieve, and for what they are. And sometimes, simply giving people credit is the essence of treating them as persons rather than as mere confluences of historical forces. Thus, Walzer (1983: 260n) says he sees Rawls as supposing that persons' "capacity to make an effort or to endure pain is, like all their other capacities, only the arbitrary gift of nature or nurture." He continues:

> But this is an odd argument, for while its purpose is to leave us with persons of equal entitlement, it is hard to see that it leaves us with *persons* at all. How are we to conceive of these men and women once we have come to view their capacities and achievements as accidental accessories, like hats and coats they just happen to be wearing? How, indeed, are they to conceive of themselves?

Part of the oddity in doubting whether Jane deserves her character is that Jane's character is not something that happened to her. It *is* her. Or if we were to imagine treating Jane and her character as separate things, then it would have to be Jane's character that we credit for being of good character, so the question of why Jane per se should get the credit would be moot. In truth, of course, it is people, not their characters, that work

hard. Thus, if we say exemplary character is morally arbitrary, it is *people*, not merely character, that we are declining to take seriously.

Martin Luther King once said, "I have a dream that my four children will one day live in a nation where they will not be judged by the color of their skin but by the content of their character." This was a dream worth living and dying for. King did not dream his children would live in a nation where their characters would be seen as accidents for which they could claim no credit. King asked us to judge his children by the content of their character, not by its causes. That was the right thing to ask, because that is how we take characters (that is, persons) seriously.

If the characters of King's children are not taken seriously, they will get neither the rewards nor the opportunities they deserve. Especially by the lights of Rawls's difference principle, this should matter, for the least advantaged can least afford the self-stifling cynicism that goes with believing no one deserves anything. Neither can they afford the license for repression that goes with the *more* advantaged believing no one deserves anything.

These remarks indicate that the possibility of deserving a chance is not mere common sense. In the end, the bottom line is in part a practical question, somewhat amenable to empirical testing: which way of talking—about what people can do to be deserving—empowers people to make use of their opportunities?

4. The Limits of Desert

We considered reasons for treating it as a real possibility that people can be deserving, and in particular, that people can deserve an opportunity. Moreover, whether Jane deserved an opportunity can depend at least partly on what she did with it. *It is crucial that the scales be balanced. It is not crucial that components of the balance be supplied in a particular order.* If X is conferred first, and the desert base is supplied later, that too is a balancing of the moral scale.

So, did Wilt Chamberlain do justice to the potential created by luck in the natural lottery? (Wilt Chamberlain was the most celebrated basketball player alive when Robert Nozick was working on *Anarchy, State, and Utopia*. Nozick mentioned him in the context of an argument against patterned principles of justice.) One possible answer is that whether he did justice to his potential is no one else's business. Wilt is not indebted to anyone for his natural assets. He did not borrow his talent from a common pool. *No account is out of balance* merely in virtue of Wilt having characteristics that make him Wilt. Still, even if it is no one else's business whether Wilt does justice to his potential, the fact remains that one way or another, Wilt will do, or fail to do, justice to it.

Part of our reason for thinking it is Wilt rather than you or me who deserves credit for the excellence of Wilt's performance is that, as David Miller (1999: 144) puts it, "the performance is entirely his." Note: the issue is not whether the performance is Wilt's rather than the Big Bang's; the issue is whether the performance is Wilt's rather than some other person's. The question of whether to credit Wilt for his performance is never a question of whether Wilt caused himself to have his character and talent. Instead, the question is whether the character, talent, or other desert-making inputs are, after all, Wilt's rather than some other person's. Thus, Charles Beitz says:

> While the distribution of natural talents is arbitrary in the sense that one cannot deserve to be born with the capacity, say, to play like Rubinstein, it does

> not obviously follow that the possession of such a talent needs any justification. … A person need not justify the possession of talents, despite the fact that one cannot be said to deserve them, because they are already one's own: the prima facie right to use and control talents is fixed by natural fact.
>
> (Beitz 1979: 138; see also Spector 2011)

If we applaud Wilt's effort, we imply that the credit is due to Wilt rather than to, for example, me. Why? Not because Wilt deserved the effort (whatever that means), or deserved the effort more than I did (whatever that means) but because the effort was Wilt's rather than mine. When we ask whether the effort is truly Wilt's, the answer sometimes is yes, in contrast to cases where we instead or in addition credit Wilt's coaches or parents for performances that contributed to Wilt's in tangible ways.

Notice that giving credit is not a zero-sum game. We do not think less of Wilt when Wilt thanks his parents. Credit due to Wilt's parents takes away from credit due to Wilt only if the implication is that the performance we thought was Wilt's was not really his. (Imagine Wilt, in an acceptance speech for an academic award, thanking his tutors for writing all those term papers.)

Because giving credit is not essentially zero-sum, desert is not essentially a comparative notion. In particular, the models of desert discussed here make room not for honoring those with advantages as compared to those without, but simply for honoring people who do what they can to be deserving of their advantages. In other words, the question about Wilt is not whether Wilt did something to deserve *more than you*, but whether Wilt did something to deserve *what he has*.

Perhaps there was never a time when an impartial judge, weighing your performance against Wilt's, had reason to conclude that Wilt's prize should be larger than yours. All that happened is that Wilt did justice to his opportunity and you did justice to yours. Should we focus on the relation, or imagine there is one, between you and Wilt? Or should we focus on a pair of relations, one between what Wilt did and what Wilt has, and another between what you did and what you have? Admittedly, neither focus captures the whole truth about justice, but the second focus (that is, on the pair of relations) is a focus on desert, where the first is a focus on something else, something more comparative, such as equality.

A central distributor, aiming to distribute according to desert, would need to judge relative deserts, then distribute accordingly. Without a central distributor, the situation is different. If Wilt worked hard for X while you worked hard for Y, there is something fitting in Wilt getting X and you getting Y. You each supplied desert-making inputs connecting you to your respective salaries. It might be impossible for a central distributor to justify judging that Wilt deserves so much more than you, but by hypothesis there was no such judgment. Perhaps no amount of desert could be enough to justify that much inequality, but that would have to be argued within the context of a theory of equality, which reminds us that there is more to justice than desert, and more to desert than deserving a chance. Thus, Olsaretti (2004: 166–8) says theories of desert cannot easily justify inequality. Indeed, it could hardly be otherwise, since the idea that inequality needs justifying is controversial, and not presupposed by theories of desert. Principles of equality treat inequality as needing justification. Principles of desert do not. (See Arneson (2006) and Kagan (2012).)

Finally, a just system works to minimize the extent to which people's entitlements fly in the face of what they deserve, but what people are entitled to and what they deserve

are different things. An entitlement system is supposed to enable people to form stable mutual expectations, while empowering them to walk away from proposed terms of engagement that fail to treat them as ends in themselves. Principles of desert do not trump principles of entitlement. Nevertheless, principles of desert can and sometimes do correct the caprices of rightful entitlements. Thus, a proprietor might know her employee is entitled to a certain wage while also seeing that the employee is exceptionally productive and (in both promissory and compensatory senses) deserves a raise. If the employer cares enough about desert, she restructures her holdings (her payroll) accordingly, benefiting not only the employee but probably her company and her customers as well. Think of the contrast this way: principles of entitlement acknowledge our status as *separate* agents. Principles of desert acknowledge our status as *responsible* agents.

Note

This essay revises, condenses, and updates portions of David Schmidtz, "How To Deserve," *Political Theory*, 30 (2002) 774–99. Reprinted by permission of Sage Publications. DOI: 10.1177/0090591702238203.

Related Topics

Luck Egalitarianism, The Difference Principle, Needs and Distributive Justice, Equality, Rights, Punishment

References

Arneson, R. (2006) "Desert and Equality," in Nils Holtug and Kasper Lippert-Rasmussen, eds., *Egalitarianism*. Oxford: Oxford University Press, pp. 262–95.
Arnold, N.S. (1987) "Why Profits Are Deserved," *Ethics* 97: 387–402.
Beitz, C. (1979) *Political Theory and International Relations*. Princeton, NJ: Princeton University Press.
Brock, G. (1999) "Just Deserts and Needs," *Southern Journal of Philosophy* 37: 165–88.
Celello, P. (2009) "Against Desert as a Forward-Looking Concept," *Journal of Applied Philosophy* 26: 144–59.
Feinberg, J. (1970) *Doing & Deserving*. Princeton, NJ: Princeton University Press.
Feldman, F. (1995) "Desert: Reconsideration of Some Received Wisdom," *Mind* 104: 63–77.
Freiman, C. and Nichols, S. (2011) "Is Desert in the Details?" *Philosophy and Phenomenological Research* 82: 121–33.
Gauthier, D. (1986) *Morals By Agreement*. Oxford: Oxford University Press.
Hayek, F.A. (1960) *The Constitution of Liberty*. Chicago, IL: University of Chicago Press.
Holmgren, M. (1986) "Justifying Desert Claims: Desert and Opportunity," *Journal of Value Inquiry* 20: 265–78.
Hurka, T. (2001) "The Common Structure of Virtue and Desert," *Ethics* 112: 6–31.
Kagan, S. (2012) *The Geometry of Desert*. Oxford: Oxford University Press.
Kleinig, J. (1971) "The Concept of Desert," *American Philosophical Quarterly* 8: 71–8.
Lamont, J. (1994) "The Concept of Desert in Distributive Justice," *Philosophical Quarterly* 44: 45–64.
Miller, D. (1976) *Social Justice*. Oxford: Oxford University Press.
—— (1999) *Principles of Social Justice*. Cambridge, MA: Harvard University Press.
Miller, F. (2001) "Sovereignty and Political Rights," in Otfried Höffe, ed., *Aristoteles Politik*. Berlin: Akademie Verlag, pp. 107–19.
Miller, R. (2002) "Too Much Inequality," *Social Philosophy and Policy* 19: 275–313.
—— (2010) *Globalizing Justice: The Ethics of Poverty and Power*. Oxford: Oxford University Press.
Moriarty, J. (2009) "Deserving Jobs, Deserving Wages," in J. Smith, ed., *Normative Theory and Business Ethics*. New York: Rowman & Littlefield, pp. 119–46.

Morris, C. (1991) "Punishment and Loss of Moral Standing," *Canadian Journal of Philosophy* 21: 53–79.
Narveson, J. (1995) "Deserving Profits," in R. Cowan and M. Rizzo, eds., *Profits and Morality*. Chicago, IL: University of Chicago Press, pp. 48–97.
Nozick, R. (1974) *Anarchy, State, and Utopia*. New York: Basic Books.
Olsaretti, S. (2004) *Liberty, Desert, and the Market*. Cambridge: Cambridge University Press.
Rachels, J. (1997) "What People Deserve," *Can Ethics Provide Answers?* Lanham, MD: Rowman & Littlefield, pp. 175–97.
Rakowski, E. (1991) *Equal Justice*. Oxford: Oxford University Press.
Rawls, J. (1971) *A Theory of Justice*. Cambridge: Belknap Press.
Scheffler, S. (1992) "Responsibility, Reactive Attitudes, and Liberalism in Philosophy and Politics," *Philosophy and Public Affairs* 21: 299–323.
Schmidtz, David (2002) "How to Deserve," *Political Theory* 30: 774–99.
—— (2006) *Elements of Justice*. New York: Cambridge University Press.
—— (2008) *Person, Polis, Planet*. Oxford: Oxford University Press.
Sher, G. (1987) *Desert*. Princeton, NJ: Princeton University Press.
Spector, E. (2011) "Do You Deserve To Be Talented?" *Utilitas* 23: 115–25.
Vilhauer, B. (2009) "Free Will Skepticism and Personhood as a Desert Base," *Canadian Journal of Philosophy* 39: 489–512.
Waldron, J. (1995) "The Wisdom of the Multitude: Some Reflections on Book 3, Chap. 11 of Aristotle's *Politics*," *Political Theory* 23: 563–84.
Walzer, M. (1983) *Spheres of Justice*. New York: Basic Books.
Zaitchik, A. (1977) "On Deserving to Deserve," *Philosophy and Public Affairs* 6: 370–88.

Further Reading

Schmidtz (2006) distinguishes two elements of the promissory model.

> Element (a): A person who receives opportunity X at t_1 can be deserving at t_2 because of what she did when given a chance.
>
> Element (b): A person who receives opportunity X at t_1 can be deserving at t_1 because of what she will do if given the chance.

Regarding the question of how to deserve punishment, see Morris (1991) and Schmidtz (2008). Regarding the differences between deserving and earning, see Schmidtz (2006). On the idea of applying historical and desert-based conceptions of justice to international relations, see Richard Miller (2010). On applications of conceptions of desert to business ethics, see Arnold (1987) and Moriarty (2009). On grounding conceptions of desert, see Hurka (2001) and Celello (2009).

39
NEEDS AND DISTRIBUTIVE JUSTICE

Gillian Brock

1. Introduction

Much moral and political discourse invokes the language of needs. We press claims about our needs all the time and we sometimes do this in a way that suggests our needs ought to have a special and different status from many other considerations. What sort of moral or political relevance (if any) should people's needs be accorded?

In this chapter we cover several issues concerning distribution according to needs. We begin with five sets of reasons to be skeptical about needs and their ability to do useful work in social and political philosophy. We then move on to consider some widely used strategies for disarming skepticism of the five kinds identified. Section 4 examines the structure of claims of need. Which needs warrant normative attention? Some influential recent accounts are discussed in Section 5. We are then well positioned to begin to discuss issues concerning needs and responsibilities, which we do in Section 6. Section 7 considers some widely discussed patterns of distribution and considers the question of whether we should prefer distribution according to equality, priority, sufficiency or desert. Is it important to ensure we have enough for a decent life, as sufficientarians maintain, or should other concerns, such as with equality or desert command our attention? How does concern with need fit alongside concerns with equality and desert? We discuss issues of what it is to distribute *according to needs* in more detail in Section 8. Finally, Section 9 identifies ways in which concern with needs extends to some other debates, such as those concerning the scope of justice, both temporal and geographic.

2. What are Needs and Why Should We Care about Them? Some Issues

Not just any need has normative force. Ted claims he needs a semi-automatic weapon to kill all the students in his school who do not want to play with him. Sally claims she needs scuba diving equipment for her resort vacation. Should we feel compelled to give Ted or Sally what they claim to need? Clearly not, or, not always, especially when morally repugnant or relatively frivolous goals are at issue. Some categories of needs are eligible for more moral or political importance than others. Which needs, if any, rightly claim our normative attention?

Before we address this question, however, we need to tackle an important concern: Should we be skeptical about *any* needs being able to perform a role in social and

political philosophy? There are two main types of skepticism about projects which aim to show needs have some political importance. One set of concerns centers on conceptual issues, such as doubt about the availability of clear criteria for *determining* needs, especially those thought to be politically salient and which can ground policy. The second area of skepticism concerns needs playing a role in policy decisions being *undesirable* or philosophically indefensible. In this section I canvas some of these sorts of complaints, but show how they are not insurmountable.

If there is to be distribution according to need, it seems we require a universal account of what our needs are. However, some are skeptical that we can derive one list of needs applicable to all people. People claim to need very different things. One person might claim to have social needs, another wants nothing but to be left in splendid isolation. Some claim to need religion in their lives to give it meaning; others believe this to be nothing but a dangerous opiate. Can we compile a list of "real" needs shared by all human beings? It looks unlikely, considering the enormous range of things people claim to need, especially if we take into account a variety of different cultures and diverse worldviews. A second worry is that there is no sharp distinction that can be drawn between needs, on the one hand, and desires, wants, or preferences, on the other. Some things we might at one time have thought of as (mere) preferences, desires or wants have, over time, come to be insisted upon as needs. Refrigeration, central heating, air conditioning, access to the internet, and television sets are some examples. Does this not show that there is, after all, no sharp distinction between needs and wants? Are needs really just things we want? A third set of concerns is that distribution according to need interferes with justice, freedom, or rights. On at least one influential account of distributive justice, namely a libertarian account, if there is to be state distribution according to need this will necessarily interfere with our freedom, rights, or entitlements. A fourth set of worries is that using needs as a criterion of distribution lends itself to paternalism, authoritarianism, abuse, or is otherwise anti-democratic. As already mentioned, it seems to be a fact of life that individuals have different needs. Should we not all be allowed the freedom to decide for ourselves what our own needs are and to pursue our needs as we see fit? If some state officials and experts are responsible for determining what our real needs are or which needs deserve state support in being satisfied, the path is open for paternalism or authoritarianism with respect to needs policy.

A final cluster of concerns deserves mention. Many of our needs are expensive to satisfy. Many medical needs (such as for renal dialysis machines or equipment required for brain surgery) or educational needs (such as training and equipment for specialist programs) can be enormously expensive. If we have responsibilities to meet needs, we will have too many responsibilities. Moreover, these demands will place significant burdens on those obligated to assist with needs-meeting. Considering the total number of people in need on the planet, if needs have normative force, we will surely be overwhelmed with pressing demands on us. If we recognize the force of need, we will be able to do little else but respond to the bottomless pit of needs that surrounds us.

3. Strategies for Disarming the Objections

Several strategies are frequently deployed to disarm such objections. The key to addressing the relativity and indeterminacy type objections is to show that there are clear boundaries to what counts as a need, and these boundaries can be drawn in principled ways (e.g. Braybrooke 1987; Doyal and Gough 1991; Wiggins 1987). Moreover, we can

show that there are reasonably clear criteria for determining the needs that will have normative implications or generate political responsibilities. Some of these objections exaggerate the difficulties involved in determining the needs that are to have political importance. The concept of needs has a core area of application that is not so essentially contested as to be completely arbitrary, relative or indeterminate. There are at least some determinate limits to what counts as a human need and, this being the case, the force of the first two complaints is softened.

The suspicions that distribution according to needs must lend itself to authoritarian purposes or involve paternalism all rest on certain dispensable assumptions. It is assumed that individuals in need will have no part in determining what they need, what form satisfaction of those needs will take, and policies surrounding their needs and how to meet them. But why assume that individuals in need will not participate in these decisions? Indeed, some have argued that an acceptable decision procedure must allow for such participation (Braybrooke 1987; Brock 1994; Doyal and Gough 1991). So, paternalism and non-democratic mechanisms are not inevitably involved in distribution according to need, but rather tied to particular decision-making procedures for arriving at policies about needs, or tied to other particular features of policies about needs which are subject to modification.

In addressing the fifth concern about needs generating too many responsibilities, it is important to point out that there are at least two questions that need not be conflated. Recognizing that *some* claims of need have normative force is a separate issue from the issues of the responsibilities that needs generate and how to distribute responsibilities for meeting needs (Braybrooke 1987; Miller 1999; Miller 2007). Furthermore, claims of need might have variable importance and might need to be especially sensitive to the resources available and the burdens that meeting them would place on others (Copp 1998).

The third set of concerns—that distribution according to need interferes with justice, rights, or freedom—requires a more sustained and substantive engagement with the ideas of what justice requires. In Section 6, I explore several common ways of arguing that, in contrast to this complaint, justice requires attention to need and, moreover, there is scope for concern with needs that does not crowd out appropriate attention to other matters of justice, freedom or rights. Next, however, more analysis of the structure of needs claims will assist in tackling the rest of the catalog of skeptical worries.

4. Claims of Need

Coherent statements of need have at least the following structure: x needs y in order to z. For a coherent needs claim we can always fill in a purpose or end-state, z, to be realized by x's having y, no matter how obvious or mundane the end-state, z, might be (Thomson 1987; Wiggins 1987). Statements of need are essentially instrumental claims. They express a relation that x's having y is needed to achieve some end, z. The normative force of the needs claim will vary depending on the end-state involved (Thomson 1987; Wiggins 1987). Clearly, not all statements of need make defensible claims on others: people might claim to need things for morally repugnant or trivial ends. It is, however, widely held in the literature that there is a group of needs that can generate normative force in a more categorical way. Our most fundamental, basic, or human needs, those central to continued human existence, can do this (Braybrooke 1987; Brock 1998; Copp 1998; Thomson 1987; Wiggins 1987). I turn now to consider some recent accounts of

how to pick out such needs. We notice there are different but importantly overlapping ways of picking out the relevant categories.

5. Some Influential Recent Accounts of Needs Salient to Distributive Justice

Various authors pick out central categories of needs as warranting normative attention. Harry Frankfurt (1998) suggests that there are two necessary conditions for a need's deserving moral importance: a need is morally important if harm typically results when the need is not met and that harm is outside the person's voluntary control. David Wiggins (1987, 1998) also believes harm arising from factors beyond one's control is central to why needs matter, when they do. To develop this idea, he introduces several terms to reflect important differences. He distinguishes the gravity of the harm that would ensue if the need is not met ("*badness*"), from the *urgency* with which harm would ensue. A need is *basic* if it results from a law of nature, an unalterable and invariable environmental fact, or a fact about human constitution. Needs can be *entrenched* when they are inflexible to modification, or *substitutable* when they are not. Using this terminology Wiggins defines vital needs as ones that are badly needed in a way that is entrenched and scarcely substitutable, and it is these vital needs that matter morally according to him (Wiggins 1987, 1998).

David Braybrooke's account (especially as developed in *Meeting Needs*) is one of the most influential in the philosophical literature. He develops an account of basic needs in terms of what is necessary for social functioning. Something is a need if, without its satisfaction, one would be unable to carry out four basic social roles, namely, those of citizen, parent, householder, and worker. By examining several lists proposed by the United Nations and others, he extracts their common elements and offers a systematic account of the needs one would have over the course of a life. The list consists of needs for a life-supporting relation to the environment; for whatever is indispensable to preserving the body intact in important respects (including food, water, exercise, and periodic rest); for companionship; for education; for social acceptance and recognition; for sexual activity; for recreation; and for freedom from harassment, including not being continually frightened. Focusing on what humans typically do (through consideration of roles) provides Braybrooke with a good reference point for compiling this list. Someone might claim that not all these needs apply to everyone (e.g., the need for sexual activity for a nun or the need for companionship for a hermit), but it is important to note that Braybrooke is interested in deriving a list of items that are plausibly needed in order to carry out the four roles he identifies. In order to have the genuine choice to perform the role (such as that of being a parent), one typically needs items on the list.

Another highly influential account which bridges philosophical and empirical disciplines is that of Len Doyal and Ian Gough (1991). Doyal and Gough's view is that needs are universalizable preconditions that enable non-impaired participation in any form of life. Chief among these preconditions will be physical health and the mental competence to deliberate and choose, or autonomy. They recognize a class of "intermediate needs," which aim to connect the two basic needs with knowledge available about basic needs in the social sciences. These are: nutritious food and clean water, protective housing, a non-hazardous work environment, a non-hazardous physical environment, appropriate health care, security in childhood, significant primary relationships, physical security, economic security, appropriate education, safe birth control, and safe

child-bearing. Their account provides important connections between the philosophical literature and the social and natural sciences, which facilitates measuring progress with respect to meeting needs in the world.

Braybrooke's and Doyal and Gough's accounts highlight important features of recent accounts of basic needs: the importance of social (not just physical) functioning in particular communities; the relevance of information about human needs collected by the natural and social sciences; and the importance of cross-cultural comparison. More generally, there are several common elements to these and other recent accounts of normatively salient needs. The needs that matter morally are those that are necessary, indispensable, or inescapable, at least with respect to human functioning in social groups (Brock 1998; Thomson 1987; Wiggins 1998). Moreover, if such needs are not met, we are unable to do anything much at all, let alone lead a recognizably human life (Copp 1998; Nussbaum 1998). Having needs satisfied is essential to our ability to function as human agents (Copp 1998; Gewirth 1978; O'Neill 1998; Shue 1996).

While one dominant approach in the needs literature is to emphasize the link between needs and human agency, other approaches can be discerned, such as, connecting our basic needs with what is required for human flourishing (Nussbaum 1998; Reader 2005). This kind of more expansive account is vulnerable to more of the catalog of skeptical concerns about basic needs canvassed in Section 2, such as that needs cannot be adequately distinguished from people's wants, preferences, or desires, since what is needed for human *flourishing*, as opposed to a more limited goal (such as *basic* human functioning), can quickly expand.

6. Needs and Responsibilities

Do we have responsibilities to meet needs? Several kinds of arguments can be discerned in the literature. Next we discuss some common argument strategies.

One strategy starts from showing that having needs met is a necessary condition for human agency, autonomy, or good human functioning and then proceeds to argue that these special necessary conditions should be protected or underwritten by communities of cooperators (Baker and Jones 1998; Copp 1998; Gewirth 1978; Wiggins 1987). Insofar as we care about the benefits of cooperation, we have reasons to protect its bases.

Another strategy commonly pursued in the literature is to show how people are especially vulnerable to coercion or having their rights violated if they do not have their needs met (Braybrooke 1987; Gewirth 1978; Goodin 1998; O'Neill 1998; Shue 1996). Being needy in certain circumstances (especially ones in which there are vast power differences between the needy and the affluent) can lead to opportunities for coercion. But whatever else we expect of a state, we minimally expect it to protect us from rights' violations and coercion, so responsibilities to remove our vulnerabilities to force and injustice and so to meet certain kinds of needs, are hereby generated.

Related to this line of argument is another common strategy—one that examines the task of government. At the very least, governments should protect vital interests and enable citizens to meet their needs for themselves (insofar as this is possible). Meeting needs is therefore a crucial part of governments' core function, according to this line of argument (Braybrooke 1987; Nussbaum 1998; Wiggins 1987). Others suggest that in addition, part of government's job is to act in certain kinds of cases as a coordinator of our moral responsibilities (Goodin 1998). If we all have some responsibilities to help

others with their needs (as we do), an efficient way to coordinate and discharge these responsibilities may well be through state-organized assistance.

According to libertarians, by contrast, the job of government is quite minimal: to ensure there are no rights violations or liberty infringements. There are many arguments that specifically target libertarian interlocutors and their conception of justice with the intention of showing how making space for meeting needs is necessary for the coherence of their own accounts (e.g. Baker and Jones 1998; Brock 1995; Sterba 1998). These arguments also often aim to show that cherished libertarian beliefs about the nature of rights to life, liberty or property, when examined, commit one to meeting needs, given a detailed analysis of the nature of the rights and liberties at issue (Baker and Jones 1998; Brock 1995; Gewirth 1978; Shue 1996; Sterba 1998). Others argue that the ideals of self-reliance with respect to meeting needs, when examined, are untenable without considerable concession in the direction of support for needs. Some needs are for services that no one can seriously be expected to provide for herself (e.g. education or health care), some needs are intrinsically social (such as the need for companionship) and some people (such as very young children) are incapable of meeting their own needs for themselves. Baker and Jones (1998) argue that the common libertarian ideal of self-reliance—that of strict individual responsibilities with respect to needs satisfaction—entails collective responsibilities to meet needs and provide social structures that are organized to ensure that individuals have real opportunities to satisfy their own basic needs.

We see, then, that several theorists argue that enabling citizens to meet needs is, or should be, a legitimate and core function of governments, across a wide ideological spectrum of political views.

There is a widespread misconception among critics about the content of the responsibility with respect to meeting needs. The responsibility to meet needs is often thought to entail directly meeting needs, for instance, supplying food for the hungry. This may or may not be the direct object of the responsibility. Rather, a preferred way of expressing the content of duties is that we have a responsibility to enable others to meet their needs themselves (in at least certain cases). Enabling is a process that involves a number of different elements depending, for instance, on the capacities of the person. If x enables y to z, this might involve a point at which y may need to be simply given goods directly to accomplish ends of the *z*-kind, a point at which y may need to be taught certain skills to accomplish z, a point at which y may be helped with opportunities to exercise the skills to accomplish z, and possibly a point at which y may need to be helped to be weaned off dependence on x. Helping with needs might focus on provision for all the stages involved in enabling one to meet one's needs oneself.

How much priority should the activity of enabling people to meet needs have relative to other activities or demands for resources? What normative significance should needs have? It is commonly held that needs have or should have a certain lexicographical priority over desires, preferences and wants (Braybrooke 1987; Frankfurt 1998; Wiggins 1987). Having a need for something makes a more compelling claim than having a desire or preference for the same thing (Braybrooke 1987; Frankfurt 1998; Wiggins 1987). Harry Frankfurt describes a Principle of Precedence which captures this insight. He notices, for instance, that people are:

> widely disposed to accept the proposition that a need for something preempts a desire for that thing. This proposition, which I shall call the Principle of

> Precedence, attributes to needs only a quite minimal moral superiority over desires. It maintains no more than that when there is competition between a desire and a need for the same thing, the need starts with a certain moral edge. That is, when A needs something that B wants but does not need, then meeting A's need is prima facie morally preferable to satisfying B's desire.
>
> (Frankfurt 1998: 20)

Claims in virtue of needs typically enjoy a prima facie advantage over claims pressed using the discourse of desires. Needs sometimes also enjoy a similar advantage when in competition with other claims. In matters of distributive justice there is the question of how needs should relate to claims pressed in virtue of desert or equality, which also have a role to play in distributive contexts. David Miller (1999) argues, for instance, that all three of these important concepts (needs, desert, and equality) should play basic roles in a theory of social justice, with their importance varying depending on what is being distributed to whom and by whom. In some contexts need seems to matter greatly, for instance when considering the distribution of medical resources. In other contexts, equality or desert might be a more salient consideration. While there certainly are connections between each of these central concepts (for instance, meeting needs might be a way of pursuing equality in certain cases), each of these concepts would seem to engage with a different core concern in matters of distributive justice and so each should play an important and distinct role in a full account of distributive justice. We discuss some of these ideas in more detail in sections below.

7. Patterns of Distribution: Should we Prefer Equality, Priority, Sufficiency or Desert?

In grappling with basic questions of justice, several questions are often distinguished, such as these three:

- To whom is justice owed?
- What should we be distributing? Should we, for instance be distributing well-being, initial opportunity for well-being, brute-luck well-being, resources, primary goods, capabilities, social/political status or freedom?
- Should distribution follow any patterns, such as, according to ideas of equality, priority, sufficiency or desert?

Egalitarian accounts of justice are concerned with equality of relevant benefits. Prioritarianism holds that it is more important to provide benefits to those who have less. Sufficientarianism holds that justice requires that everyone get an adequate (or sufficient) amount of specified goods, such as what is required to meet basic needs, and that aiming at equality beyond that sufficiency threshold is not typically the appropriate focus (Frankfurt 1998).

While prioritarianism and sufficientarianism generally give priority to those who are worse off, desert-based approaches, by contrast, distribute according to a criterion of desert typically thought not to include concern for neediness. Desert criteria vary, but the criteria of effort and contribution are often taken as salient desert bases: what x deserves should correlate with how much effort or contribution x made.

Frankfurt maintains that we often misidentify our reasons for concern in situations

where some have much less, as involving concern with inequality. However, in such cases it is frequently really the importance of having enough and insufficiency that does the normative work. What troubles us is that some fall well short of a threshold of adequacy while others are greatly beyond it. Unmet need often supplies much of the normative force when claims are pressed in virtue of concern for inequality. While some inequality does command our attention, not all inequality should. We need not be troubled by inequalities between millionaires and billionaires, and the fact that they are both well beyond some threshold of having enough for a decent life operates as an important consideration. Appreciating this would mean we should endorse sufficientarian accounts of obligations of distributive justice, especially those that take needs as salient in preference to egalitarianism (Frankfurt 1998).

8. Distribution According to Need

What is it to distribute *according to needs*? David Miller (1999) offers one of the most detailed and sustained contemporary analyses of this idea. He is concerned to interpret this ideal in circumstances in which there is scarcity and we cannot meet all needs. How are we to mediate in such cases among conflicts with people's claims of need?

Like other authors that we have discussed in Section 5, Miller argues that we can give a reasonably clear account of what is needed for a decent life within a particular society, and so needs can be given the necessary determinacy and objectivity to function sufficiently clearly as a criterion for distribution. Miller identifies two distribution problems. First, there is the problem of mediating between needs claims and non-needs claims. This problem involves the weight claims of needs should enjoy compared with claims made in virtue of competing considerations such as desert, productive contribution, rights, and other important projects which we, as a society, might deem worthy of resource allocation, such as protecting the environment or investing in infrastructure. Second, various different needs claims compete for normative attention and there is the problem of mediating among these claims, such as in the case of the need for medical attention versus the need for housing. There might be competition among different types of needs claims, and the claims might vary in intensity, extent, urgency, kind, and the like. Alice needs medical attention, Bob needs housing. Who is in greater need or to which need should we direct our attention if we cannot do both?

Putting problems of commensurability to one side, let us consider cases of only one type and assume commensurability so that we can place people all on one scale. One way to interpret "to each according to need" is to allocate resources *in proportion to need*. So on this account, if Carol scores −50 on the scale of need while David scores only −10, then Carol should receive five times as many resources as David does (Miller 1999: 214). But we could challenge this. This proposed distribution does not take account of *the effect the resources have on the needs*, so we could modify this first proposal by stipulating that we distribute so that everyone enjoys a *proportionate increase in the capacity to function*, thereby satisfying everyone's needs to the same (proportional) extent.

But alternatives could also be proposed: Why not give more weight to Carol's greater need, giving her priority at least until she is on a par with David? Should we give strict priority to the more needy? The practice of triage suggests this would be unwise. Devoting all our resources to the most needy might mean there are none left for the less needy, but those who might have greater prospects of recovery. Triage is a well-known practice for using scarce resources most efficiently, for instance in battle situations. Casualties

are divided into three categories: those who are badly wounded but whose chance of recovery is small even with extensive medical resources, those who are in need of medical attention and if and only if they receive it, have good prospects for recovery, and those who will recover eventually without treatment or resource use. Under a triage policy, the second group is prioritized, even though they have less urgent needs than the first group, because scarce resources allocated here have the biggest impact on favorable outcomes.

Distributing according to triage principles is another plausible way of interpreting the injunction to distribute according to needs. Furthermore, other ways of distributing in the face of need are plausible. If twenty trapped miners are in need of food and there is only one loaf of bread, a plausible way to meet needs here (assuming no further relevant differences between parties) is to give each person one twentieth of the loaf, even though everyone will still be hungry afterwards and have some unfulfilled need for food. Consider also the case of there being one kidney available for transplant and twenty people who need a kidney, whose medical situation is equivalent. In such cases deciding who gets the kidney via a lottery would be a procedurally fair way to decide whose need is to be satisfied.

The cases canvassed above show that the following are all plausible ways to negotiate situations of competition among claims of need: (1) strict priority to the most needy, (2) priority to those whose needs can be efficiently or effectively met with scarce resources, (3) equality in provision for needs even though all remain needy, and (4) meeting needs via lotteries in cases of extreme scarcity. This range of ways of addressing unmet needs is not necessarily any kind of problem; rather it points to complexity in the face of (1) diverse needs, (2) diverse ways of satisfying needs, (3) needs that have different urgencies, (4) differences in whether or not needs can be partially satisfied, (5) diverse capacities to convert resources into states of need satisfaction, and the like. Similar difficulties arise with distribution in accordance with other salient considerations including desert, productive contribution, and effort.

According to Miller, the slogan to distribute "to each according to needs" encourages us to do the following:

> [W]hen goods are fully divisible and transferable, it tells us to get people as close as possible to one another's position on the scale of need: if my needs are half-satisfied, so should yours be. Usually this will mean giving priority to those whose needs are greatest, but … not always. Difficult problems arise when resources are not fully divisible or transferable. Here we may have to choose between meeting needs equally (comparative justice) and not wasting or withholding need-satisfying resources (noncomparative justice). Our decision will depend on the absolute magnitude of the cost involved in withholding the resource. Wherever the cost is high, we should look for a fair procedure for assigning the resource. In this way, though we cannot treat everyone's needs equally as the principle ideally requires, we can at least show everyone a form of equal respect.
>
> (Miller 1999: 221)

The obligation to meet needs or enable people to meet needs is often thought to be subject to something like a threshold criterion of taking responsibility for meeting needs. Of course it is no easy matter to decide for which activities people can properly be

held responsible. However, there will be community views about what is a matter of fate and what a matter of individual responsibility that can assist in making such determinations. Furthermore, if we supply resources for someone to meet needs which are then squandered, the obligation to meet needs in virtue of justice has been discharged, notwithstanding the fact that the resources have not succeeded in meeting needs. In general we do not have an obligation to supply further resources, not at any rate in virtue of justice, though we might still have a duty in virtue of humanitarian considerations.

9. Connections to Other Debates and Conclusions

The debate concerning meeting needs extends in various directions. Importantly, there is the issue of the scope of justice: whose needs deserve consideration when we are distributing according to need? Are we implicitly talking about compatriots only or does the concern extend across borders to include non-compatriots (Brock 2009; D. Miller 1995)? And what about future generations? Should their needs act as constraints on how we may permissibly distribute resources now?

Here I have space to sketch the way in which the debate expands in just one direction and so, briefly, do this for the case of the extension across state borders, typically thought to limit our responsibilities in important respects. While most philosophers agree that there is some duty to assist with global poverty, there is vigorous debate concerning the grounds for this duty and also its content (D. Miller 2007; R. Miller 2010; Pogge 2002; Rawls 1999). For instance, according to utilitarians if the situation of some of the world's poor can be improved through transfers from wealthier nations, which result in no significant sacrifice to wealthier nations, then this is what is required (Singer 1972). Many argue that such positions leave inadequate scope for personal projects and prerogatives that also make up meaningful individual lives (Cullity 2004; Miller 2010).

A dominant deontological approach to duties of global justice is one which emphasizes protection of basic rights, such as Henry Shue does in his important work *Basic Rights* (1996). There he argues that if we accept that all people have a right to security (as is commonly assumed when we endorse the so-called "negative rights" to be free from assault and non-interference more generally), then all people also have rights to subsistence. Shue argues that there is no morally salient difference between negative and positive rights: they both require positive and negative actions for adequate protection. Consider the right to security. In order to protect the right to security, agents must both refrain from assault (a so-called "negative duty") and offer them protections against standard threats such as having police officers would provide (standardly termed a "positive duty").

Despite this important critique, the distinction between negative and positive rights continues to have a powerful influence on people's conception of what we owe one another. However, Thomas Pogge (2002) marshals a significant argument that even if we have only negative duties to refrain from harming one another, we can still generate strong obligations to the global poor in virtue of the many ways in which we have harmed and continue to harm them through the global institutional order we sustain. There are many grounds on which critics challenge such conclusions. Some challenge the idea that the global order really harms the poor or question the individual's contribution to such harm (Risse 2005). Others argue that in the absence of ways of enforcing justice, or special relationships of the right kind among people, there are no duties

of justice, though there may be humanitarian duties (D. Miller 2007; R. Miller 2010; Nagel 2005).

It is worth noting that basic needs have played an important role in global public policy matters, for instance, in the so-called Basic Needs Approach, introduced by Paul Streeten in the 1970s (Streeten *et al.* 1982). The idea was to identify universal basic needs and then to provide the means to meet these to communities in an attempt to address global poverty. However, the implementation of this program suffered from several avoidable problems, such as excessive paternalism and commodity-focus, with the result that the capabilities approach is often thought to be superior, especially in its ability to avoid these dangers. Whether or not the basic needs and capabilities approaches are necessarily distinct is also subject to debate, as is the matter of whether the dangers often associated with the Basic Needs Approach must always accompany it (Reader 2005).

In this chapter we have addressed several debates concerning needs and distributive justice. We saw that among needs-theorists there is convergence on what counts as a need and the needs claims that give rise to appropriate claims of justice. Clarity concerning the needs claims of normative import removes a variety of sources of skepticism concerning needs playing a role in social and political philosophy and policy. Needs can be a source of powerful responsibilities grounded in concern for justice, according to a wide range of theories of justice. There is a role for needs to play in distributive justice that does not displace concern with several rival concepts such as desert or equality. Though there are several ways of interpreting the advice to distribute "to each according to his needs," some important guidance on how to establish priorities among different needs claims is also available.

Related Topics

The Difference Principle, Desert, The Capability Approach (and Social Justice), Human Rights and Cosmopolitanism, Global Justice and Politics, Rights

References

Baker, J. and Jones, C. (1998) "Responsibility for Needs," in G. Brock (ed.), *Necessary Goods: Our Responsibilities to Meet Others' Needs*, Oxford: Rowman & Littlefield.

Braybrooke, D. (1987) *Meeting Needs*, Princeton, NJ: Princeton University Press.

Brock, G. (1994) "Braybrooke on Needs," *Ethics* 104: 811–823.

—— (1995) "Is Redistribution to Help the Needy Unjust?" *Analysis* 55: 50–60.

—— (ed.) (1998) *Necessary Goods: Our Responsibilities to Meet Others' Needs*, Oxford: Rowman & Littlefield.

—— (2009) *Global Justice: A Cosmopolitan Account*, Oxford: Oxford University Press.

Copp, D. (1998) "Equality, Justice, and the Basic Needs," in G. Brock (ed.), *Necessary Goods: Our Responsibilities to Meet Others' Needs*, Oxford: Rowman & Littlefield.

Cullity, G. (2004) *The Moral Demands of Affluence*, Oxford: Clarendon Press.

Doyal, L. and Gough, I. (1991) *A Theory of Human Need*, Houndsmills: Macmillan.

Frankfurt, H. (1998) "Necessity and Desire," in G. Brock (ed.), *Necessary Goods: Our Responsibilities to Meet Others' Needs*, Oxford: Rowman & Littlefield.

Gewirth, A. (1978) *Reason and Morality*, Chicago, IL: University of Chicago Press.

Goodin, R. (1998) "Vulnerabilities and Responsibilities: An Ethical Defense of the Welfare State," in G. Brock (ed.), *Necessary Goods: Our Responsibilities to Meet Others' Needs*, Oxford: Rowman & Littlefield.

Miller, D. (1995) *On Nationality*, Oxford: Oxford University Press.

—— (1999) *Principles of Social Justice*, Cambridge, MA: Harvard University Press.

—— (2007) *National Responsibility and Global Justice*, Oxford: Oxford University Press.
Miller, R. (2010). *Globalizing Justice: The Ethics of Poverty and Power*, Oxford: Oxford University Press.
Nagel, T. (2005) "The Problem of Global Justice," *Philosophy and Public Affairs* 33: 113–147.
Nussbaum, M. (1998) "Aristotelian Social Democracy," in G. Brock (ed.), *Necessary Goods: Our Responsibilities to Meet Others' Needs*, Oxford: Rowman & Littlefield.
O'Neill, O. (1998) "Rights, Obligations, and Needs," in G. Brock (ed.), *Necessary Goods: Our Responsibilities to Meet Others' Needs*, Oxford: Rowman & Littlefield.
Pogge, T. (2002) *World Poverty and Human Rights*, Cambridge: Polity.
Rawls, J. (1999) *The Law of Peoples*, Cambridge, MA: Harvard University Press.
Reader, S. (ed.) (2005) *The Philosophy of Need*, Cambridge: Cambridge University Press.
Risse, M. (2005) "How Does the Global Order Harm the Poor?" *Philosophy and Public Affairs* 33: 349–376.
Shue, H. (1996) *Basic Rights: Subsistence, Affluence and US Foreign Policy*, 2nd edn., Princeton, NJ: Princeton University Press.
Singer, P. (1972) "Famine Affluence and Morality," *Philosophy and Public Affairs* 1.
Sterba, J. (1998) "From Liberty to Universal Welfare," in G. Brock (ed.), *Necessary Goods: Our Responsibilities to Meet Others' Needs*, Oxford: Rowman & Littlefield.
Streeten, P. *et al.* (1982) *First Things First*, Oxford: Oxford University Press. World Bank Research Publication.
Thomson, G. (1987) *Needs*, London: Routledge.
Wiggins, D. (1987) "Claims of Need," in D. Wiggins, *Needs, Values, Truth*, Oxford: Oxford University Press.
—— (1998) "What Is the Force of the Claim That One Needs Something?" in G. Brock (ed.), *Necessary Goods: Our Responsibilities to Meet Others' Needs*, Oxford: Rowman & Littlefield.

40
THE CAPABILITY APPROACH (AND SOCIAL JUSTICE)

Ingrid Robeyns

In its most general description, the capability approach is a flexible and multi-purpose normative framework, rather than a precise theory of well-being, freedom or justice. At its core are two normative claims: first, that the freedom to achieve well-being is of primary moral importance, and second, that freedom to achieve well-being is to be understood in terms of people's capabilities, that is, their real opportunities to do and be what they have reason to value. This framework can be used for a range of evaluative exercises, including most prominently the following: (1) the assessment of individual well-being; (2) the evaluation and assessment of social arrangements, including assessments of social and distributive justice; and (3) the design of policies and proposals about social change in society. In all these normative endeavors, the capability approach prioritizes (a selection of) peoples' beings and doings and their opportunities to realize those beings and doings (such as their genuine opportunities to be educated, their ability to move around or to enjoy supportive social relationships). This stands in contrast to other accounts of well-being, which focus exclusively on subjective categories (such as happiness) or on the means to well-being (such as resources such as income or wealth).

This chapter will be limited to the capability approach to justice, which is the area within political philosophy where the capability approach is the furthest developed and has up to now made most impact. Yet it is instructive to note that there are other debates in social and political philosophy where the capability approach has played a role, such as providing a metric for an all-things-considered evaluation of institutions (Claassen 2009), or as the basis to conceptualize key concepts, such as development (Sen 1989) or education (Robeyns 2006).

This chapter will address the following main questions. First, which metric of justice does the capability approach defend, and how does that compare with other metrics of justice? Second, has the capability approach succeeded in providing an alternative to John Rawls's theory of "Justice as Fairness"? Third, which capabilities are relevant to questions of justice? Fourth, since the capability approach endorses a multidimensional metric of justice, how should these dimensions be aggregated to come to an overall judgment, or are there ways to circumvent the need for aggregation? Finally, which other

theoretical "modules" need to be added to the metric of justice for a complete capability theory of justice?

1. Capabilities as a Metric of Justice

A theory of distributive justice can be compared with an engine, in which the capability approach provides only one part, or "module" (Arneson 2010: 103). Not all theories of justice have exactly the same modules, yet theories of distributive justice must specify at least two things: a metric and a distributive rule (Anderson 2010: 81). The "metric of justice" (also called "the currency of justice") is the good that is to be compared between individuals when making statements of distributive injustices, hence the dimensions in which the interpersonal comparisons that are an integral part of claims of justice are made. The distributive rule specifies what justice requires in terms of the distribution of that good between people. Typical distributive rules are equality, sufficiency (minima that everyone should be situated above), or priority to the worst off. Moreover, there are other modules to the full theory of distributive justice, such as the grounds for its principles of justice, or a specification of the duties of justice rather than merely the rights of justice (see Section 5). As Arneson (2010: 103–4) points out, one of the theoretical difficulties of comparing the capability approach to justice with other theories is that one limits the comparison to only one module, whereas the comparison really requires an assessment of the entire engine. We can't make an overall comparison of, say, cars, by only comparing their wheels or their maximum speed. Thus, while some political philosophers mistakenly think the capability approach is a theory of equality or a theory of justice, the capability approach *only* specifies a metric of justice, and hence a variety of capability theories are possible, depending on the other modules that are integrated.

Capability theories of justice argue that when making those interpersonal comparisons needed for justice, we should focus on people's functionings and capabilities. Functionings are "beings and doings," that is, various states of human beings and activities that a person can undertake. Examples of the former (the "beings") are being well-nourished, being undernourished, being housed in a pleasantly warm but not excessively hot house, being educated, being illiterate, being part of a supportive social network, being part of a criminal network, and being depressed. Examples of the second group of functionings (the "doings") are travelling, caring for a child, voting in an election, taking part in a debate, taking drugs, killing animals, eating animals, consuming lots of fuel in order to heat one's house, and donating money to charity. To every functioning corresponds a capability, being the real opportunity one has to achieve that functioning. If we say that person A has a capability to functioning X, we are saying that if A chooses to (be or do) X, then A will succeed in (doing or being) X. Hence, capabilities stand to functionings as an opportunity stands to outcome, or the potential stands to the realized.

What reasons are there to focus on the evaluative space of functionings and capabilities, rather than other possible metrics of justice, such as Rawlsian primary goods or happiness? The capability approach holds that judgments of justice should focus on what really matters to people, which is their freedom to achieve well-being. Thus, we should focus on the ends rather than the means of well-being, and the capability approach postulates that these ends are what people are able to be and to do, rather than their happiness.

The capability approach argues against subjective metrics of justice on at least two grounds (Anderson 2010: 85–7). First, subjective metrics are sensitive to problems of

adaptation, that is, the problem that people adapt their subjective well-being to their deprived circumstances or oppressive social norms. The opposite kind of adaptation, whereby one gets used to high levels of affluence or the privileges caused by social hierarchies, is possible too—a phenomenon known as "expensive tastes" in theories of equality and justice (Dworkin 1981a).

Arguments against subjective metrics of justice could be used to motivate a shift to resources, seen as the means of justice, as has been done in the theories of justice defended by John Rawls (1971) and Ronald Dworkin (1981a, b). Yet the problem with shifting to means is that people differ in their ability to convert means into valuable opportunities (capabilities) or outcomes (functionings) (Sen 1992: 26–8, 36–8). Since ends are what ultimately matter when thinking about well-being and the quality of life, means can only work as reliable proxies of people's opportunities to achieve those ends if all people have the same capacities or powers to convert those means into equal capability sets. Capability scholars believe that these inter-individual differences are far-reaching and significant, and that theories that focus on means tend to downplay their normative relevance. The sources of inter-individual differences to convert means into ends can be personal (e.g. impairments), social (e.g. social norms) or environmental (e.g. living in an area affected by malaria).

2. Capabilities Versus Social Primary Goods

Of all the possible metrics of justice, philosophers defending the capability approach have most explicitly targeted John Rawls's metric of justice, the "social primary goods" (Nussbaum 2006; Sen 1980, 2009). In his 1979 Tanner lecture entitled "Equality of What?" Sen (1980) argued that "the primary goods approach seems to take little note of the diversity of human beings." He continued:

> If people were basically very similar, then an index of primary goods might be quite a good way of judging advantage. But, in fact, people seem to have very different needs varying with health, longevity, climatic conditions, location, work conditions, temperament, and even body size. … So what is being involved is not merely ignoring a few hard cases, but overlooking very widespread and real differences.
>
> (Sen 1980: 215–16)

A person with a disability, however severe, would not have a claim to additional resources grounded in his impairment under Rawls's two principles of justice. Sen argues that Rawls's difference principle would not justify any redistribution to the disabled on grounds of disability. Rawls's strategy has been to postpone the question of our obligations towards the disabled, and exclude them from the scope of his theory. Rawls certainly does not want to deny our moral duties towards the people that fall outside the scope of his theory, but he thinks that we should first work out a robust and convincing theory of justice for the "normal" cases and only then try to extend it to the "more extreme cases" (Rawls 2001: 176).

Sen's critique in his Tanner lecture, however, was not only about the case of the severely disabled. Sen's more general critique concerned what he saw as the inflexibility of primary goods as a metric of justice. Sen believes that the more general problem with the use of primary goods is that it cannot adequately deal with the pervasive differences

among people. Primary goods, he argues, cannot adequately account for differences among individuals in their abilities to convert these primary goods into what people are able to be and to do in their lives. Primary goods are among the valuable means to pursue one's life plan. But the real opportunities or possibilities that a person has to pursue her own life plan, are not only influenced by the primary goods that she has at her disposal, but also by a range of factors that determine to what extent she can use these primary goods to generate valuable states of being and doing. Hence, Sen claims that we should focus on the extent of substantive freedom that a person effectively has, i.e. her capabilities.

Yet Rawls has defended the exclusion of the disabled from his theory. Rawls stressed, especially in his later work, that in his theory "everyone has physical needs and psychological capacities within the normal range," and therefore he excludes people with severe physical or mental disabilities from the scope of justice as fairness (2001: 170–6). In his earlier work (Rawls 1971), Rawls justified the restriction by arguing that a theory of justice should in any case apply for "normal cases"—if the theory is inconsistent or implausible for such cases, then it will certainly not be an attractive theory for the more challenging cases, such as people with severe disabilities. We could postpone the question of how to treat people with disabilities to one of the later (legislative) stages of the design of the basic structure of society, though, of course, even in his earliest discussions of this Rawls thinks that the final theory of justice must deal adequately with the claims of people whose abilities fall outside the normal range, and that any theory that cannot do so should be rejected on those grounds. In later work Rawls (2001: 176) no longer argued that the case of justice towards the disabled had to be postponed to the legislative phase, but rather that we had to try to extend justice as fairness to include those cases. Rawls has not pursued this task systematically himself, though he has emphasized the role that his conception of the person possessed of the capacities for a sense of justice and a conception of the good plays in justice, and has argued (2001: 176–8) that this conception enables him to deflect accusations of "fetishism" about the primary goods.

A more recent wave of philosophical enquiry has highlighted how complex the comparisons between Rawls's theory of justice and the capability view are (Pogge 2002; Robeyns 2009; Brighouse and Robeyns 2010). One reason is that the capability metric is a general metric of well-being freedom, whereas the social primary goods metric emerges as one element of an integral and complex theory of political justice (rather than social justice more broadly, let alone the even wider category of moral evaluations). Also, Rawls's theory of justice is an ideal theory of justice since it tries to outline the conditions of a completely just (yet "realistic") utopia, which the capability metric does not aspire to do. This means that it is very hard to compare Rawls's work on justice with the philosophical work on the capability approach, since their scope and theoretical aims are not the same (Robeyns 2009, 2011).

Regarding scope, Rawls's theory of justice is limited to (1) the basic structure of society (that is, the set of most important social institutions), (2) to liberal democratic societies rather than also to nondemocratic and illiberal societies, (3) and to the principles of justice insofar as they apply to people in their capacity as citizens. The scope of the capability approach can be summarized as "justice applies everywhere"—that is, it applies to all human beings independently of their country of birth or residence, and not only to social institutions but also to the social ethos and to social practices.

Regarding theoretical aims, one can safely say that most capability scholars tend to disfavor top-down theorizing, and prefer to find out how theory or philosophy can help

us make the actual world, a social institution or a practice more just, rather than to work more abstractly on the principles of justice and their justification. This last difference has been a main point of attention in Sen's more recent work (Sen 2009).

Rawlsians have criticized the capability approach too, and not all of their critiques have been sufficiently rebutted (Pogge 2002; Kelly 2010).

First, the capability approach is claimed to be endorsing a particular comprehensive moral view, which Rawlsians find objectionable. Rawls aims to stay away from a perfectionist account of justice, and the question is whether this is possible for a capability theory of justice. This is an important area of dispute, to which we will return briefly in Section 3.

Another main Rawlsian objection to the capability approach concerns the publicity criterion, which stipulates that the conception of justice must be public and the necessary information to make a claim of injustice must be verifiable by all, and easily accessible. Rawlsians argue that a theory of justice needs a public standard of interpersonal comparisons, as otherwise the obtained principles of justice among citizens with diverse conceptions of the good life will not prove stable. The suggestion is that as capabilities are very hard to measure or assess in such a public fashion, and as they would require very large amounts and difficult sorts of information, the capability approach is unworkable as a theory of justice. Clearly, not everyone would agree that this is a valid complaint. For example, Elizabeth Anderson (2010: 85) has argued that the capabilities metric does meet the publicity criterion, while Richard Arneson (2010: 114) has argued that concerns of justice overrule concerns of publicity: if social justice can only be achieved by relying on measures that violate the publicity criterion, then that is a price worth paying. Finally, it has also been questioned whether the social primary goods metric, which prominently includes opportunities, can itself meet the publicity requirement, in which case the publicity critique would lose much of its force (Robeyns 2009: 409).

3. Which are the Capabilities Relevant for Justice?

A major challenge for a capability theory of justice is the question which capabilities matter. In answering this question, philosophers have had two different notions of justice in mind. One sees the question of justice as a question about truth, sharply distinguished from questions about implementation, justice-enhancing policies, feasibility, and other practical concerns. G.A. Cohen (2008) is an important representative of this line of work. However, very few articles analyzing or defending the capability approach to justice take this line. An exception is Peter Vallentyne, who argues that all functionings should be included when considering issues of justice. According to Vallentyne (2005: 362), "given that any functioning could, under some circumstances, enhance (or otherwise affect) the quality of someone's life, it is a mistake to exclude some functionings from consideration. To do so would leave out something that is relevant for justice." However, Vallentyne adds that this is not to deny that when designing policies, we need to select the most important capabilities. Since virtually all capability theorists implicitly or explicitly understand "justice" as a practical concept, that is, as a concept that will help us in telling what we ought to do or how we ought to shape social institutions, they take some (minimal) feasibility constraints into account. Those working in a more practical line of political philosophy have argued that considerations of justice require that we demarcate morally relevant from morally irrelevant and morally bad capabilities (Nussbaum 2003; Pogge 2002; Pierik and Robeyns 2007). Put differently,

any capability account of justice will have to tell us which capabilities are relevant and which are not for purposes of justice.

Amartya Sen (2004, 2009: 242–3) notoriously has refused to answer this question, claiming that processes of public reasons and democratic deliberation should lead to the selection of relevant capabilities. Yet this "democratic route" to selecting the relevant capabilities requires a specific account of the deliberative processes that are needed, and that specific account has not been provided by Sen. While several capability theorists have debated issues of democratic deliberation in the context of development questions or other policy decisions, within the context of distributive justice this work remains to be done.

The second way to select the relevant capabilities for the purpose of justice is the "criteria route," whereby the criteria that the selected capabilities should meet are proposed and defended. A prominent example of the "criteria route" is Elizabeth Anderson's (1999) theory of democratic equality. Anderson, who aims to develop the outline of a *political* theory of justice (rather than a theory of social justice that encompasses all spheres of life), argues that people should be entitled "to whatever capabilities are necessary to enable them to avoid or escape entanglement in oppressive social relationships" and "to the capabilities necessary for functioning as an equal citizen in a democratic state," without giving a complete list of which capabilities are meeting these criteria (Anderson 1999: 316).

The third way to select the relevant capabilities for the purpose of justice is the "objective-list route." Following Derek Parfit's notion of objective-list theories, being theories claiming that "certain things are good or bad for us, whether or not we want to have the good things, or to avoid the bad things" (Parfit 1984: 493), the capability theorist proposes an objective list of well-being which will be the concern of distributive justice. Richard Arneson (2010) defends this version of the capability approach, which he dubs the "perfectionist capability theory"—without, however, specifying an account of well-being in terms of an objective list.

Martha Nussbaum's (2000, 2006, 2011) minimal theory of justice is the most well-known version of the capabilities approach which relies on an objective list of well-being. Nussbaum's theory of social justice is comprehensive, in the sense that it is not limited to an account of political justice, or to liberal democracies. Rather, her account holds for all human beings, independently of whether they are living in a liberal democratic regime, or of whether they are severely disabled. The main demarcation of Nussbaum's account is that it provides only "a partial and minimal account of social justice" (Nussbaum 2006: 71) by specifying thresholds of a list of capabilities that governments in all nations should guarantee to their citizens. Nussbaum's theory focuses on thresholds, but this does not imply that reaching these thresholds is all that matters for social justice; rather, her theory is partial and simply leaves unaddressed the question what social justice requires once those thresholds are met. Nussbaum's well-known list contains capabilities that are grouped together under ten "central human capabilities": life; bodily health; bodily integrity; senses, imagination and thought; emotions; practical reason; affiliation; other species; play; and control over one's environment (Nussbaum 2006: 76–8; 2011: 33–4).

Nussbaum (2000: 70–7; 2006: 78–81) justifies her list by arguing that each of these capabilities is needed in order for a human life to be "not so impoverished that it is not worthy of the dignity of a human being" (2000: 72). She defends these capabilities as being the moral entitlements of every human being on earth. She formulates the list

at an abstract level and advocates that the translation to implementation and policies should be done at a local level, taking into account local differences. Nussbaum argues that this list can be derived from a Rawlsian overlapping consensus and stresses that her list remains open-ended and always open for revision (Nussbaum 2000: 77), yet other philosophers have taken issue with her claim that this would result in a form of political liberalism, claiming that she is a perfectionist liberal after all (Barclay 2003).

4. Aggregating Capabilities

The selection of relevant capabilities is one major challenge for the capability approach to justice, yet another is the question how to aggregate the different capabilities that are judged to be relevant for issues of justice. If judgments of justice require us to make interpersonal comparisons of people's overall freedom to achieve well-being, don't we need a way to aggregate the value of the different capabilities into an overall value?

In many cases, this will be true. Yet the capability-justice literature is remarkably silent on this question. There are some proposals for aggregating capabilities either using social choice procedures (Chakraborty 1996) or else equating the value of a capability by its contribution to a person's happiness (Schokkaert 2007), but these are made in the context of welfare economics, rather than as part of an attempt to assessing justice. Amartya Sen deliberately refuses to investigate the question how such aggregation should or could be done, since he believes that the quest for aggregation is driven by a concern for complete rankings, yet to his mind striving for complete rankings is a mistake (Sen 2009).

The capability literature thus seems to leave us with empty hands. Yet the problem of aggregating the dimensions of the metric of justice is of course relevant for all multidimensional metrics of justice, including Rawls's social primary goods metric, or Dworkin's resource egalitarianism. Could the capability approach borrow the aggregation mechanisms available in Dworkin's or Rawls's theory?

On one interpretation of Rawls, the aggregation problem is an unsolved problem in his theory of justice, since the social primary goods are noncommensurable. On another interpretation, the social primary goods of basic liberties, opportunities and the social basis of self-respect, will be distributed equally if Rawls's first principle of justice (the principle of equal liberties) is met, implying that for distributive questions the relevant social primary goods reduce to income and wealth (assuming the social basis of self-respect is best taken care of by the realization of the principle of equal liberties). If it is problematic to reduce to only income and wealth the full range of social primary goods, then there is no satisfactory solution to the aggregation problem in Rawls's theory of justice either. So the problem of aggregating would then be a problem for the capability approach to justice, but as well for the social primary goods metric.

Dworkin (1981b) proposes the so-called "envy test" to make comparisons between the "resources" that people hold, which in Dworkin's conceptualization includes not only their material possessions and leisure time, but also their skills, talents and handicaps. If one is willing to take a pill to trade with the place of another person, taking the entire "package deal" of their life, one has a justified complaint that one has not received one's fair share in life. Yet while Dworkin's envy test might be a useful heuristic device, the envy test entails impossible epistemological requirements, and operates against a number of background assumptions that make Dworkin's theory highly idealized (Pierik and Robeyns 2007). Moreover, while it might be a useful device in

micro-situations or for thought experiments, it offers little to political leaders or policy makers who cannot make these envy tests on behalf of all the citizens. Still, the point which capability theorists of justice should take home is that something like a Dworkinian envy test will have to be developed if one wants the capability metric to do the work of a complete theory of distributive justice.

A capability theory of justice that endorses sufficiency as the distributive rule, such as Martha Nussbaum's theory (2000, 2006, 2011), can avoid the problem of aggregation, since such a theory specifies that justice requires only that all people meet a certain threshold-level for each capability. Yet even such a theory has two major problems to solve. First, we will need to know where the thresholds are set, or which person or which foundational principles decide on the thresholds; basically, the normative decisions that confront the selection of relevant capabilities re-emerge here, albeit in a slightly different guise. Second, in an unjust world, the theory of transitional justice or the non-ideal theory of justice will need to tell us which capability to prioritize if not all people are above the thresholds of all relevant capabilities. Should we prioritize education, health, or being able to hold a decent job? Nussbaum (2011: 37–8) has argued that the impossibility to get all people above the thresholds for all capabilities involves a tragic choice, which should prompt us to ask the question how we can work towards a future where this is no longer the case. For Nussbaum, this strategy will be sufficient, since "If the whole list has been wisely crafted and the thresholds set at a reasonable level, there usually will be some answer to that question" (Nussbaum 2011: 38). Yet one could wonder whether this response does solve the problem: many of the one billion most deprived people on earth are below reasonable thresholds on most of the capabilities on Nussbaum's list. Which one should a justice-seeking organization or government prioritize?

5. A Family of Capability Theories of Justice

The capability approach is often wrongly taken to be an egalitarian theory or a theory of social or distributive justice, yet its core doesn't entail more than the two normative propositions stated at the beginning of this chapter. The capability approach specifies what should count for interpersonal evaluations and thus provides *one* important aspect (or "module") of a theory of social or distributive justice, yet more is needed before one can speak of a theory of justice (Robeyns 2011).

Nussbaum (2000, 2006, 2011) offers us a capability theory of justice, but her theory too doesn't amount to a full theory of social justice. Moreover, it would be a mistake to think that there can be only one capability theory of justice; on the contrary, the open nature of the capability approach allows for the development of a family of capability theories of justice. But this prompts the question: what is needed to develop a full capability theory of justice, and which of these aspects have already been developed by capability theorists? Assuming that the capability theory of justice has addressed the issues outlined above—that is, the selection of the relevant capabilities, and the question of aggregation—another number of theoretical choices remain to be made and defended.

First, a theory of justice needs to explain on what basis it justifies its principles or claims of justice. For example, in Rawls's theory of justice the two principles of justice are justified by the thought-experiment of the original position and the more general social contract framework on which this is based. Dworkin's egalitarian justice theory starts from the meta-principle of equal respect and concern, which he then develops in

the principles that the distribution of burdens and benefits should be sensitive to the ambitions that people have but should not reflect the unequal natural endowments with which individuals are born. One could also develop a capability theory of justice arguing that the ultimate driving force is a concern with autonomy or with human dignity, or with human vulnerability, or with a combination of these. If capability scholars want to develop a full theory of justice, they will also need to explain on what bases they will justify their principles or claims. As mentioned earlier, Nussbaum starts from a notion of human dignity, whereas the Senian strand in the capability approach stresses the importance of what people have reason to value, hence an account of public reasoning. However, little work has been done so far to flesh out this embryonic idea of "having reason to value," and it therefore remains unclear whether the capability approach has a solid unified rationale on the basis of which a full account of justice could be developed.

Second, a capability theory of justice needs to take a position on the "distributive rule" that it will endorse: will it argue for plain equality, or for sufficiency, or for prioritarianism, or for some other (mixed) distributive rule? Both Martha Nussbaum's and Elizabeth Anderson's theories are sufficiency accounts (Anderson 1999, 2010; Nussbaum 2006), but from this it does not follow, as one sometimes reads in the secondary literature, that the capability approach entails a sufficiency rule. Sen might have given the (wrong) impression of defending straight equality as a distributive rule, by asking the question "Equality of what?" (Sen 1980), though a careful reading shows that he was merely asking the question "*If* we want to be defending equality of something, then what would that be?" In fact, Sen has remained uncommitted to one single distributive rule, which probably can be explained by the fact that he is averse of building a well-defined theory of justice but, rather, prefers to investigate how real-life unjust situations can be turned into more just situations, even if perfect justice is unattainable (Sen 2009). The capability approach clearly plays a role in Sen's work on justice, since when assessing a situation he will investigate inequalities in people's capabilities and analyze the processes that led to those inequalities. Yet Sen has an eclectic approach to theorizing, and hence other notions and theories (such as human rights or more formal analyses of freedoms from social choice theory) also play a role in his work on justice. The presence and importance of the capability approach in Sen's work is thus undeniable, but should not be seen as the only defining feature.

Third, when developing a capability theory of justice we must decide whether we want it to be an outcome or an opportunity theory, that is, whether we think that we should assess injustices in terms of functionings, or rather in terms of capabilities, or a mixture. At the level of theory and principles, most theorists of justice endorse the view that justice is done if all have equal genuine opportunities, or if all reach a minimal threshold of capability levels. Translated to the capability language, this would imply that at the level of theory and principles, capabilities are the relevant metric of justice, and not functionings. However while most theorists defend opportunities rather than outcomes, the focus on opportunities is not entirely uncontested (e.g. Fleurbaey 2002; Phillips 2004; Wolff and De-Shalit 2007).

Fourth, a capability theory of justice needs to specify where the line between individual and collective responsibility is drawn, or how it will be decided, and by whom, where this line will be drawn. There is a remarkable absence of any discussion about issues of responsibility in the capability literature, in sharp contrast to political philosophy and welfare economics where this is one of the most important lines of debate, certainly since the publication of Ronald Dworkin's (1981b) work on justice and equality which

led to what Elizabeth Anderson (1999) has called "luck-egalitarianism." An exception is Peter Vallentyne (2005: 365), who has argued that the relevant metric of justice is "brute luck capabilities," being the capabilities that one has as a matter of brute luck only; those that one has as a matter of option luck (that is, due to one's own choices) are not a concern of justice. While Vallentyne endorses a strict separation between theorizing about justice and policy and institutional design, the question of responsibility also has important effects for the more practical approach to justice. Indeed, whether one wants to discuss it explicitly or not, *any* concrete capability-based policy proposal can be analyzed in terms of the division between personal and collective responsibility; but this terminology is largely absent from the capability literature. There might be plausible explanations for why this issue is not discussed in the capability literature, but if a capability theory of justice wants to be applicable to questions of justice, then it cannot but confront the question about the just division between personal and collective responsibility (Pierik and Robeyns 2007: 148–9).

This brings us to a related issue: a theory of justice generally specifies rights, but also duties. However, capability theorists have remained largely silent on the question of who should bear the duties for the expansion of the selected capabilities. Nussbaum passionately advocates that all people all over the world should be entitled, as a matter of justice, to threshold levels of all the capabilities on her list; but apart from mentioning that it is the governments' duties to guarantee these entitlements (2006: 70), she remains silent on the question who precisely should bear the burdens and responsibilities for realizing these capabilities. Yet as Onora O'Neill (1996: ch. 5) has argued, questions of obligations and responsibilities should be central to any account of justice.

This short and presumably incomplete list of the "modules" which a complete capability theory of justice would need to entail, makes clear that a capability theory of justice is theoretically much more demanding than the basic presupposition of the capability approach to distributive justice, namely its claim that "functionings" and "capabilities" are the best metrics of justice. While much has been written on the capability approach in recent years, by an increasing number of scholars, including philosophers, much of the philosophical work needed for turning the open-ended capability approach into a specific theory of justice remains to be done.

Related Topics

Contractualism and Political Liberalism, Utilitarianism and Consequentialism, Perfectionism, Pluralism, Luck Egalitarianism, The Difference Principle, Needs and Distributive Justice, Global Justice and Politics, Equality

References

Anderson, E. (1999) "What is the Point of Equality?" *Ethics* 109(2): 287–337.

Anderson, E. (2010) "Justifying the capabilities approach to justice," in H. Brighouse and I. Robeyns *Measuring Justice: Primary Goods and Capabilities*, Cambridge: Cambridge University Press.

Arneson, R. (2010) "Two cheers for capabilities," in H. Brighouse and I. Robeyns *Measuring Justice: Primary Goods and Capabilities*, Cambridge: Cambridge University Press.

Barclay, L. (2003) "What Kind of Liberal is Martha Nussbaum?", *SATS: Nordic Journal of Philosophy* 4(3): 5–24.

Brighouse, H. and I. Robeyns (eds.) (2010) *Measuring Justice: Primary Goods and Capabilities*, Cambridge: Cambridge University Press.

Chakraborty, A. (1996) "On the Possibility of a Weighting System for Functionings," *Indian Economic Review* 31: 241–50.
Claassen, R. (2009) "Institutional Pluralism and the Limits of the Market," *Politics, Philosophy and Economics* 8(4): 420–47.
Cohen, G.A. (2008) *Rescuing Justice and Equality*, Cambridge, MA: Harvard University Press.
Dworkin, R. (1981a) "What is Equality? Part 1: Equality of Welfare," *Philosophy and Public Affairs* 10: 198–246.
Dworkin, R. (1981b) "What is Equality? Part 2: Equality of Resources," *Philosophy and Public Affairs* 10: 283–345.
Fleurbaey, M. (2002) "Development, Capabilities and Freedom," *Studies in Comparative International Development* 37: 71–7.
Kelly, E. (2010) "Equal opportunity, unequal capability," in H. Brighouse and I. Robeyns *Measuring Justice: Primary Goods and Capabilities*, Cambridge: Cambridge University Press.
Nussbaum, M. (2000) *Women and Human Development: The Capabilities Approach*, Cambridge: Cambridge University Press.
Nussbaum, M. (2003) "Capabilities as Fundamental Entitlements: Sen and Social Justice," *Feminist Economics* 9(2/3): 33–59.
Nussbaum, M. (2006) *Frontiers of Justice: Disability, Nationality, Species Membership*, Cambridge, MA: Harvard University Press.
Nussbaum, M. (2011) *Creating Capabilities. The Human Development Approach*, Cambridge, MA: Harvard University Press.
O'Neill, O. (1996) *Towards Justice and Virtue*, Cambridge: Cambridge University Press.
Parfit, D. (1984) *Reasons and Persons*, Oxford: Oxford University Press.
Phillips, A. (2004) "Defending Equality of Outcome," *Journal of Political Philosophy* 12(1): 1–19.
Pierik, R. and I. Robeyns (2007) "Resources versus Capabilities: Social Endowments in Egalitarian Theory," *Political Studies* 55(1): 133–52.
Pogge, T. (2002) "Can the Capability Approach be Justified?" *Philosophical Topics* 30(2): 167–228.
Rawls, J. (1971) *A Theory of Justice*, Cambridge, MA: Harvard University Press.
Rawls, J. (2001) *Justice as Fairness: A Restatement*, Cambridge, MA: Harvard University Press.
Robeyns, I. (2006) "Three Models of Education: Rights, Capabilities and Human Capital," *Theory and Research in Education* 4(1): 69–84.
Robeyns, I. (2009) "Justice as fairness and the capability approach," in Kaushik Basu and Ravi Kanbur (eds.), *Arguments for a Better World. Essays for Amartya Sen's 75th Birthday*, Oxford: Oxford University Press.
Robeyns, I. (2011) "The Capability Approach," *Stanford Encyclopedia of Philosophy*, online at http://plato.stanford.edu/entries/capability-approach/.
Schokkaert, E. (2007) "Capabilities and Satisfaction with Life," *Journal of Human Development* 8(3): 415–30.
Sen, A. (1980) "Equality of What?" in S. McMurrin (ed.), *Tanner Lectures on Human Values*, Cambridge: Cambridge University Press.
Sen, A. (1989) "Development as Capability Expansion," *Journal of Development Planning*, 19: 41–58.
Sen, A. (1992) *Inequality Re-examined*, Oxford: Clarendon Press.
Sen, A. (1993) "Capability and well-being," in Nussbaum and Sen (eds.), *The Quality of Life*, Oxford: Clarendon Press.
Sen, A. (2004) "Capabilities, Lists and Public Reasons: Continuing the Conversation," *Feminist Economics* 10(3): 77–80.
Sen, A. (2009) *The Idea of Justice*, London: Allen Lane.
Vallentyne, P. (2005) "Debate: Capabilities versus Opportunities for Wellbeing," *Journal of Political Philosophy* 13: 359–71.
Wolff, J. and A. de-Shalit (2007) *Disadvantage*, Oxford: Oxford University Press.

41
INTERGENERATIONAL DISTRIBUTIVE JUSTICE

Clark Wolf

1. Intergenerational Justice in Theory and Practice

Our present actions and institutions can influence the lives of people who will live in the distant future. These people's lives might not overlap our own, and they have no way to reward us for benefits we may bestow or to punish us for harms we may inflict upon them. But because we can sometimes gain advantages for ourselves at cost to future people, the relationship between people who live at different times and in different generations raises questions of justice: is it just to gain advantages for ourselves in the present, if we know that our actions impose serious risk of harm for people who don't yet exist? Can we make sense of the idea that present actions might, under some circumstances, violate the rights of people who live in the future? These concerns are not hypothetical: many people worry that our present use of energy and environmental resources could harm our descendants or leave them impoverished. When our activities benefit us as members of the present generation but also impose costs and risks on people who will live in the distant future, it is appropriate to ask whether we may justly discount their interests and favor our own. This chapter will not articulate a theory of intergenerational distributive justice. Instead, it will examine alternative approaches one might take and tools one might use to develop such a theory.

2. Rights, Community Standards, and Intergenerational Saving

Theorists frequently classify theories of justice as *Libertarian*, *Liberal*, or *Communitarian*. *Libertarians* often frame the problem in terms of underlying rights to liberty and property. Some libertarians hold that the only obligations that can legitimately be enforced are negative obligations, and obligations people freely accept when making contracts or agreements with others. But since future persons do not exist, they cannot have any negative or contractual rights. Such assumptions quickly lead some to conclude that libertarian justice cannot extend across generations (Beckerman and Pasek 2001; Beckerman 2003). Of course, acceptance of this view does not necessarily imply lack of concern for future generations, nor does it imply that they will be badly off. Some advocates argue that the market institutions libertarians favor are most likely to provide benefits for future generations (Gauthier 1986; Beckerman 2004; Cowen 2007).

Communitarians hold that norms of justice are local, and apply within the communities in which they arise. But distant future people do not seem to be members of our

present community. Consequently, some communitarians argue that norms of justice do not apply to future generations. In this spirit, de-Shalit (1995) argues that our obligations to future people are obligations of *humanity* but not of justice. Those who find the communitarian ideal appealing but who are unable to accept this implication might instead consider the sense in which future—even distant future—people, may be part of our community even though they do not presently exist. In this spirit, Burke writes that the social "ends" of society cannot be obtained or created in a single generation, and that society should therefore be considered a contractual partnership "not only between those who are living, but between those who are living, those who are dead and those who are to be born" (Burke 1790/2001 paragraph 165).

There is no reason in principle why either a libertarian or a communitarian account of justice could not include provisions to protect distant future generations. In practice, however, representative libertarian and communitarian views have not usually been developed with distant future generations in mind.

Liberal theories of justice place a high value on individual rights and liberties, but also include a role for *distributive* justice, specifying the way in which the burdens and benefits of social cooperation or interaction should be carried by different persons. Liberal theorists have often framed the problem of intergenerational justice as a problem of *saving* or *investment* in the future. For example, John Rawls (1971, 1993, 1999) argues that the central question for a theory of intergenerational justice is to select a just rate of saving that specifies what resources present generations should preserve for the future. It is easy to see why this might seem an appropriate strategy: future generations will have at their disposal whatever we leave for them to use, so our obligations to them might reasonably be framed as a duty concerning what we leave behind. Just treatment of future generations may require that we divide the world's resources into fair intergenerational shares, and avoid using more than our share. But this strategy faces difficulties: We don't know how many people there will ever be, so it is difficult to divide the world into shares. Worse, the total number of present and future people is probably very large, and if we were to try to allocate each present and future person a fair share of the present wealth of the world, each share might be infinitesimal. Fortunately, many resources grow or recover over time, and others are intergenerationally durable, unlikely to be depleted as they are used. Perhaps intergenerational distributive justice won't involve the indefinite subdivision of a finite-sized pie, but instead the non-destructive employment of productive resources that can be used by one generation and then passed on to the next.

3. Material Sustainability: Modeling Resource Use Over Time

The concept of intergenerational justice is often identified with the thought that environmental resources should be used *sustainably* so that we accommodate the needs of future generations. (Barry 1989; Solow 1974, 1993) By "sustainable use," people often mean non-depletionary use: using a portion of the resource and maintaining (or investing) the balance, so that subsequent generations will have an undiminished supply. One of the simplest and clearest models of non-depletionary use is articulated in the work of Partha Dasgupta (1974a, 1974b).

Dasgupta asks that we consider a community that starts with a finite set amount of some essential and enduringly valuable renewable resource K, the quantity of which at time t is K_t. Each generation faces a choice about how much to consume now, and the choice made will determine how much will be available for subsequent generations.

Since renewable resources grow or recover after use, we can let α be the rate at which saved resources grow over time from one period to the next. Suppose the first generation to be represented in our model possesses a finite quantity K_t of resource K at time t, and consumes C_t. Then the amount of the resource available for the subsequent period at time t + 1 will be:

$$K_{t+1} = \alpha(K_t - C_t) \tag{1}$$

That is, people in the first period start with K_t but consume C_t leaving the remainder, $(K_t - C_t)$. We can represent this value, $(K_t - C_t)$, as the rate of investment, since it will grow from one period to the next at a rate of α, leaving $\alpha(K_t - C_t)$ available in period t + 1. In the case of overstressed environmental resources, investment of this kind might involve forgoing consumption to let natural stocks recover from present use. For example, we might lower the rate of present fishing harvest to allow fish stocks naturally to grow back, in which case α represents the rate at which these stocks recover.

While economic models typically assume that resources grow (or recover) over time, this will not be true for all resources. If $\alpha = 1$, then the amount of resource left over for the next generation will diminish at the rate that it is consumed by earlier generations. In that case, $K_{t+1} = (K_t - C_t)$ and the amount of resource available will continually decrease over time. If $\alpha < 1$ then the resource decays over time and the first generations might need to consume it quickly before it spoils. In the worst case, if $\alpha = 0$, then there is not even time to consume the resource before it spoils. But as long as $\alpha > 1$, then whatever is saved by one generation grows for the benefit of the following generation. We might suppose that each generation would make a decision about how much to consume and how much to save based on the value of α, the value they place on present resource consumption, and the value they place on consumption by future generations.

Over time and across generations, saving and consumption can be represented with a simple accumulation equation:

$$\begin{aligned} &K_{t+1} = \alpha(K_t - C_t) \\ &K_t \geq C_t \geq 0 \quad \text{for } t = 0, 1, 2, 3, \ldots \\ &K_0 \text{ given.} \end{aligned} \tag{2}$$

After the first period, the choice for subsequent periods will be constrained within the limit presented by the saving done previously. This allows us to refer to intertemporal allocation or consumption sequences, so that the sequence $(C_t) = (C_0, C_1, C_1, \ldots C_n)$ will include the consumption allocation for each period. The set of *feasible* intertemporal allocations is the set of sequences C_t that satisfy equation (2) above. If people invest more in earlier periods, there will be more to consume later. If they invest less, then the future will be poorer.

Is there a minimum saving rate required by justice, such that saving less would violate obligations to future people? If people in the earlier generations gratuitously consume all of the available resource, then people in later generations will be left with nothing at all. On the other hand, if people save too much they might needlessly impoverish themselves for the sake of their rich descendants. On this model, a theory of intergenerational justice would, among other things, provide a principled method for choosing an appropriate rate of saving.

It might be easier to identify undesirable saving plans than to identify a unique optimal plan. Even if there *is* no unique optimum, it might be possible to rule out the worst alternatives. In this spirit, Dasgupta suggests that we should eliminate those schemes that tend toward zero consumption, and that it would "not be very just" for early generations to leave nothing for later ones (Dasgupta 1974a: 415). He suggests that justice requires that intergenerational consumption and saving should, in this sense, at least be sustainable over time. But the possibility of sustainable production and consumption will obviously depend on the sizes of α and K_t. Suppose we define "sustainability" as the condition that $K_{t+1} = K_t$ for all times t. This means that the amount of K available to later generations is no less than the amount available to earlier generations. In order to obtain this condition, we must have $K_t = \alpha(K_t - C_t)$. This gives us a condition on C that will insure nondecreasing resource availability over time and across generations:

$$C_t \leq (K_t - K_t/\alpha) \tag{3}$$

Where resources fit the assumptions of the model, savings plans that consistently consume faster than the sustainable rate will tend toward zero consumption in the long run. For this reason, it has seemed to some that the sustainable consumption rate is a minimal requirement of intergenerational justice.

However, we should not be too quick to identify a sustainability requirement as a necessary condition for intergenerational justice. Dasgupta's model focuses on resources that meet people's needs, but non-depletionary use might be neither necessary nor sufficient for meeting future needs. In the case of non-renewable resources such as oil and coal, sustainable use would mean no use at all. In such cases we might consider tradeoffs among different kinds of resources. For example, when we use non-renewable resources such as oil (where $\alpha = 1$) we can be sure the total quantity of these resources that will be available to future generations will be less. On the other hand, our present consumption might support the development of substitute technologies that might take the place of oil. What is a "resource" for one generation might no longer be needed by later generations, if new technologies provide an alternative. We might know little about future needs and preferences. Future people might not need or value the resources we save for them. While it would be unreasonable to assume that proximate future people will be so different from us that they might not need a breathable atmosphere or other environmental basics, the things we might choose to save for future generations could reflect a systematic misunderstanding of their needs and values.

There are additional reasons to question the simple view of intergenerational justice as non-depletionary resource use: merely sustainable consumption may, in some circumstances, be *insufficient* to meet basic needs if population size is increasing over time. And sustainability might not be *necessary* for meeting human needs if population size were to diminish. If the goal of sustainable resource use is to insure that future generations are as well-off as present generations (Solow 1993), it will be necessary to take into account expected rates of population growth, as well as our uncertainty about future needs. Because of this, an account of non-depletionary resource use will not be a full theory of intergenerational justice, even if it might be part of such a theory (Wolf 2010).

4. The Model of Preferences and Needs: Intergenerational Social Choice

An alternative strategy to represent intergenerational choice focuses on preferences and needs. Many theorists have modeled intergenerational distribution by considering the preferences of a hypothetical chooser who is ignorant of the position she or he will occupy in society. This strategy has been suggested and developed by a number of different writers, including Harsanyi (1955), Vickrey (1960), Rawls (1971), Arrow (1973), Mueller (1974) and Dasgupta and Heal (1979). Ignorance of one's own identity or position in society is understood as a way to extend equal consideration to different members of society. While there are difficulties in modeling intergenerational social choice in this way, it is useful to consider how far this notion can be extended to a theory of intergenerational justice.

Consider a simple case involving two generations, or to simplify, two people a and b who live at different times and whose consumption is represented by Ca and Cb, respectively. The impartial chooser is assigned the task to select a distributional scheme that will determine the distribution of goods between a and b. In the two-person-two-generation case, the choice will be an ordered pair <Ca,Cb> where Ca is the amount of resource available for consumption in period a, and Cb is the amount available in period b. We can specify the range of feasible distributions by graphing them on a coordinate system, where Ca and Cb are the two dimensions (Figure 1). The range of feasible (or possible) distributional alternatives can be represented as the points that lie within a line that is concave with respect to the origin, FF, which represents the set of ordered pairs that *completely* distribute all of the available resources. FF is sometimes called a production possibility frontier (PPF).

On Figure 1, the preferences of the impartial chooser can be represented as *indifference curves* (a_1–a_1) and (a_2–a_2), which provide a ranking of the feasible distributions. Feasible distributions include all of the ordered pairs that lie within the PPF. Each indifference curve is a set of points (ordered pairs) that are regarded as *equally good* (or equally bad) by the choosing agent. In this case, the indifference curves represent the preferences of an impartial chooser who does not know when she will exist, and hence does not know which period's consumption will be *her own* consumption. She will be equally

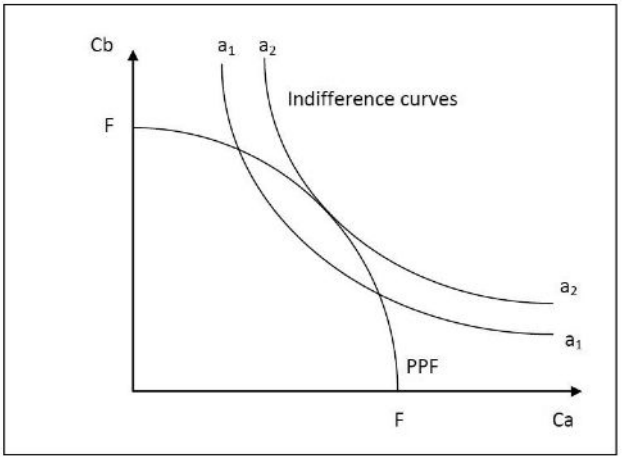


Figure 1 Two-period production possibility frontier with indifference curves

concerned with both. Alternately, Ca and Cb could represent the resources available to each of two *generations*, where the impartial chooser does not know to which generation she herself belongs. The problem of intergenerational distribution, then, is reframed as a question about the shape of the indifference curves of an impartial chooser.

Dennis Mueller (1974) develops an account of intergenerational distributive justice by positing alternative possibilities for the indifference curves of an impartial chooser like the one posited in Harsanyi (1955) and Rawls (1971). The three diagrams in Figure 2 illustrate alternatives considered by Mueller. As in the previous diagram, the two axes represent the amount available for consumption in the two periods, and the curved line is a PPF. Notice that the PPF is not symmetrical with respect to the 45 degree ray extending from the origin. This models the thought that saving in "period a" (reflected in a lower value on the horizontal axis) can be expected to increase resource availability later in "period b" (vertical axis) as resources grow over time. According to Mueller, intergenerational justice is reflected in the choice of an impartial chooser charged to consider whether she would be willing to be worse off if she were to live in an early generation for the sake of being better-off if she were to live in a later one.

Figure 2-A shows an indifference curve for a risk-neutral chooser who aims to maximize total resource availability regardless of who will enjoy it when it is available. For such an individual, the indifference function a–a is a straight line which will touch the PPF at point A, quite close to the vertical axis. This point represents an intergenerational distribution that provides much more for the later period than the earlier one, while maximizing total resource availability during the two periods. Figure 2-B represents the values of a Rawlsian maximin chooser, who wishes to maximize the consumption of the worst-off generation. Such an individual will have indifference curves b–b parallel to the axes, with a sharp 90 degree angle on the 45 degree ray extending from the origin. This angular indifference function will select point B where the PPF intersects the 45 degree ray. This has the effect of equalizing consumption between the two generations, but would prevent tradeoffs that would involve relatively little cost to the worse-off earlier generation and which might have provided great benefits for those who will live later. Figure 2-C illustrates the preferences of a normally risk-averse chooser,

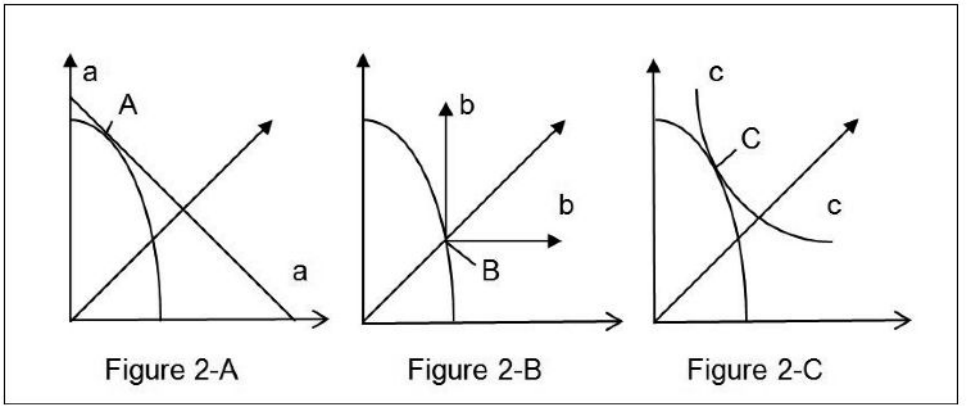


Figure 2 Dennis Mueller, alternative indifference functions for two-period intergenerational distribution

interested to avoid deprivation but willing to sacrifice consumption in an earlier period when it would provide great benefits later. Indifference curves c–c pick out a point on the PPF that is between points A and B.

Mueller urges that the risk-neutral choice sacrifices too much consumption in generation a, while the maximin choice forgoes benefits for generation b that could be achieved at minimal cost to those in generation a. According to Mueller, it is quite plausible to think that an agent who is ignorant of her own identity would be risk averse—and unwilling to accept outcomes that involve too much deprivation—but willing to trade greater consumption in the future at the cost of lesser-but-adequate consumption in the earlier period. Following Harsanyi (1955), he takes this as an indication that such a preference structure should be a strong candidate for a theory of intergenerational justice.

5. Evaluating Intergenerational Social Choice

While the model above includes only two generations, it can be extended to include an indefinite number. But it is worthwhile to consider some of the objections that can be raised against it, as well as the limitations inherent in the methodology. One concern involves the indifference curves we might impute to the impartial chooser, and social choice theory methodology involved in their use: whose indifference curves are these, and what is our real basis for choosing among the alternative possibilities? One concern is that they may simply reflect the arbitrary preferences or values of the modeler. In that case we may have no reason to regard one set of proposed indifference curves as more appropriate or as preferable from the moral point of view. Some theorists are skeptical about indifference curves that do not expressly represent the preferences of some actual person, and urge that the exercise to represent a social welfare function in this way is simply a sneaky way for social theorists to represent their own private preferences and values as if they represented the impersonal public good (Montgomery 1999: 49).

The objection is an important one, and it is crucial for social theorists to guard against the temptation to represent private values as if they were public. But raising the objection is not sufficient if objectors do not take the time to evaluate the reasons given for structuring social choices. Those who defend models like those of Mueller and Rawls urge that they may avoid inappropriately attributing their private preferences to the impartial observer if they articulate good reasons that explain and justify the choices that are prescribed. The reasons themselves must be considered and evaluated. If the reasons aren't adequate, then the model must fail, but it will not be sufficient simply to dismiss all such models as public imposition of private values.

Another common objection involves the use of "consumption" as the object to be distributed. Many have argued that the currency of distributive justice should be something other than consumption. For example, capability theorists argue that distributive justice models should focus on capabilities instead of on goods—on what people can do or be, and not on what they can consume (Sen 2010; Nussbaum 2000, 2011). Others urge that distributive models should focus on a complex basket including various different kinds of rights and goods (Rawls 1971), or on opportunities (Arneson 1989). As a first thought concerning this objection, it is worth noting that the models employed above do not specify the value to be promoted or distributed. It would be perfectly consistent to specify that the "production frontier" in Mueller's model might represent capability or opportunity, or whatever one takes distributive justice to distribute. There

need not be any special association with the production or consumption of physical commodities.

On the other hand, any plausible theory of justice must take seriously the availability of at least some key resources. As Sen (2010) and Nussbaum (2011) have insisted, capabilities cannot effectively be exercised if people are destitute or hungry, or if they lack access to basic physical necessities. Thus while a capabilities approach will not focus on goods as having primary significance, they must be attentive to the distribution of necessary goods that make possible the exercise of fundamental capacities. The significance of fundamental needs can easily be incorporated into a social choice model: for example, one might suppose that the impartial chooser would employ a maximin strategy up to the point where basic needs have been met, and that from that point on intergenerational tradeoffs would be more acceptable.

But even if Mueller's social choice model can be defended against these charges, there is another that could be more damaging: the methodology employed might precipitously *rule out* certain kinds of theories of intergenerational distributive justice. Some otherwise reasonable social preference orderings cannot be neatly reflected as smooth indifference curves covering the space in question. To show this, it is sufficient to identify one plausible candidate theory of intergenerational distribution that *cannot* be represented in such a model.

One such model is a *basic needs* model, which places priority on the goal to minimize the number of people who suffer deprivation with respect to basic needs. Suppose there is some minimal level of provision P, such that anyone who has less than P will be severely deprived. We might suppose that an impartial chooser would be more risk averse about alternatives that involve extreme deprivation in one of the two periods, and would become less risk averse when such fundamental needs are not at stake. Indifference curves reflecting this preference might be coextensive with the maximin curves until a minimal level of provision has been reached. These are represented by curve m_1–M–m_1 in Figure 3. After that point, they might curve more gently, like line a_1–a_1

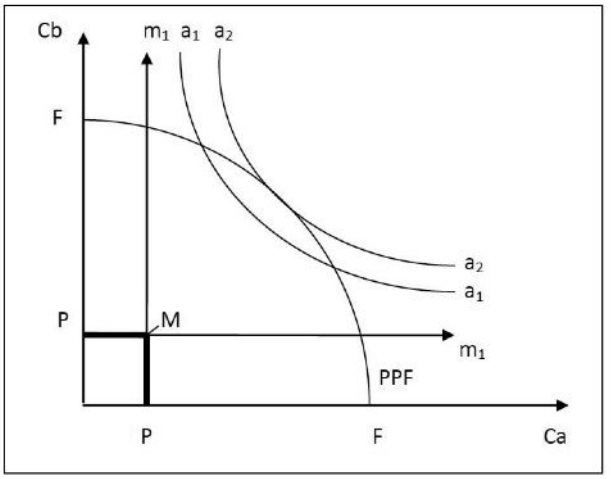


Figure 3 Two-period production possibility frontier with indifference curves and including a minimal necessary level of provision, PMP

reflecting the willingness to trade greater consumption in one period for lesser-but-adequate provision in another.

The view that intergenerational social choice should prioritize the goal to minimize deprivation with respect to basic needs is sufficiently plausible that it has sometimes been defended (Rawls 1993; Wolf 2009). It is not unreasonable to suppose that an impartial chooser would be more risk averse concerning outcomes that leave some with basic needs unmet. However, such a chooser might also be less concerned about other intergenerational inequalities, willing to sacrifice in one period for the chance of greater benefits in another after needs have been met. A view that treats basic need provision as lexically prior to other goals is simply a limiting case for views of this type. In the spirit of such a proposal, we might identify principles that define such a choice function. Call this view *triage prioritarianism*:

Triage Prioritarianism:

1) *Triage Principle:* Minimize the number of people who lack P (the minimal level).
2) *Worst-Off Prioritarianism:* Up to the point where everyone has (at least) P, outcomes are better when they are better for those who are worst off.

This is a *triage* strategy, because within square PMP it prescribes triage tradeoffs: if the choice is between (i) bringing both persons higher but neither as high as the minimal level P, or (ii) bringing one person up to level P while ignoring the person who lives in the other period, it recommends the latter. There is nothing intrinsically irrational about a "triage" strategy with respect to basic needs provision, which recommends minimizing the *number of people* who possess less than P. Such a triage strategy is fully consistent with the additional principle that once the number of people lacking P has been minimized, we should give prioritarian consideration to those who are worse off. By themselves, these two principles provide a complete ordering of alternative distributions that lie on or within PMP, but are consistent with a variety of different principles for ordering outcomes that lie outside PMP.

Since P represents the minimal level of provision, a person who aims to minimize the number of people who lack P would prefer all points on line PMP to points that lie within the boundary of line PMP. Prioritarian concern for the worse off implies that points on PMP that are closer to M are universally preferred to those that are further from M. Such a person will be indifferent between points on PMP that are equidistant from either axis, since they will be mirror images of the same distribution. There is no way to draw smooth, continuous, differentiable indifference curves on this diagram that will describe the values of the "triage" chooser described here. Perhaps this reflects the limitations of the modeling method, and not a problem with the principles under consideration. The triage strategy frames the choice around different values from the ones represented in the model. Since standard social choice models cannot represent conceptions of justice like the triage model, the use of such models may precipitously and inappropriately constrain the range of theoretical alternatives.

6. Justice between Persons not "Generations?"

There is another serious objection to models like the one discussed above: when the axes represent alternative "generations," as they are often specified, such models assume

that *time* is relevant from the perspective of justice, and that "generations," understood as *groups* of people who live at different times, are the relevant subject of distributive justice. We might instead employ a model that represents distributive justice as a matter of distribution among individuals who live at different times, instead of among generations. To see why this is different and how it might be done, note that we can arrange human generations into an array, where each generation, or birth cohort, is represented as a collection of individuals:

$$
\begin{array}{llll}
U_{1,1} & U_{1,2} & U_{1,3}\ldots & U_{1,n} \\
U_{2,1} & U_{2,2} & U_{2,3}\ldots & U_{2,n} \\
U_{3,1} & U_{3,2} & U_{3,3}\ldots & U_{3,n} \\
. & . & . & . \\
. & . & . & . \\
. & . & . & . \\
U_{\delta,1} & U_{\delta,2} & U_{\delta,3}\ldots & U_{\delta,m}
\end{array}
$$

Each row of this array includes individuals, not goods. Since economists typically identify *utility levels* as the relevant measure of individual well-being, we can use 'U' to represent the individuals listed. Each element of this matrix could then represent the utility level of some person, indexed by generation, and by number within that generation, so that $U_{3,6}$ represents the level of well-being enjoyed by the 6th person in the 3rd generation. Utilitarian distribution, however, is only one among many different possibilities. For that reason, the values included in this matrix should be understood to represent *whatever it is that distributive justice is understood to distribute*. Alternative accounts of the *currency* of distributive justice will involve different interpretations, but it is not necessary to resolve that issue here.

We might stipulate that the rows in this array represent generations, G1, G2, G3, … Gn …:

$$
\begin{array}{llll}
G_1 = (U_{1,1} & U_{1,2} & U_{1,3}\ldots U_{1n}) \\
G_2 = (U_{2,1} & U_{2,2'} & U_{2,3}\ldots U_{2n}) \\
G_3 = (U_{3,1} & U_{3,2} & U_{3,3}\ldots U_{3n}) \\
. & . & . & . \\
. & . & . & . \\
. & . & . & . \\
G_\delta = (U_{\delta,1'} & U_{\delta,2} & U_{\delta,3}\ldots, U_{\delta,m})
\end{array}
$$

But how should generations be distinguished? In principle $U_{1,1}$ could represent any individual person, past or present. But for our purposes here, we can arbitrarily stipulate that the first generation to be modeled includes everyone who was alive at 12:00 am on January 1 of the year 2000, arranged along the row according to birth order with the oldest listed first. We could further stipulate that each subsequent generation includes

the *next* 7 billion people to come into existence, so by this stipulation n = 7 billion. Thus the first generation includes all of the (roughly) 6.1 billion people alive at the specified time, plus the next 0.9 billion people born after 2000. While these stipulations are intentionally arbitrary, they make successive generations equinumerous. They also have the implication that new generations arrive on the scene more quickly when global fertility rates are high, and more slowly when fertility rates are low. Since we do not know the total number of human generations so defined, I have identified the last generation as Gδ. In this way, we can refer to this last generation (and to each of the intervening generations) without knowing its number.

This matrix, then, represents all of the people now in existence, everyone who has died since the year 2000, and every person who will ever exist in the future. It follows Dasgupta and Heal (1979) in the assumption that the total number of persons who will ever exist is finite. Not everyone regards this assumption to be obvious, but perhaps it is likely. The last generation might disappear when the sun eventually explodes to destroy the earth, or when the universe finally reaches an entropy equilibrium. All generations except the last will be equinumerous. I have represented the number of the last generation as m.

With such an array in mind, we can again consider alternative ways to model intergenerational distributive justice. For example, we *could* define a function that measures tradeoffs between the different "generations" (G1 … Gn) like Mueller's model, discussed above. But this is now revealed to be a peculiar way to model intergenerational relations, since the way generations have been defined is entirely arbitrary. Since other methods we might use to individuate generations would involve similarly arbitrary assumptions, the notion that intergenerational distributive justice should take as its subject alternative allocations between *generations* should now seem arbitrary. We should instead consider the allocation of burdens and benefits among the individuals who comprise different generations, not between the generations themselves. In developing a theory, we might specify an arbitrary array of generations like the one above, and then evaluate alternative principles of distributive justice based on the implications they would have for the individuals represented. Because we do not know the total number of people who will ever exist, we do not know many particular features of the array that applies to our world. In spite of this, we can still evaluate alternative principles and policies in terms of their consequences for *any* of the alternative arrays that might apply. Since our choices could *change* the number of persons who will exist in the future, we might need to compare alternative policies that involve different numbers of future generations. Comparisons that involve different numbers of people and generations will raise special problems: it is difficult to compare the benefits or costs a future person could experience under different alternatives if that person would not have existed at all in one of them. Different writers have developed alternative methods for making such comparisons, but none of them are uncontroversial (Broome 2004; Mulgan 2006; Wolf 2004, 2009). A full theory of intergenerational justice would need to include principles for making such comparisons.

While it might not be possible to identify a single intergenerational/interpersonal distribution as optimal from the perspective of justice, we might once again endeavor to rule out certain kinds of intergenerational tradeoff as *prima facie* unjust. For example, in the case of justice among contemporaries, we typically regard it to be *prima facie* unjust to sacrifice the basic needs of some people in order to satisfy the non-need wants of others. Were we to adopt such a principle in the intergenerational case, we could formulate it as an action guide for members of the present generation, whichever generation

that may be at the time when the principle is employed. Such a principle might have implications similar to the *triage prioritarian* view considered earlier, and might look something like this:

> Needs Principle: Present actions, policies, and institutions are *prima facie* unjust if their effect is to undermine the ability of future people to meet their basic needs, in order to satisfy the non-need wants of present and temporally proximate people.

Except for the focus on persons instead of generations, this principle is practically identical to the famous *Brundtland Definition of Sustainability*, which states that institutions are sustainable when they "meet … the needs of the present generation without compromising the ability of future generations to meet their own needs" (WCED 1987: 43).

7. Conclusion

The development of a theory of intergenerational distributive justice would require comparison and evaluation of alternative principles and their respective grounds for support. But in evaluating the alternatives, it is important also to keep alive the possibility that no theory of intergenerational distributive justice will be fully satisfactory. As Parfit (1984) has shown, future generations raise special problems and paradoxes for standard moral and political theories, and there might be no theory that fully satisfies all of the various requirements we might wish to impose. For some, this has seemed a good argument for skepticism about intergenerational justice (Beckerman and Pasek 2001; de-Shalit 1995; Gauthier 1986) Others have not given up on the attempt to balance and qualify competing principles and interests in the effort to articulate a theory that appropriately represents our obligations to future generations.

Related Topics

Liberalism, Environmentalism, Contractarianism, Contractualism and Political Liberalism, Utilitarianism and Consequentialism, Left Libertarianism, Libertarianism, Needs and Distributive Justice, The Capability Approach (and Social Justice), Social Choice Theory

References

Arneson, R. (1989) "Equality and Equal Opportunity for Welfare," *Philosophical Studies*: 56: 77–93.
Arrow, K. (1973) "Rawls' Principle of Just Savings," *Swedish Journal of Economics* 75: 232–5.
Barry, B. (1989) *Democracy, Power, and Justice*, Oxford: Oxford University Press.
Beckerman, W. (2003) *A Poverty of Reasons*, Oakland, CA: The Independent Institute.
—— (2004) *A Poverty of Reasons: Sustainable Development and Economic Growth*, Oakland, CA: The Independent Institute.
—— and J. Pasek (2001) *Justice, Posterity, and the Environment*, New York: Oxford University Press.
Broome, J. (2004) *Weighing Lives*, Oxford: Oxford University Press.
Burke, E. (1790) [2001] *Reflections on the Revolution in France*, Stanford, CA: Stanford University Press.
Cowen, T. (2007) "Caring about the Distant Future: Why It Matters and What it Means," *University of Chicago Law Review* 74: 5–40.
Dasgupta, P. (1974a) "On Some Alternative Criteria for Justice between Generations," *Journal of Public Economics* 3: 405–23.

—— (1974b) "On Some Problems Arising from Professor Rawls' Conception of Distributive Justice," *Theory and Decision* 4: 325–44.
—— and G. Heal (1979) *Economic Theory and Exhaustible Resources*, New York: Cambridge University Press.
de-Shalit, A. (1995) *Why Posterity Matters*, New York: Routledge.
Gauthier, D. (1986) *Morals by Agreement*, New York: Oxford University Press.
Harsanyi, J. (1955) "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility," *Journal of Political Economy* 63: 309–21.
Montgomery, W. (1999) "A Market-Based Discount Rate," in P. Portney and J. Weyant, eds., *Discounting and Intergenerational Equity*, Washington, DC: Resources for the Future, pp. 45–50.
Mueller, D. (1974) "Intergenerational Justice and the Social Discount Rate," *Theory and Decision* 5: 263–73.
Mulgan, T. (2006) *Future People*, Oxford: Oxford University Press.
Nussbaum, M. (2000) *Women and Human Development*, New York: Cambridge University Press.
—— (2011) *Creating Capabilities*, Cambridge, MA: Harvard University Press.
Parfit, D. (1984) *Reasons and Persons*, New York: Oxford University Press.
Rawls, J. (1971) *A Theory of Justice*, Cambridge, MA: Harvard University Press.
—— (1993) *Political Liberalism*, New York: Columbia University Press.
—— (1999) *The Law of Peoples*, Cambridge, MA: Harvard University Press.
Sen, A. (2010) "The Place of Capability in a Theory of Justice," in H. Brighouse and I. Robeyns, eds., *Measuring Justice: Primary Goods and Capabilities*, New York: Cambridge University Press.
Solow, R. (1974) "Intergenerational Equity and Exhaustible Resources," *Review of Economic Studies*, Symposium Issue: 29–45.
—— (1993) "Sustainability: An Economist's Perspective," in R. Dorfman and N. Dorfman, eds., *Selected Readings in Environmental Economics*, New York: Norton.
Vickrey, W. (1960) "Utility, Strategy, and Social Decision Rules," *Quarterly Journal of Economics* 74: 507–35.
Wolf, C. (2004) "O Repugnance, Where is thy Sting? On the Neutral Level of Existence, and Parfit's Repugnant Conclusion," in Jesper Ryberg and Torbjörn Tännsjö, eds., *The Repugnant Conclusion*, Dordrecht: Kluwer.
—— (2009) "Intergenerational Justice, Human Needs, and Climate Policy," in A. Gosseries and L. Meyer, eds., *Intergenerational Justice*, New York: Oxford University Press.
—— (2010) "Intergenerational Justice and Saving," in J. Lamont, G. Gaus, and K. Favor, eds., *Values, Justice, and Economics*, Stanford, CA: Stanford University Press.
World Commission on Environment and Development (WCED) (1987) *Our Common Future*, New York: Oxford University Press.

Part V

THE NATIONAL STATE AND BEYOND

42
NATIONALISM

Margaret Moore

This chapter is on nationalism in political theory. Although nationalism raises many interesting and important normative issues—such as the justifiability of the central nationalist methods and projects of national self-determination and nation-building—this chapter does not consider these. Rather, it is focused on the interface between nationalism theory and political theory, the ways in which nationalism has been (on the whole, inadequately) theorized in political theory, and the challenges that face theorizing about nationalism.

There are, broadly speaking, three main views that political theorists have had of nationalism, which will be discussed in later sections. Before moving to the central argument of the chapter, however, it is important to say something about what nationalism is, and what nations are, in the next section. This sets the groundwork for the later discussion about the extent to which theorists address the concerns of nationalists, or justify nationalist political projects while misunderstanding nationalism. The following section outlines different negative views, held by liberals, democrats and Marxists, and the basis of their largely negative judgments, which are based on both principle and expected consequences. Later, the chapter focuses briefly on the fact that many theorists do not say anything at all about nationalism, and what can be inferred from this. Also, many political theorists, who have said something about nationalism—from Rousseau to John Stuart Mill—have been read by political theorists for their other contributions. Indeed, nationalism may be unique in the contrast that it provides between its political power and its philosophical poverty: as Anderson notes (1991: 5): "unlike most other isms, nationalism has never produced its own grand thinkers: No Hobbeses, Tocquevilles, Marxes, or Webers." Nevertheless, recently, there has been sustained philosophical attention to the phenomenon and aspirations of nationalists, and in the final section, the chapter analyses these and identifies a more positive line of argument about nationalism, or at least an attempt to identify some goods that attach to (certain kinds of) nationalism. These three positive arguments are Miller's argument from trust to democracy and justice; Taylor's politics of recognition argument; and Kymlicka's societal cultures and autonomy argument.

1. What is a Nation?

People engage in nationalist politics when they mobilize on the basis of their national identities, to achieve their national aspirations. This is not the same as promoting the interests of their particular state, since, as Walker Connor (1994) has argued, the terms "state" and "nation" are not synonymous. A nation refers to a community of people,

many members of which believe is a nation, and aspire to be collectively self-governing, and who claim special relations to a particular territory or "homeland." David Miller lists five characteristics which together constitute a nation: it is, he writes, "a community (1) constituted by shared beliefs and mutual commitments, (2) extended in history, (3) active in character, (4) connected to a particular territory, and (5) marked off from other communities by its distinct public culture" (Miller 1995: 27). As Levy and Norman have argued, plausible sociological definitions of a nation, which include political mobilization aimed at collective self-determination as an important criterion for defining a nation, *presuppose* the nationalist political program (Levy 2004: 160; Norman 2006: 5). This circularity is embedded in the subjective component of the definition of nationhood, and particularly the idea that the aspiration for collective self-determination is a criterion for defining the social group as a nation. As Gellner has argued, emphasizing this subjective element: "Nations are the artefacts of men's convictions and loyalties and solidarities" (Gellner 1983: 7).

States, by contrast, are institutional structures, and nations and states are not coextensive. States can contain more than one nation, such as Canada, Spain and the United Kingdom; and a nation can extend beyond state borders to encompass more than one state, as the Kurdish nation does, for example. Nations are also not the same as ethnic groups: there are many ethnic groups, distinguished by language, culture and ethnic descent, and, while some nations may have developed from an ethnic "core," the defining feature of a nation is the commonly held aspiration to be collectively self-governing (Smith 1986).

2. Nationalism as Horror Show

The first, and probably dominant view of nationalism, held by most political theorists, is a negative one: they view nationalism as hostile to the basic moral equality of persons; as based on, indeed creating and perpetuating, morally arbitrary divisions; and as having negative consequences. Marxists typically hold such a negative view: they are concerned with class oppression and note that nationalism tends to serve ruling class interests, divide the working class, and obfuscate the true interests that should unite them. As James Connolly, a Marxist writing in 1913 Ireland, put it: Nationalism is the result of "the perfectly devilish ingenuity of the master class" who use this ideology to further their own ends (Connolly 1975: 2). The standard Marxist or socialist reaction to the institutional recognition of national identity is almost uniformly critical, on the grounds that nationalism is reactionary and romantic (Hobsbawm 1990); that consociationalism is elitist; and that both consociationalism and ethno-federalism institutionalize divisions and reify identities (Lenin 1913; Denitch 1994). At times, however, Marxists, including Lenin, have had to negotiate with nationalists and create ethnofederations to maintain power (Connor 1984), but these were viewed by Marxists as temporary measures *en route* to a non-national socialist future.

Liberals, too, have been unsympathetic to nationalism. Much contemporary liberal political theory is focused on the moral value of individuals; on the meaning and implications of the ideals of equal rights to concern and respect that lie behind the liberal state, which is implicit in the idea of the rule of law, in constitutional rights and processes, international human rights, theories of distributive justice, and so on. All these are explicable in terms of the moral importance of the individual and the compelling vision of the moral equality of persons that lies behind it. Liberals who emphasize

the autonomy component of the liberal doctrine have difficulty giving importance to forms of identity that seem to be unchosen. Liberal egalitarians are uncomfortable with nationalism because it stands in tension with egalitarian aims. As Simon Caney has argued, "If one thinks, as egalitarian liberals do, that it is unjust if persons fare worse because of their class or ethnic identity one should surely also think that it is unjust if persons fare worse because of their nationality" (Caney 2005: 123). Brian Barry offers a classic liberal egalitarian rejection of all types of group-differentiated policies and practices, and applies this critique to minority nationalists (not nationalists who already have a state): they are equated with cultural groups and viewed as regressive, divisive and threatening to the only kinds of equality that it is possible for the state to achieve: equality under the law, equal treatment, and equal rights (Barry 2001: 278–84).

Democratic theory, too, in arguing for institutions of equal voice and equal political rights, presupposes the moral value of individuals, without any regard for the constitution of the democratic collective. Democratic institutions have largely been interpreted as operating within the domain of the state, but there is no principle, internal to democratic theory, to explain how to arrive at this domain of democratic theory. The status quo (statist) bias is under-theorized. The classic statement of this problem was articulated by Ivor Jennings, in relation to Woodrow Wilson's appeal to the concept of self-determination in 1919. According to Jennings: "On the surface, it [the principle of self-determination] seemed reasonable: let the people decide. It was in fact ridiculous because the people cannot decide until somebody decides who are the people" (Jennings 1956: 56).

Although this was presented by Jennings as a problem for national self-determination, it is also a problem connected to democratic theory, which presupposes a defined people who can constitute a majority for the purposes of democratic decision-making, but does not offer a standard internal to democratic theory for defining the relevant territorial unit, and hence people, who should exercise collective self-government.

One of the objections that liberals make to nationalism is a consequentialist one, which doesn't require one to embrace particularly liberal political principles. Brian Barry, for example, claims that multicultural policies (and he includes here institutional recognition of national identities under that rubric), are disastrous: if they succeed, they will increase social isolation and perpetuate cultural differences indefinitely, whereas egalitarian policies will help to facilitate integration into the mainstream of society and improve equality of opportunity (Barry 2001: 317–28). John Dunn offers a more extreme view of the possible negative consequences of nationalism, perhaps bearing in mind a different kind of case. "Nationalism," Dunn writes (1979: 5), "is the starkest political shame of the twentieth century, the deepest, most intractable and yet most unanticipated blot on the political history of the world since the year 1900." Daniel Philpott (1998: 79), focusing on the nationalist aspirations for collective self-determination, writes:

> If you want to see the problem with self-determination, drive through the Bosnian countryside and look at the villages. War's destruction there is checkered: houses of rubble and weeds lie next to wholly intact houses, ones with trimmed rows of bushes, lights inside, mothers and children walking in and out. The wrecked houses are the ones where the Muslims lived when the Serbs came to town, where the Croats lived when the Muslims came to town, or where the inhabitants were simply of the wrong ethnicity at the wrong time.

3. Nationalism Ignored

Of course, many theorists, especially those who wrote prior to 1990, expressed no view at all about nationalism, so it is safe to conclude either that they had a serious blind spot about nationalism, or that their views were largely negative and so nationalism had no place in their idealizations about a just society. Rawls, for example, famously adopted several convenient simplifying assumptions in working out *A Theory of Justice* (1971): that society "is a cooperative scheme in perpetuity," that its membership is "fixed," that there is "no immigration or emigration" and the concern for "difference" or "diversity" was mainly focused on philosophical diversity, the idea that we can reasonably expect people to have a range of (reasonable) conceptions of the good life. The theory is silent on controversial political issues of relevance to nationalists, such as the appropriate public policies and constitutional arrangements for societies that are linguistically, ethnically or nationally diverse. In many cases, of course, this silence about nationalism might simply reflect a negative view of it: after all, if their focus is on "ideal theory," on articulating a blueprint of a just society, it makes sense to abstract from the issue of national division and national aspirations (because nationalism is viewed as a distinctly non-ideal sentiment and institutional arrangement). Of course, to the extent that they do abstract in this way their theories lack relevance and are unable to tackle many of the pressing political issues facing us today, which concern the appropriate response to diversity, to demands for constitutional rights or privileges, for autonomy within the state, or power-sharing in the executive.

While the individualist language of those critical of and silent about nationalism is dominant in contemporary political theory, the actual practice of politics and of states across the world is somewhat different. This is not only because many of them fail to live up to these principles—they are undemocratic, intolerant, and often violate human rights—but, more interesting, at least from a theoretical perspective, is that the world seems to be organized along other principles, of nationality and national political community. Despite a great deal of rhetoric, and also the genuine reality, of globalization, the current world order is dominated by territorial states; indeed, the entire usable earth surface is divided up between jurisdictional units, and in each of these jurisdictional units, the political authorities claim to represent "their nation" and are also frequently engaged in nationalizing projects—that is, attempts to forge or reinforce a nation, by which I mean a community of people who are unified by myths of shared values or shared descent or shared political aspirations, over a territory, over which they exercise collective self-government. Moreover, many people, while they accept a basic general principle of moral equality of persons, and also often equate justice with the sort of distributive justice exemplified by Rawls and others, also often express an attachment to their political community, understood not simply as an institutional structure or constitution, but also as an attachment to territory and co-nationals. Indeed, in many parts of the world, there are nationalist political movements—movements for forms of collective self-determination against the state order in which they are encompassed—from Chechnya to Kashmir, Quebec to Scotland to the Basque country, Kurds in Iraq, Iran and Turkey, to name a few. The ubiquitousness of "nations," both the state-created kind, and nations that are either wholly encapsulated within a state that is dominated by another national community, or nations that straddle state borders, is either dismissed as unfortunate or not directly theorized in much contemporary political theory.

4. Nationalism Justified (in at least some of its versions and projects)

There is, however, a group of more applied political theorists who have turned their attention to nationalism and there is now a significant body of literature, developed since the 1990s, which identifies some core values that nationalism either promotes extrinsically or is associated with. This view can be identified with a distinct methodological approach in applied political theory, which argues that it is important to engage with significant social or political forces and policies, and assess them normatively. Although Jacob Levy, in relation to multiculturalism,

> does not celebrate ethnic and cultural identities, the beauty of diversity they add to the world, or the meaning they add to the lives of many … neither does [he] … advocate or even look toward the transcendence of ethnic or cultural identities in favor of a cosmopolitan ethic.
>
> (Levy 2000: 7)

This practical approach is also at the core of Moore's *Ethics of Nationalism* (2001), which explicitly argues that even if nationalism is socially constructed, by opportunistic elites, and by misguided state policies of assimilation that give rise to counter-nationalism, it does not follow that it is easy or possible to de-construct a nationalist group, once mobilized. This leaves her work, like Levy's, occupying the curious space of asking what we should do if people are nationalists, but not advocating that people become nationalists. This is also clear in Wayne Norman's book *Negotiating Nationalism*, which begins from "empirically uncontroversial starting points," and then asks "how we ought to act, and what sorts of institutions we ought to give ourselves, when we live in societies where nationalism is current" (Norman 2006: Preface, x), but does not ask the more fundamental question, of whether nationalism is part of a good way of life. Once this stance is adopted, the central questions become: What is the best (institutional) way to deal with national divisions? And can we distinguish between positive and negative kinds of nationalism?

This third, more positive view begins from the sociological observation that, in fact, people hold these aspirations and share these identities, and then asks how we should think about them. It tends to distinguish between different kinds of nationalism, or different ways in which nationalism might facilitate normatively positive goals. It is possible to isolate three dominant positions. First, the views of David Miller, following in the footsteps of J.S. Mill, who makes extrinsic arguments to show that nationalism can support important values, principally that it facilitates trust which is in turn important for both democratic governance and the stability of redistributive (social justice) practices. The second view, put forward by Charles Taylor, views nationalism in terms of a Hegelian story of recognition of collective identities. The third view, largely associated with Kymlicka, tends to associate nationalism with other types of multicultural policies, and argues for group-differentiated rights which tend to meet (some of) the main goals of nationalists. This chapter argues that the first view justifies state-framed nationalism, but is difficult to apply to minority (stateless) nationalism, and that it assumes the goodness of the current status quo (statist) baseline. The second and third positions are not guilty of this problem, and they do seem to offer justifications for some kinds of nationalist programs and policies, but they do so in part, I will argue, either by being over-inclusive and so not explaining clearly what kind of institutional "recognition"

is justified, or by misunderstanding nationalism and falsely equating it with cultural distinctiveness.

5. Miller, Democracy, Trust and Social Solidarity (Nationalism)

In *On Nationality*, David Miller argues that there is not one single type of "bad" nationalism, but that nationalism, in the sense of a sentiment of attachment to a nation, can be both good and bad, and he then attempts to isolate some of the goods that flow from nationalist attachments and bonds of solidarity.

One of the most interesting aspects of Miller's argument is the conception of ethics that undergirds it. Unlike many of the liberal views canvassed above, which equate the moral point of view with the standpoint of impartiality, Miller appeals to an associative view of morality, according to which moral action involves the person in weighing a number of values and considerations, including obligations deriving from relationships, in deciding what is the right thing to do. Miller writes:

> We are tied in to many different relationships—families, work groups, voluntary associations, religious and other such communities, nations—each of which makes demands on us, and there is no single overarching perspective from which we can order or rank these demands. In case of conflict … I simply have to weigh [the] … respective claims, reflecting both on the nature of my relationship to the two individuals and on the benefits that each would get from the help I can give.
>
> (Miller 1995: 53)

On this view, we have associative obligations—obligations derived from membership in particular communities—and our obligations to co-nationals are just one of many of these, which need to be weighed. This associative obligation view of nations suggests that we have obligations both to state-framed nations (nations that have states) and also to nations that do not have a state or any institutional recognition of their identity, but who aspire to this, like the Kurds. Although Miller does not discuss this directly, it is implied by the associative obligations component of the argument. Instead of discussing this issue, however, he moves on to note that whether we have associative obligations or not seems in part to depend on whether the association is valuable, and Miller also is at pains to point out that there are objective values that attach to nations, in addition to the subjective value that it embodies.

At one level, Miller employs the idea of democracy to identify the kinds of obligations that might attach to shared nationality. That is to say, it is unclear what obligations a Kurd might have to her fellow Kurds, or a Basque person to fellow Basques. To address this issue, Miller deploys the concept of public culture, and critical debate and reflection within a public culture, to explain what specific obligations might flow from shared national identity, and to identify the moral goods that are internal to national communities. Miller appeals to the idea of a public culture, and justification within such a culture is a procedural mechanism to arrive at the collective goods and obligations, with the caveat that the public culture must be characterized by a process of open critical debate and reflection (Miller 1995: 69). As he writes, "national identities and the public cultures that help to compose them are shaped by processes of rational reflection to which members of the community can compete on an equal footing" (Miller

1995: 70). He then goes on to identify two moral goods that tend to be associated with common nationality.

The first is the view that a common national identity is necessary for democratic institutions to function properly. This view counts among its intellectual progenitors John Stuart Mill, who argued, in *On Representative Government*, that common nationality, which he seemed to equate with common language, is necessary for social unity, which, in turn, is a necessary condition for the proper or well functioning of representative democracy. This view was also held by Rousseau, who, in the *Discourses on Political Economy* and *The Government of Poland* (Rousseau 1973) addressed the problem of achieving unity and maintaining stability. Rousseau argued that, if the people were to be sovereign, they needed to have a corporate identity, and it was therefore necessary to forge common bonds of membership and solidarity to unite the people. In this way, Rousseau seemed to think that a nation-building project (by the state) would facilitate mutual trust and help create a common national identity. Miller makes a similar argument, appealing to the kind of deliberation that is possible in a shared public culture, and identifying shared public culture with the concept of a national community.

In addition to the claim about democracy, Miller has also linked common national identity with trust, and trust with the good functioning of distributive justice schemes. This is a complicated argument, and critics have pressed on the empirical claims implicit in each element of the nexus: they have questioned whether nationalism, or shared national sentiment, generates trust, and the difficulties testing for this, and also the nexus that links trust, or greater levels of trust, with redistributive justice. In this section, I do not attempt to examine the empirical links of this complex argument, except to note that, for Miller, the positive consequences that attach to nationhood, or to shared national identity, only seem to obtain when state and nation are co-extensive, when the democratic and redistributive institutions of the state correspond to the national (deliberating) community. Although the vision of associations and relationships giving rise to obligations seems to suggest that this could justify obligations for both types of national communities, the moral goods that he identifies with nations and which give value to this kind of association suggest that it has a status quo (i.e. statist) bias. Many national communities aspire to be states, aspire to have their own redistributive practice, their own democratic community, but if these are smaller than the existing state, and so destructive of existing institutions of redistributive justice or democratic conversation, it seems that these national communities must be theorized as hostile or inimical to social trust within the society, and to democratic deliberations.

The arguments of Kymlicka and Taylor do seem to treat equally sub-state nationalism with state-framed nationalism, and in this sense are not subject to one of the criticisms that could be made of Miller's justificatory argument for (state-framed) nationalism, but, I argue below, while they provide good arguments for why we might accord political autonomy or rights to nations, they do so by, to some extent, at least, misunderstanding nationalism and the nation. In Taylor's case, they are over-inclusive—they apply to many different types of collective identities, not specifically national identities, and it is not clear what institutional forms should flow from this argument for "recognition." In Kymlicka's case, the argument for collective autonomy does seem to apply to national groups, but it does so only if we conceive of national groups as linked by shared culture, an assumption, which, I argue below, is itself problematic.

6. Nationalism and the Politics of Recognition

In "The Politics of Recognition," Charles Taylor argues that people's well-being is bound up with the formation and maintenance of identities. Some of these identities are collective in the sense that they involve membership in a group, and he also notes that people's self-understanding is formed to some extent dialogically, through interactions with others. It follows from this analysis of the importance of collective identity to well-being that people can suffer non-trivial harm by *misrecognition*: "a person or group of people," Taylor writes (1994: 25), "can suffer real damage, real distortion, if the people or society around them mirror back to them a confining or demeaning or contemptible picture of themselves." As this applies to minority nationalists and indigenous peoples, it seems to imply a politics of recognition, a justification for public policy and constitutional provisions which affirm the identity of the group and promote respect for members of the group *as members of the group*. Although Taylor himself was primarily concerned with Quebec nationalist politics, and wrote in the context of Canada's constitutional wrangling over what is known as "the distinct society clause" in the Meech Lake Accord, the analysis is not confined to members of minority nations. As Nancy Fraser (1995) has noted, most societies have two kinds of status hierarchies: an economic hierarchy, defined by money, in which class position is important; and a status hierarchy, in which certain people are defined as "normal" and people who deviate from the hierarchy can be defined as lower on the hierarchy than others. The status hierarchy is, to some extent, independent of the economic hierarchy and requires different kinds of remedies. The analysis is not simply applied to minority nations but to all oppressed groups; as Young (1993) notes, many groups suffer from oppression in contemporary Western capitalist societies: gays and lesbians, Muslims, Hispanic peoples, native Americans, disabled people, to name a few. This is not necessarily a problem for this theory, but it does suggest that the politics of recognition, and the kinds of arguments on which it is based, is over-inclusive in that it seems to apply to many different collective groups, and, relatedly, it fails to indicate what kinds of remedies are appropriate. Although it can explain why collective identity misrecognition is harmful and liberal non-discrimination policies are insufficient, it cannot explain what group-differentiated policies in particular might be justified, and the extent to which a state is required to accommodate these various different oppressed groups. It cannot justify, in particular, political autonomy rather than, say, power-sharing in the executive branch of government, or constitutional rights aimed at "recognizing" symbolically the distinct society of the group in question.

7. Nationalism as Societal Cultures and Liberal Autonomy

One of the central, perhaps even distinguishing features of nationalism is that it aspires to collective self-government. One of the more prominent arguments in the multicultural literature is an argument for collective rights to self-determination for territorially concentrated cultural groups (which are equated with nations). On this argument, liberal rules are justified in terms of their role in facilitating personal autonomy; and culture is an important background condition for the exercise of autonomy. Collective self-determination is then justified as a mechanism to protect culture, particularly for minority groups who find that their culture is threatened, or at least out-bid through normal democratic majoritarian politics.

Kymlicka's argument begins by identifying a range of interests that are at stake in the recognition of people's culture and relating these interests to an expansive understanding of the conditions that are necessary for the exercise of autonomy. Crucially, it turns out that there are a number of public goods that are necessary conditions for the exercise of autonomy, including a shared public culture. The claim that there is an internal relationship between culture and autonomy is developed by examining the conditions under which individuals can be said to be autonomous. In Kymlicka's view, "individual choice is dependent on the presence of a societal culture, defined by language and history" (Kymlicka 1995: 82–3). Not only does the culture provide the options from which the individual chooses but it infuses them with meaning. This is important to the argument: the autonomous ideal of a self-choosing, self-forming being presupposes some conception of value according to which the life is constituted, and this conception of value is provided by a national or societal culture. At this stage of the argument, the emphasis is on providing a coherent conception of value, a "societal culture" in Kymlicka's terms. The integrating or societal nature of the cultural option is necessary to locate the various options within a coherent overall conception of what is good, or what is valuable.

The next step in the argument is the claim that, since a rich and flourishing culture is an essential condition of the exercise of autonomy, liberals have good reason to adopt measures that would protect culture. At this point, the argument has only shown that the existence of a (or some) flourishing cultural structure is necessary to the exercise of autonomy, but not a particular culture. However, he makes the empirical point that "most people have a very strong bond to their own culture" (Kymlicka 1995: 8). Kymlicka then points to the equality principle to justify rights for minority groups: it is unfair for majorities to have the protection of their culture which comes from being a majority in the state, for this places an unfair burden on minorities, who find that they have to bear additional costs for maintaining their culture. This supports the conclusion that minority national (or societal) cultures should be supported as a context in which personal autonomy is exercised.

This argument has been extensively discussed in terms of its resolution of liberal autonomy with the protection of culture. Because, on this argument, culture is valuable insofar as it contributes to the exercise of autonomy, rights to the protection of culture are justified as long as these are consistent with protection of personal autonomy. Kymlicka proposes that external protections, which allow the group to protect their culture vis-à-vis majority groups in the state, are justified, but not internal restrictions on the autonomy of the members of that culture. The internal–external distinction is intended to reconcile concern with the fairness of access to culture and protection of culture with protection for individuals in the exercise of their autonomy.

This argument provides an important justification for ensuring that jurisdictional authority is held by members of a particular cultural group. It argues that there is an important link between jurisdictional authority, protection of culture, and the background conditions for the exercise of autonomy. It also suggests a mechanism for drawing the boundaries of such communities, namely, that borders should be drawn consistent with people's cultural identifications. Although this argument is usually interpreted as a minority protection argument, the idea that political boundaries should be drawn *around* such groups, to enable them to be collectively self-governing, could also be used to support, in certain circumstances, the secession of some cultural groups.

The most interesting aspect of this argument, from the point of view of this chap-

ter, is that, while it justifies a central nationalist goal, namely, that of political self-determination, it does so in a way that may misunderstand national groups. It justifies political rights for cultural groups, and, to the extent that we understand national groups as cultural groups, the argument applies to them. However, it is not clear that national groups should be understood in this way. In Section 1, nations were identified primarily as political communities, not cultural ones, and while there may be cultural or ethnic markers to maintain boundaries between national groups, it is not at all clear that there is a one-to-one correspondence between "nation" and "culture" that is suggested by the structure of the argument. After all, Canadians and Americans evince very different national orientations, and want distinct political communities, but there is not much difference between parts of Canada and parts of the United States, and the cultural difference between different parts of Canada or different parts of the United States might not be objectively greater. Similarly, surveys of cultural similarity in Northern Ireland show that while Northern Irish Catholics and Protestants demonstrate a great deal of cultural similarity across groups, and while they are much more similar than their "national" compatriots in either the Republic of Ireland or Great Britain, this did not correspond to national sentiments, where almost all Catholics identified as "Irish" and almost all Protestants identified as "British" (Harris 1972: 131). There have been similar findings in Bosnia, just prior to the outbreak of violence there, which show that rural Serbs and Croats were culturally more similar to each other than to their urban counterparts (Malcolm 1994: 282; Bennett 1995: 247). This suggests that the underlying assumption of the argument—that there is a close connection between nation and cultural community—is questionable.

8. Conclusion

Nationalism is a potent political force. In spite of the increased practice and rhetoric of globalization, states still dominate the political landscape: indeed, the entire usable earth is divided up between states; and within these states there are policies designed to effect nation-building and also, in many cases, national self-determination movements threaten the unity and integrity of many existing states. A variety of attitudes have been adopted by political theorists towards nationalism: many of these are critical of nationalism, or silent about it, but this chapter has analyzed three arguments in contemporary normative political theory which identify moral goods associated with shared national identity. In Miller's argument, the goods seem to attach primarily to state-framed nations, even though his view of moral reasoning seems to imply that it should apply to both kinds of nations. Taylor demonstrates why collective identities matter but it is unclear which collective identities we should focus on when they conflict, or what (institutional) recognition follows from the aspiration for "recognition." Kymlicka's argument seems to apply more specifically to national communities (and he has a parallel argument for "polyethnic rights" for immigrant groups) but his approach, I argue above, works by misunderstanding national groups. All this is to suggest that political theory has been and still is deeply uncomfortable with the kind of claims and normative appeals made by nationalists, and neither Kymlicka nor Taylor even attempts to address the question of what might ground the claim to a particular land or territory claimed by nationalists. All this suggests that the universal aspirations of much contemporary political theory sit uncomfortably with the particularist claims of nationalists, and so one of the most important political movements of our time is largely under-theorized.

Related Topics

Hegel, Mill, Liberalism, Multiculturalism, Equality, Freedom, Autonomy, Justice and Borders

References

Anderson, B. (1991) *Imagined Communities. Reflections on the Origins and Spread of Nationalism*, 2nd edn, London: Verso.

Barry, B. (2001) *Culture and Equality*, Cambridge, MA: Harvard University Press.

Bennett, C. (1995) *Yugoslavia's Bloody Collapse: Causes, Course and Consequences*, New York: New York University Press.

Caney, S. (2005) *Justice Beyond Borders: A Global Political Theory*, Oxford: Oxford University Press.

Connolly, J. (1975) [1913] "Ulster," in *Ireland upon the Dissecting Table*, Cork Workers' Club 1975, Transcription: Einde O'Callaghan.

Connor, W. (1984) *The National Question in Marxist Leninist Theory and Strategy*, Princeton, NJ: Princeton University Press.

—— (1994) *Ethnonationalism: The Quest for Understanding*, Princeton, NJ: Princeton University Press.

Denitch, B. (1994) *Ethnic Nationalism and the Tragic Death of Yugoslavia*, Minneapolis: University of Minnesota Press.

Dunn, J. (1979) *Western Political Theory in the Face of the Future*, Cambridge: Cambridge University Press.

Fraser, N. (1995) "Recognition or Redistribution? A Critical Reading of Iris Young's Justice and the Politics of Difference," *Journal of Political Philosophy* 3: 168–70.

Gellner, E. (1983) *Nations and Nationalism*, Ithaca, NY: Cornell University Press.

Harris, R. (1972) *Prejudice and Tolerance in Ulster*, Manchester: Manchester University Press.

Hobsbawm, E.J. (1990) *Nations and Nationalism since 1780*, Cambridge: Cambridge University Press.

Jennings, I. (1956) *The Approach to Self-Government*, Cambridge: Cambridge University Press.

Kymlicka, W. (1995) *Multicultural Citizenship. A Liberal Theory of Minority Rights*, Oxford: Oxford University Press.

Lenin, V.I. (1913) "Lenin's Letter to Shawnyan of November, 1913," Vol. XVIII: p. 90, Collected Works Index, *Marxist Internet Archive* at www.marxists.org.

Levy, J. (2000) *Multiculturalism of Fear*, Oxford: Oxford University Press.

—— (2004) "National Minorities without Nationalism," in Alain Dieckhoff, ed., *The Politics of Belonging: Nationalism, Liberalism and Pluralism*, Lanham, MD: Lexington Press.

Malcolm, N. (1994) *Bosnia: A Short History*, New York: New York University Press.

Miller, D. (1995) *On Nationality*, Oxford: Oxford University Press.

Moore, M. (2001) *Ethics of Nationalism*, Oxford: Oxford University Press.

Norman, W. (2006) *Negotiating Nationalism: Nation-building, Federalism and Secession in the Multination State*, Oxford: Oxford University Press.

Philpott, D. (1998) "Self-Determination in Practice," in Margaret Moore, ed., *National Self-Determination and Secession*, Oxford: Oxford University Press.

Rawls, J. (1971) *A Theory of Justice*, Cambridge, MA: Harvard University Press.

Rousseau, J.J. (1973) *The Social Contract and Discourses*, eds. G.D.H. Cole, J.H. Brumfitt, John C. Hall, London: Everyman.

Smith, A. (1986) *The Ethnic Origins of Nations*, Oxford: Oxford University Press.

Taylor, C. (1994) "The Politics of Recognition," in A. Gutmann, ed., *Multiculturalism and 'the Politics of Recognition'*," Princeton, NJ: Princeton University Press.

Young, I.M. (1993) *Justice and the Politics of Difference*, Princeton, NJ: Princeton University Press.

43
HUMAN RIGHTS AND COSMOPOLITANISM

David A. Reidy

1. Introduction

Since World War II, human rights have moved center stage within the theory and practice of international law and politics and have received increasing attention within moral and political philosophy. The rise of human rights and the corresponding decline of the so-called Westphalian system (of sovereign states oriented only toward their own national interest and bound by their voluntary undertakings only prudentially) has breathed new life into cosmopolitan ambitions.

The natural rights tradition of the seventeenth and eighteenth centuries prefigures in certain ways the modern human rights movement. But it focused on a few relatively abstract moral rights that all persons were claimed to have in a state of nature prior to or apart from any institutional context, were invoked mainly within national contexts as a way of checking or setting limits to domestic political authority, and were justified by appeal to moral first principles and systematic moral theory, often theistic. The idea at the heart of the modern human rights movement initiated by the 1948 Universal Declaration of Human Rights (UDHR), by contrast, is one of numerous and specific rights clearly oriented toward the institutional realities of the twentieth century. They are moral rights in the sense that they are proper objects of international concern and action apart from their current status within positive law, domestic or international. They are universal in the sense that they are possessed by all human persons prior to and apart from any voluntary undertaking by the states to which they belong. But the UDHR drafters offered no further general characterization of these rights and no justification or philosophical ground for them beyond the claims that their recognition serves the cause of international peace and security and expresses and answers to the dignity and worth of human persons.

The cosmopolitan ambitions revitalized by the contemporary human rights movement were prefigured in certain respects by Kant's essay "Perpetual Peace" (1795) as well as older ideals belonging to Stoic and Christian thought. But these earlier cosmopolitanisms differ significantly from their contemporary descendants, not least in that the latter are typically responsive to institutional, economic, and cultural patterns and forces of globalization and global governance absent and probably unimaginable centuries ago.

Anyone committed to human rights, on almost any understanding of them, is a cosmopolitan in some sense, for they recognize some sort of universal moral commu-

nity to which all persons belong. But cosmopolitans today typically commit to more than human rights. They may commit to ideals of global economic or distributive justice, to the subordination of nation-states to a complex network of transnational and international institutions, or to some sort of democratically structured system of global governance. But since they all commit themselves to human rights, that is a good place to start.

2. Human Rights: Some Preliminaries

Whatever else they are or purport to be, human rights are rights. Rights are deontic in their logic. They issue reasons for action in the language of required, forbidden and permissible, not the language of good, better and best. But rights are distinctive in their deontic logic. To say that a particular action is required (or forbidden) and so right (or wrong) is not yet to say that any particular person has *a right* against any other particular person to it (or to be free from it). To say the latter, to point to individual rights rather than simply what is right, is presumably to say more than that an action is required (or forbidden). What this more is, however, is a matter of some controversy. Nevertheless, because human rights are rights, a complete theory of human rights must include an account of their special status (as rights) within the more general deontic field of the right. It must account for the distinctive nature of the deontic reasons they express.

It must also include an account of the nature of the more general deontic field of the right and its relationship to the axiological domain of the good, for both issue in reasons for action. Here what is needed is an account of whether the deontic field of the right is best explained as arising out of a rational demand to maximize the good or as arising out of some other ground. If the latter, then what is needed is an account of the normative relationship between the right and the good, including an account of the ways in which the right conditions or constrains the reasons issuing from the good.

These important matters of moral theory are beyond the scope of this essay. Nevertheless, it bears noting that disagreements over how best to understand human rights sometimes turn on just these sorts of issues.

It is often said that human rights are the rights all and only humans possess simply because they are human. But some rights typically identified as human rights are rights that only human *persons* could plausibly be said to hold. For example, the rights to own and dispose of property, to join trade unions and to hold political office are not typically thought to be held by infants or the very seriously cognitively impaired, who are said to be humans but not persons. Of course, one might argue that *really* human rights are only the rights that all and only humans possess simply because they are human. But which rights are these? The list is likely short. Perhaps the right to be free from intentionally inflicted suffering for which there is no moral justification, or not to be treated as if merely a thing. But what else? Further, it seems likely that any reason for assigning rights to all humans simply because they are humans (rather than because they are persons) would also justify assigning those same rights to some non-human animals. Other than controversial theological doctrines (e.g., God loves and holds in uniquely special esteem the human species) what other reason could there be for singling out of the natural order all and only members of the human species for the special status of holding human rights? (See Wolterstorff 2010.)

With respect to the contemporary idea of human rights, the claim that human rights are the rights all and only humans possess simply because they are human is probably

best understood by reference to the fact that human rights were introduced into international relations and law after World War II to serve as a check against the sort of evil worked by states assigning a lesser or perhaps even no moral status to some human groups within their populations. The claim need only be taken, then, to assert that every state must accord all human members in its population a certain threshold moral status irrespective of group-based membership.

Two distinctions common within discussions of human rights merit special notice as troublesome. One is the distinction between so-called negative and positive rights. Negative rights are more or less rights to be left alone. They impose only duties of non-interference. Positive rights on the other hand are rights to some particular object or positive action. They impose duties requiring more than non-interference. It is sometimes said that positive rights depend for their existence on some sort of voluntary undertaking or special relationship so that human rights, as universal moral rights that depend on neither, are limited to negative rights. But this is a mistake. Most human rights have both negative and positive dimensions. The right to life, for example, imposes duties of non-interference (don't kill people unless you have a recognized moral justification such as self-defense for so doing) as well as duties to take reasonable steps to protect people from being killed and to redress unjustified killings, at least when these things can be done without significant cost. Without these duties to act, which seem not to depend on voluntary undertakings or special relationships, the right to life is left incomplete or hollow.

The other troublesome distinction is between so-called civil and political rights, on the one hand, and economic, cultural and social rights, on the other. The former are typically said to be the liberty rights essential to civil or political or citizenship status under the rule of law, whereas the latter are said to be the welfare rights essential to a certain minimally acceptable level of well-being. But as with the distinction between negative and positive rights, the distinction between civil/political and economic/cultural/social rights is, while not without some basis, too flimsy to do any real work. The human right to own and to transfer property, for example, seems to be essential to civil or political or citizenship status as well as to a minimally acceptable level of material well-being. The same holds for the right to speak one's own language or to a basic education.

Henry Shue argues that to enjoy anything *as a right*, rather than at the sufferance of others, one must have basic liberty, security and subsistence rights, each of which has both negative and positive dimensions (Shue 1996). For example, to enjoy political participation as a right, one must also have a right to be free from physical assault and to basic subsistence. Otherwise one is able to participate in politics only at the sufferance of others. And these security and subsistence rights will have both positive and negative dimensions. For what would the security right to be free from physical assault come to if no one was duty bound to take any positive action to discourage or prevent assaults, to aid those under assault, or to take steps to redress completed assaults. Shue's reasoning casts doubt on the idea that the distinctions between negative and positive rights and between civil/political and social/cultural/economic rights run very deep.

Some commentators and activists, moved by the foregoing considerations, have declared the *essential unity of all human rights*. The idea is that every human right presupposes and is bound up with every other human right. But this seems little more than rhetoric. To be sure, as Jim Nickel has emphasized, human rights are often linked in various dependency relations and "linkage arguments" often constitute good and

compelling arguments in favor of particular human rights. For example, under modern conditions subsistence rights likely depend on property rights (Nickel 2007). But it does not follow that there is an essential unity of all human rights.

Though our focus here is on human rights as moral rights, a distinctive feature of human rights as moral rights, at least on the contemporary understanding, is that they aspire to the status of legal rights within international law. Further, one prominent approach to theorizing human rights as moral rights proceeds by way of a constructive interpretation of international law and politics as an emergent practice. Accordingly, a few preliminary remarks are in order about human rights as legal rights within international law.

The UDHR is not a treaty and does not establish the rights it sets out as legally recognized or enforceable within international law. Its significance is primarily political, not legal. It constitutes a first step toward the development of a system of human rights as international legal rights. There is some debate over whether some of the rights it sets out have become international law by achieving the status of customary practice within international relations. There is also some debate over whether some of them are *jus cogens* and so binding on all states, even those that have resisted emergent customary practices within international relations. But the main way in which the rights it sets out become international law is by being agreed to by states party to treaties (e.g., the Conventions on Civil and Political Rights, and on Economic, Social and Cultural Rights).

Treaties bind only the states party to them and then only after they ratify in accord with their respective constitutions. States often sign and ratify treaties subject to certain reservations. When they do, they are not bound by the provisions with respect to which they express reservation. To determine the legal status of the rights set out in any treaty, one must determine which states signed and ratified and then identify the reservations they expressed.

There is no settled, centralized, global enforcement mechanism with respect to human rights. Adjudication and enforcement are matters typically left to state parties to settle for themselves. They often defer these matters to optional protocols. Often human rights treaties include no provisions for adjudication or enforcement or include only limited reporting requirements. This makes it easier for states to sign and ratify human rights treaties for reasons other than their being fully committed as a matter of international law to honoring the rights they establish. And it helps to explain the gap between the rhetoric and politics of human rights and their effective legal enforcement.

To be sure, human rights are sometimes enforced through familiar modes of legal adjudication, domestic and international, yielding judicial orders capable of inducing compliance by a noncompliant state. But more often they are enforced through formal and informal diplomatic pressure, economic and other sanctions or incentives, targeted forms of non-cooperation (for example, the refusal to honor extradition requests from countries that violate prohibitions on cruel and unusual punishment), as well as so-called 'jaw-boning' (Nickel 2007). In rare cases they are enforced by military interventions, with or without United Nations Security Council authorization.

States belonging to the United Nations are obligated to seek Security Council approval for military interventions. The Security Council has clear authority to authorize the use of force only to secure international peace and security. So human rights violations that do not threaten international peace and security present a challenge. While there is growing support for the permissibility of coercive intervention, even without Security Council authorization, to bring an end to grave and systematic human

rights violations, there is little consensus beyond this, and not much consensus, except for the most obvious cases of genocide, on what counts as a grave and systematic human rights violation. Without Security Council authorization, coercive interventions in the name of human rights are typically controversial.

In 2005 the United Nations member states participating in the World Summit endorsed what has come to be known as the "Responsibility to Protect" (R2P) doctrine. It requires states to protect their citizens against certain mass atrocities such as genocide, ethnic cleansing and war crimes and, should they fail, it assigns the international community the responsibility (not just discretionary authority) to take the necessary steps to do so. In 2006 the Security Council affirmed the doctrine. In 2009 the Secretary General issued a report with guidelines on implementation. The report was subsequently discussed by the General Assembly. There was general agreement that the doctrine needs to be more formally entrenched and then implemented, but little agreement beyond this. The R2P doctrine does not yet have the status of international law, but it is emerging as an important political norm of international relations and reflects the cosmopolitan thrust of the contemporary human rights movement.

The European Community has achieved perhaps the most fully developed system for adjudicating and enforcing human rights within and among its member states. The extent to which its achievement might be expanded to or reproduced in other regions remains unclear. What seems likely is the development of diverse regional models of adjudication and enforcement, each, like the European system, affirming a "margin of appreciation" doctrine that allows states and multi-state regions non-trivial discretion when it comes to interpreting and fulfilling human rights within their own diverse institutional, historical, cultural and demographic contexts.

3. Human Rights: Overcoming Skepticism

The idea of human rights has long met with skepticism. One positivist form runs back at least to Jeremy Bentham (Bentham 2010). For Bentham, the paradigmatic case of a right was the institutionally realized legal right. It exists as a matter of objectively and empirically ascertainable social fact. Only by so existing does it structure social expectations and constitute a special case in the general deontic domain of the right. Bentham regarded universal moral rights such as the natural rights articulated in the French Declaration of the Rights of Man and of the Citizen (1789; subsequently prefixed to the 1791 French Constitution) as mere shadows of legal rights, little more than pretend rights. If they existed at all, they did not exist as objectively and empirically ascertainable social facts and so could not structure social expectations in the way that legal rights did. Bentham declared them "nonsense on stilts." Of course, he did not mean to deny that putative natural rights could express valid moral claims, even claims registering in the deontic domain of right. He meant to insist that a valid moral claim, even one registering in the domain of right, did not by itself constitute *a right*. If it did then there would be virtually no limit to the range, content and number of rights and no way for rights to perform their distinctive function within social life, marking off just part of the more general domain of the right for special interpersonal treatment. Bentham meant also to draw attention to the fact that there is very often considerable disagreement over the validity of moral claims registering in the domain of the right, and invoking the language of rights is often merely a rhetorical move aimed at sidestepping this disagreement. Better, Bentham concluded, to restrict rights talk to the paradigm case of institutionally realized legal rights.

Unlike Bentham, Kant was broadly sympathetic to the idea of natural rights. But he too held that to have a right was to have more than a valid moral claim (Kant 1996). Kant argued that in the absence of any common public mechanism to govern their specification, adjudication and enforcement, each person was left, indeed required, to judge for herself the content of her natural rights, the duties they entail, the persons on whom those duties fall, how to adjust her various natural rights one to another and to the natural rights of others, and the appropriate mechanisms for their enforcement and remedies for their violation. And, of course, each was then left to act on her own, presumably reasonable, judgment. This was not, for Kant, a mere "inconvenience" of the state of nature, as it was for Locke. It was a metaphysical obstacle to the realization of natural rights as rights. This is because without some public authority to determine rights there is no path to a system of rights that does not involve the arbitrary subordination of some person's private reasonable judgment and will to that of another. But that is itself a violation of right. Kant concluded that in the absence of a constitutional polity subject to the rule of law with a common public authority to which the reasonable private judgments of all may be equally subordinated, the natural rights of each would remain forever defective as rights. To be perfected as rights, they must be realized as legal rights within a constitutional republic. Kant further concluded that if the world was to be divided into distinct and independent constitutional republics, they would in turn each have to be subordinated to the common public authority of an international federation, otherwise their relations one to another would be the result of arbitrary subordinations of the reasonable private judgments and willing of some to others.

Bentham and Kant suggest that there are good reasons to be skeptical about rights understood simply as valid moral claims. Each points toward established social institutions or practices as the site within and through which ostensibly valid moral claims are transformed into rights. But caution is in order here. The idea that rights and so human rights exist only insofar as they achieve official legal recognition, or analogous institutional embodiment, even if only in conventional morality, is problematic for it calls into question the very idea of moral rights as a normative ground for criticizing existing institutions, practices and beliefs. And that may seem too conservative a view of rights. After all, as moral rights, human rights are often invoked as a normative ground for criticizing existing institutions and conventional morality.

A more familiar ground for skepticism about human rights is the relativist idea that there are no universal criteria of moral truth, validity or justification. There is a good deal of moral disagreement worldwide. And this may be thought to suggest the absence of such universal criteria. But much of this disagreement is reasonable disagreement between intelligent and informed persons of good will who seem to be reasoning in ways free of obvious error or bias and so on. And reasonable disagreement can be explained without any appeal to relativism. Of course, not all moral disagreements are reasonable. But unreasonable disagreements are probably best explained by reference to such common failures of reasoning as bias and dogmatism. Finally, while there is a good deal of moral disagreement worldwide, it seems to have been steadily, if slowly and episodically, narrowed over the centuries. Over time, unreasonable disagreements are exposed. This suggests the availability of some criteria of moral truth, validity or justification universal in reach.

Another familiar ground for skepticism about human rights is the realist idea that the international arena is a kind of Leviathan-free Hobbesian state of nature within which nation-states pursue only their own self-interest. As a descriptive claim, this picture is hardly compelling. Today, more than ever, states (and other international and

transnational agents) cooperate subject to common moral norms across wide domains of their interaction. While there is no Leviathan-like world state, there is a complex decentralized system of global governance that is increasingly effective. Of course, one might say that all this is merely calculated self-interest. But this seems unlikely. To be sure, all states properly pursue their national interests. But increasingly many do so recognizing that their national interests are partly given and fully constrained by their citizens' shared moral interests. These include living in a world within which human rights are universally secure and within which states cooperate on terms each can affirm as fair. No doubt there is a good deal of disagreement, much of it reasonable, about what is involved in satisfying these interests. And these disagreements can persist and prove destabilizing. But none of this tells, descriptively or normatively, in favor of the realist's vision of an international Hobbesian state of nature. (See Beitz 1979; Rawls 1999.)

4. Human Rights: Rival Conceptions

Two general sorts of conceptions of human rights have emerged in recent years. On the first, human rights are more or less continuous with the natural rights tradition. On the other, they are largely a *sui generis* mid-twentieth-century development. There are many varieties of each sort of conception. All address themselves to the nature, function, content, and justification of human rights. James Griffin and Alan Gewirth each offer a distinctive account of human rights as more or less continuous with natural rights.

Griffin does not come to human rights with a comprehensive, systematic moral view or set of foundational first principles already in hand (Griffin 2008). In this regard he departs from the older natural rights tradition. Griffin theorizes particular domains of moral experience taken on their own terms. He is open to discovering, but not a priori committed to, the need for and the possibility of stitching together the results of his piecemeal work into a more fully comprehensive and systematic outlook resting on something like foundational first principles.

Somewhat surprisingly, to theorize human rights Griffin focuses not on international law and politics, but rather on rights talk as it has developed within interpersonal relations in an increasingly globalized civil society. This talk he interprets as descending from the natural rights tradition. But having been cut free of that tradition's doctrinal and institutional context (theism, national struggles against absolutist rulers), contemporary human rights talk, is at risk of degenerating into a Tower of Babel. To be rescued it must be reinterpreted and reconstructed in a new way in light of today's doctrinal and institutional context.

Griffin proposes that human rights are a conventional rational social means to promoting the basic, general interest shared by all persons in developing and exercising the capacity for normative agency. This capacity is exercised and expressed when persons consider, evaluate, choose and revise their own course in life. Griffin argues that other things being equal, persons' lives go better to the extent that they develop and exercise this capacity, not merely because so doing is a means to their good, but rather because it is a constitutive ingredient of their good. On Griffin's view, then, human rights express and answer to considerations of the good. His view is, accordingly, teleological (though not utilitarian).

Griffin analyzes the capacity for normative agency into the capacity to make choices (autonomy), to act on choices (liberty), and to succeed in acting on choices (material welfare). All persons have weighty fundamental interests tied to each. These interests,

conjoined with general practical facts about the human condition, constitute the normative underpinnings of human rights, which in turn function as an efficient social means to, and so as the conventional expression of our common rational orientation toward, the goods realized when these interests are fulfilled.

The human rights to autonomy, liberty, and material welfare are very abstract. Griffin is not altogether clear about how these are to be rendered more determinate (e.g., how are we to determine whether, in the contemporary context, any candidate human right, such as the right to same-sex marriage or to democratic government, follows from or is part of one of the three abstract rights?). To avoid the sort of unconstrained proliferation of rights that Griffin thinks plagues contemporary human rights talk and that might plague his own view (for the list of things that promote autonomy, liberty and human welfare is long), Griffin emphasizes that human rights answer to only that which is essential to, rather than that which merely promotes, the development and exercise of normative agency. But it is not clear that this notion can bear the weight Griffin puts on it. Were slaves in Ancient Greece denied something essential to their normative agency? An affirmative response suggests they were not able to achieve normative agency. But that seems wrong. But a negative response suggests that there is no human right to immunity from the form of slavery to which they were subject. That too seems wrong. One is tempted to say that their human right to freedom from slavery is a matter of promoting their interest in normative agency. But then the floodgates to a proliferation of human rights seem to be open (Raz 2010).

On Griffin's view, human rights are weighty claims arising out of a great universal good. But they compete with other weighty claims arising out of other great goods, such as social justice, personal virtue, and communal solidarity. When human rights conflict with these other claims we can do no better, Griffin argues, than assess the balance on a case-by-case basis. Here human rights have no special moral weight. To be sure, given their moral significance we should decide in favor of human rights unless it is very clear that the balance tips in favor of the other claim or good. Of course, if the other claims or goods are also of great moral significance, the same will hold for them. In either case, while we must be certain that the balance tips, it need only slightly tip. Human rights function in the economy of moral discourse and deliberation as just one important sort of reason for acting. All reasons for acting ultimately derive from considerations of the good, and so, given a plurality of great goods, any reason might carry the day in a particular case, though we should be careful when the goods are especially significant to ensure we have all the needed information and a clear view before we assess the balance.

Alan Gewirth also conceives of human rights in a manner largely continuous with the natural rights tradition, but he offers an anti-teleological account (Gewirth 1996). Gewirth's basic argument is that to be a person is to be the sort of being that acts for reasons or purposes, but in order to act for reasons or purposes, one must have a certain amount of freedom and material well-being. Each person must regard her own freedom and material well-being, then, as necessary goods, and that means she must think of others as normatively prohibited from interfering with her freedom and well-being. But to be logically consistent she must recognize that everyone else is in the same position. If others are constrained to respect her freedom and well-being, she is constrained to respect theirs. Human rights, then, arise out of these logically reciprocal, interpersonal, and pre-institutional deontic demands to freedom and well-being.

For Gewirth, the justificatory ground of human rights is the structure of practical rationality or agency. Human rights arise as a matter of logic independent of any

particular good at which we aim when we act with purpose or for reasons. Having derived human rights from considerations independent of the good, Gewirth holds that they also generally trump or constrain considerations of the good. Considerations of the good do not generate reasons for acting until one has given considerations of right their due. So, Gewirth's view is an anti-teleological or deontological view that affirms the priority of the right over the good.

Gewirth faces a challenge similar to that faced by Griffin when it comes to specifying the determinate rights entailed by or ingredient in his general abstract rights to freedom and well-being. This is a challenge for any view that assimilates contemporary human rights talk to the natural rights tradition.

Charles Beitz and John Rawls largely set aside the natural rights tradition and ask what human rights must be on the best constructive interpretation of the emergent international institutions and practices, often legal, in which they now figure so prominently (Beitz 2010; Rawls 1999). They argue that human rights express an international commitment to ensuring, through international action and law if necessary, that certain important interests are secured for all persons against standard threats posed by modern states. If we ask ourselves, what is the point and purpose of the emergent post-World War II international order, the best or most compelling answers will all refer to this commitment. The interests served by human rights need not derive from a single tap root (such as our interest in the good of normative agency or the logical structure of status as normative agents or even the idea of a minimally decent human life). Human rights are likely to serve many diverse, though all important and widely shared, interests.

Neither Beitz nor Rawls claim that the emergent international system has as its sole point and purpose the realization of human rights. It no doubt has many aims. They claim only that there is no plausible constructive interpretation of it that will not include among its basic aims the realization of human rights understood as an international institutional commitment to protecting certain well-defined interests shared by all or nearly all persons from standard threats posed by modern institutions, especially the state.

Rawls suggests one line of constructive interpretation. He imagines representatives of contemporary liberal democratic polities placed in an international original position behind a veil of ignorance and required to reach unanimous agreement on principles to constitute and govern cooperative international interaction. He argues that they would all agree, inter alia, to a principle requiring respect for human rights understood as limited to those set out in or implied by Articles 3–18 of the UDHR as well as the conventions on genocide and apartheid. Taken together, these "human rights proper" condition the recognition or standing of a state as a free and independent potential cooperator within international relations, with a presumptive right to nonintervention, on its being organized as a constitutional republic subject to the rule of law.

Of course, states so organized may voluntarily take on additional commitments to human rights by treaty as a matter of international law. For example, liberal democracies may so commit to a human right to democracy within international law. But by so doing they cannot bind states that are not party to the treaty. Rawls argues that this conception of "human rights proper" is one that could be affirmed by any "decent people," including hierarchical Islamist states organized as constitutional republics subject to the rule of law, and so is not in any way parochial or distinctively "Western." This notion of "human rights proper," belongs, then, to the best constructive interpretation of the institutional international order that has emerged since World War II, for

it captures the idea that recognition and standing in the international order, with an attendant right to nonintervention, is predicated upon achieving political institutions that aim at the common good of all members subject to the rule of law, and that the progressive liberalization and democratization of all polities meriting recognition and standing is something to be realized through their voluntary undertakings and without force in violation of the right to nonintervention.

Beitz affirms a somewhat more expansive list of human rights, though he too stops short of claiming that there is a human right to something like liberal democracy. This is in part because he pays closer attention than Rawls to the many roles human rights play within international relations and law beyond informing the principle of nonintervention and setting limits on permissible pluralism, from informing political discourse across multiple domains within global civil society to framing and giving direction to voluntary international treaty-making. So Beitz's constructive interpretation of the emergent institutional international order includes a commitment not only to a somewhat more expansive list of human rights, but also to a more permissive stance toward various forms of non-military or soft coercion (e.g., financial incentives) aimed at inducing widespread compliance with that list.

5. Cosmopolitanism(s) Trailing in the Wake

In some sense, all those committed to human rights are cosmopolitans. But the cosmopolitanism trailing in the wake of a commitment to human rights will vary with the nature of that commitment. A more robust and demanding cosmopolitanism trails in the wake of Beitz's commitment to human rights than trails in the wake of Rawls's. And a still more robust and demanding cosmopolitanism potentially trails in the wake of Griffin's and Gewirth's commitments to human rights, depending on the list of specific rights generated from their abstract general rights.

Those already cosmopolitan by virtue of their commitment to human rights will differ over what else a cosmopolitan ought to be committed to. To be cosmopolitan is, as a matter of etymology, to be committed to some ideal of a "global or worldwide polity" or an ideal of "global citizenship." As a matter of contemporary usage, it is to be committed to a world within which the moral status of all persons is fully secured as an object of universally shared moral concern. This clearly entails a commitment to human rights. But it's not clear what else it entails. For some, it entails a commitment to something like a decentralized, demilitarized, liberal democratic world "state" the members of which prioritize their global citizenship over more local attachments and realize in their shared institutions an egalitarian conception of distributive justice, economically and otherwise. For others, it entails a commitment to a complex worldwide system of transnational and international institutions and norms within which semi-autonomous and diverse constitutional republics, some perhaps neither liberal nor democratic but all capable of drawing the free allegiance of their members, cooperate across multiple dimensions for their common good, but without aiming at any more substantial ideal of distributive justice, economic or otherwise. For still others, it entails only a commitment to an international rule of law. For some, it entails mainly a commitment to an open, dynamic, multicultural global civil society within which voluntary associations are unconstrained by national borders and identities, and persons affirm their fidelity to a common universal interpersonal morality. And for still others, it requires a world within which the morally arbitrary fact of one's nationality plays no role in determining one's life prospects.

Contemporary cosmopolitans face two key challenges. They must determine what follows from their common commitment to recognizing and respecting the basic and inviolable moral status of all human persons. And they must do so in a way that does not run roughshod over the local, particular, and historical dimensions of social life from which persons draw their self-understandings, assess their options, justify themselves to one another, and find meaning and direction in their lives.

Related Topics

Natural Law and Rights Theory, Nationalism, Global Justice and Politics, Authority and Legitimacy, Rights, Toleration

References

Beitz, C. (1979) *Political Theory and International Relations*, Princeton, NJ: Princeton University Press.
—— (2010) *The Idea of Human Rights*, Oxford: Oxford University Press.
Bentham, J. (2010) "Anarchical Fallacies," in *The Works of Jeremy Bentham*, vol. 2, ed. J. Bowring, Charleston, SC: Nabu Press.
Gewirth, A. (1996) *The Community of Rights*, Chicago, IL: University of Chicago Press.
Griffin, J. (2008) *On Human Rights*, Oxford: Oxford University Press.
Kant, I. (1996) "The Doctrine of Right," in *The Metaphysics of Morals*, ed. Mary Gregor, Cambridge: Cambridge University Press.
—— (1970) "Perpetual Peace," in *Kant: Political Writings*, ed. H.S. Reiss, Cambridge: Cambridge University Press.
Nickel, J. (2007) *Making Sense of Human Rights*, 2nd edn, Oxford: Blackwell Publishing.
Rawls, J. (1999) *The Law of Peoples*, Cambridge, MA: Harvard University Press.
Raz, J. (2010) "Human Rights Without Foundations," in *The Philosophy of International Law*, eds. S. Besson and J. Tasioulas, Oxford: Oxford University Press.
Shue, H. (1996) *Basic Rights: Subsistence, Affluence and U.S. Foreign Policy*, 2nd edn, Princeton, NJ: Princeton University Press.
Wolterstorff, N. (2010) *Justice: Rights and Wrongs*, Princeton, NJ: Princeton University Press.

Further Reading

Buchanan, A. (2007) *Justice, Legitimacy and Self-Determination: Moral Foundations of International Law*, Oxford: Oxford University Press.
Donnelly, J. (2003) *Universal Human Rights*, 2nd edn, Ithaca, NY: Cornell University Press.
Fine, R. (2007) *Cosmopolitanism*, New York: Routledge Publishing.
Ignatieff, M. (2003) *Human Rights as Politics and Idolatry*, Princeton, NJ: Princeton University Press.
Moellendorf, D. (2002) *Cosmopolitan Justice*, Boulder, CO: Westview Press.
Morsink, J. (2009) *Inherent Human Rights: Philosophical Roots of the Universal Declaration*, Philadelphia: University of Pennsylvania Press.
Perry, M. (1998) *The Idea of Human Rights: Four Inquiries*, Oxford: Oxford University Press.
Pogge, T. (2002) *World Poverty and Human Rights*, Oxford: Blackwell Publishing.
Talbott, W. (2005) *Which Rights Should be Universal?*, Oxford: Oxford University Press.

44
MULTICULTURALISM

Chandran Kukathas

1. Introduction

When Plenty Coups, the last great chief of the Crow nation, recounted the history of his people to Frank Linderman, he declined to speak about anything in the life of the Crow after they were confined to the reservation. He said: "When the buffalo went away the hearts of my people fell to the ground, and they could not lift them up again. After this nothing happened" (Linderman 2002; Lear 2006: 2). In saying this, Plenty Coups was giving voice to the idea that a traditional form of life had ended. Yet why should it matter that one form of life has ended if another has begun? One answer might be found in the account, given by an elderly Crow woman named Pretty Shield, of the shame she felt for using a saddle-strap on her disobedient grandchild, who had run off to a dance "with a bad young man." She had grown up in a society in which one never struck a child. "I felt ashamed," she said, before adding: "I am trying to live a life I do not understand" (Lear 2006: 61). Pretty Shield knew she had saved her granddaughter from "one of the emerging degradations of reservation life," but did so in the context of a life she did not understand and did not know how to evaluate (Lear 2006: 61). Even if it is possible for humans to live by many and varied traditions, it is not easy—for some it may even be impossible—to move from one familiar form into another, alien, way of life. Plenty Coups, as leader of the Crow, addressed this issue as a practical problem: how could the Crow as a people manage the transition to the modern world? It is also the problem that, admittedly with less urgency, theorists of multiculturalism address in more general form: how should the diversity of traditions within a society be handled, given that minorities will surely struggle and suffer (both individually and collectively) in their efforts to accommodate themselves to the realities of their condition, and to flourish?

This is an old question in political theory, presented in this instance from the perspective of the minority tradition. It was addressed in the past by the Spanish scholastics who debated the question of the proper treatment of the Indians of the new world coming under the rule of a colonizing empire, and under the power of settlers looking to transform the foreign landscape to suit their own cultural preferences. Thinkers such as Francisco de Vitoria thought that there were limits to the extent to which the Spanish could exercise dominion over peoples from distant lands, though it was permissible to preach to or trade with them (Vitoria 1991a, b). Others, such as Bartholomé de Las Casas argued that it was wrong to enslave such people, and justifiable only to leave them alone, or to introduce them to the blessings of European society, while protecting them from the harmful consequences of colonization (Las Casas 1999, 2003; Hanke 1974, 2002). Minorities ought not to be exploited, but protected.

Yet there is another perspective from which this whole matter might be viewed, which was also the standpoint adopted by the critics of thinkers like Las Casas. This is the perspective not of the minority culture or tradition but of the dominant culture or the wider society. It asks not what is good or best for the minority, but whether or not society can be expected to accommodate diversity or difference in its midst if its own flourishing and the well-being of people generally are at stake. Indeed, if there are universal values to be considered, perhaps even tolerating many minority traditions is a mistake (Kukathas 2006a).

The concerns evident in these two perspectives dominate the debate about multiculturalism that has been a part of politics for the past 40 years, and central to political theory for more than two decades. The minorities in question today are indigenous peoples in post-colonial societies, and (broadly speaking) immigrant groups in once-colonized and ex-colonizing societies alike, as well as groups who hold minority status for reasons of religion, language, or culture. Multiculturalism had its beginnings as a public policy in Canada and Australia. It supplanted earlier policies, which aimed at assimilating both indigenous peoples and immigrants to create a homogeneous and cohesive population. Those earlier policies had provoked hostility from indigenous peoples, who had begun to call for recognition of their cultural distinctiveness and redress of historical wrongs. They also led to disaffection among immigrant groups, who expressed a strong desire to hold on to some, if not most, of their customs, their cultural traditions, and their languages. Some groups went further to argue that difference should not only be tolerated but also preserved and protected.

The central issue in the theoretical debates that ensued has been whether, how, and to what extent, diversity should be embraced. Much of that debate has taken place against the backdrop of broadly liberal thinking, for the problem of multiculturalism has been an issue primarily in liberal democracies. There are two questions that need to be resolved. The first is whether minority groups ought to be given cultural protection. The second is whether minority groups ought to be required or made to conform to (at least the most important) ethical standards of the majority society. Theories of multiculturalism have thus tended to embrace one of four possible positions: (1) *favoring* group protection and also *favoring* a measure of cultural conformity; (2) *against* group protection but *favoring* cultural conformity; (3) *favoring* group protection but *against* requiring cultural conformity; and (4) *against* group protection and *against* requiring cultural conformity.

2. Group Rights with Liberalization

The best-known advocate of the first position is Will Kymlicka, who defends the view that cultural minorities should be protected from domination by the majority by being granted group-differentiated rights (Kymlicka 1989, 1995, 2001, 2007). Kymlicka's theory is the most influential in the literature of multiculturalism and is generally credited with initiating the debate that has ensued. His primary thesis is that liberal political theory has erred in assuming that political institutions should look primarily (or only) to protecting *individual* rights when it was evident that how well people fared depended not only on whether they enjoyed individual protection under the law but also on whether the cultural communities of which they were part were similarly protected. People whose cultures were threatened were in danger of suffering from the destruction of their cultural environments, without which they could not flourish. Having the

standard individual rights might be enough for members of the cosmopolitan majority; but minorities needed more than guarantees of free speech, religious liberty, and civil rights generally. They needed group-differentiated rights.

Kymlicka's argument is presented as a contribution within liberal political theory and rests on a claim that what matters is that individuals live autonomous lives whose value they themselves can endorse. For this they need not only certain freedoms, and the resources necessary for living any kind of good life, but also a secure cultural structure within which they can learn what a good life can be, and how to live such a life autonomously. The liberal ideal, he thinks, is a society of free and equal individuals. People can only enjoy a form of freedom and equality they value if it is freedom and equality within their own "societal culture." This means they value the integrity of their nation-state, with its own distinctive national culture, and the integrity of their more particular cultural groups. Unfortunately, Kymlicka argues, "the fate of ethnic and national groups around the world is in the hands of xenophobic nationalists, religious extremists, and military dictators" (Kymlicka 1995: 195). If liberalism is to take hold in these places, he thinks, liberal theory must explicitly address the needs and concerns of such minorities. It can do so by recognizing the rights not only of individuals but of groups. Though liberal theory has generally been reluctant to consider groups as rights-bearers, Kymlicka insists that this does not require any departure from liberalism, which has always assumed the centrality of one particular kind of group: the nation or state. Liberals are generally liberal nationalists (Kymlicka 1995: 93; Tamir 1990: 139).

The liberal nation-state, in Kymlicka's theory, should be conceived on the model of a treaty among different cultural communities, for the nation is often made up of different peoples. The ideal he has in view is portrayed by Edward Hicks in "The Peaceable Kingdom," the painting that graces the cover of Kymlicka's *Multicultural Citizenship*. It depicts the signing of a treaty between Pennsylvania Quakers and three Indian tribes, against the foreground of a peaceful gathering of animals, including the wolf, the lamb and the lion, resting as children play among them. Untroubled by Woody Allen's quip (that the lion will lay down with the lamb but the lamb won't get much sleep), Kymlicka suggests that this image provides the ideal metaphor for a multicultural society.

What is crucial for Kymlicka is that multiculturalism accommodate not only immigrant peoples but also indigenous peoples and what he calls (non-immigrant) "national minorities." Societies like Canada contain not only the largely cosmopolitan majority and immigrant groups of recent arrivals, but also indigenous peoples such as the Inuit, and communities such as the Quebecois. His treaty model is conceived as a way of theorizing the relationship of these different groups in a single state. According to this model, everyone should enjoy fundamental liberal rights as citizens of a democratic state. Minority groups, however, should enjoy further rights to protect group interests, though not all minorities can plausibly claim the same kinds of rights. Immigrant minorities should be granted legal protections to help them preserve some of their traditions and customs—Sikhs should be exempted from laws that might require them to remove their turbans, for example, and some groups should be granted rights and resources to help them preserve their languages—but cannot demand the right to govern themselves. They can only justifiably ask for "polyethnic rights." National minorities, however, should be granted a further set of rights of self-government, for they are in a different position to recent immigrants. They have governed themselves in the past, before being incorporated into the nation or the federation. These groups, which include Puerto Ricans, Chicanos and Hawaiians in the USA, Quebecois in Canada, and Maori in New

Zealand, possess what Kymlicka calls a "societal culture." These people were involuntarily incorporated and so have a legitimate claim to self-government, while recent immigrants have moved voluntarily and so cannot make any such demand. Both kinds of groups might in some circumstances be entitled to "special representation rights," if it would be advisable to make special provision to ensure that otherwise under-represented minorities take their place in governmental institutions. But only national minorities can be self-governing within a liberal multicultural framework.

Kymlicka's theory has been subjected to voluminous and searching criticism. The very emphasis on cultural groups as bearers of rights has been questioned from the outset (Kukathas 1992a, b; Phillips 2007). The distinction between ethnic groups and national minorities has also been viewed as problematic (Benhabib 2002: 62–3; Kukathas 2003: 80–4; Festenstein 2005: 70–84), as has been his notion of a "societal culture" (Nathan 2010: 25–30). His classification of cultural rights has been scrutinized and revised (Levy 2000: 125–60). The success of his theory in protecting cultural minorities rather than hastening their demise has also been called into question (Spinner 1994: 103–4). Nonetheless, Kymlicka's theory has been influential, not only because it has been the starting point for philosophical discussions of multiculturalism but also because many liberal theorists have turned out to be broadly sympathetic to his viewpoint.

The most important distinction Kymlicka makes in developing his theory of minority rights is the distinction between two kinds of protection. Groups, according to Kymlicka, ought to be assured of *external* protections against the power of the majority in society, but cannot also claim the right to control the behavior of their members by placing *internal restrictions* on individual conduct. External protections promote fairness between groups, but internal restrictions "limit the right of group members to question and revise traditional authorities and practices" (Kymlicka 1995: 152). One obvious concern this view raises is that it does not give minority groups the kind of autonomy his treaty model promises. If groups wish to maintain illiberal practices (say, by restricting religious liberty or denying education to girls) they would get no support from Kymlicka's doctrine. Given his general concern to protect minority groups from intervention by the majority society, Kymlicka is reluctant to sanction interference in illiberal communities; but equally reluctant to give theoretical succor to illiberalism.

In spite of the tensions in Kymlicka's theory, it remains a view that is widely embraced since political philosophers who are sympathetic to the claims of cultural minorities are generally reluctant to tolerate illiberal groups. Some, however, have objected vigorously to granting any kind of protection to illiberal minorities and reject group-differentiated rights altogether.

3. Liberalism without Multiculturalism

Among Kymlicka's most vigorous critics are liberals who reject multiculturalism as an idea that is inconsistent with egalitarianism. Kymlicka had defended his version of multiculturalism as a liberal theory of equality that had sought to put individuals on a more equal footing by recognizing that group membership had a vital effect on people's relative standing. Political theorists such as Brian Barry, however, argued that pandering to cultural groups ultimately undermined equality while giving unwarranted protection to various forms of illiberalism (Barry 2001). Groups should be neither granted external protections nor allowed to impose internal restrictions on their members. Individuals have individual rights, and should be protected in the possession of those—and those alone.

For Barry, treating people equally means treating individuals equally. In a liberal polity, this means that the state acts as a neutral or impartial agent that does not favor one community, culture, language or religion over any other. Laws should apply to all equally, without exemptions for any group or people. If some people wish to associate in groups (or remain in groups to which they belong), they should be free to do so; but can no more demand public support for this preference than a citizen can demand special privileges or a greater share of society's resources because he has expensive tastes. There is nothing special about one's taste for one's culture (Barry 2001: 40, 2002: 215; Quong 2006).

On this view, then, there is no warrant for exempting Sikhs from motorcycle helmet laws, for permitting Amish parents to take their children out of school at a younger age, for tolerating halal and kosher butchery, or for subsidizing religious groups either directly or through tax relief. The liberalism of John Stuart Mill had no need for groups, Barry insists, and neither should liberalism today. Multiculturalism comes not to fulfill liberalism but to ruin it.

Barry is not hostile to the people existing in, or coming together to form, groups. He does not, however, wish either to subsidize them or permit them to operate under different rules. People should certainly be free to associate; but they should also be free to leave their groups, which are obliged not to place obstacles in the way of dissenters who wish to leave. To the extent that groups make it difficult for people to leave because their customs leave their members ill-equipped for life in the outside world, the groups should bear the costs of helping those members make their way into life beyond their cultural community. Groups should bear all the consequences of their beliefs (Jones 1994). For example, if they cannot tolerate meat that is not from animals that are cruelly slaughtered, then they should turn vegetarian and not eat meat at all. They cannot demand exemption from the laws governing the humane treatment of animals.

Barry's hostility to the very idea of exemptions for cultural groups is difficult to understand. Law is riddled with exemptions for all kinds of people (such as the disabled) and circumstances, and Barry has rightly been criticized for his blanket condemnation of the rule and exemption approach (Festenstein 2005: 111–13). His insistence that groups bear the costs of enabling members to exit is also questionable to the extent that this simply transfers power from some elements of the group to others (Kukathas 2002). Nonetheless, Barry seems to be on strong ground when he argues that multiculturalism involves a departure from liberal equality inasmuch as it offers protection or succor to groups whose practices are internally illiberal. How, he asks, can liberals condone groups whose very existence sets back liberal gains in securing rights for women and the protection of the vulnerable, including children? The claims neither of religion nor of culture provide sufficient warrant for this.

This view is expressed with equal vigor by some feminist critics of multiculturalism, who, like Barry, argue that there is no justification for granting groups external protections, or tolerating their internal restrictions on their members. Susan Okin, the most prominent of these critics, insists that since most cultures are in fact unsympathetic to the interests of women, protecting such cultures to help them flourish means trying to preserve the conditions against which feminism has been struggling (Okin 1998, 1999, 2002). Far from encouraging or subsidizing such groups, the liberal state should look to re-educating their members and, if necessary, punishing them if their practices seriously disadvantage women. Indeed she courted controversy by suggesting that it would be better if some groups that were particularly harmful to women became "extinct" (Okin

1999: 23), even if ideally they should be transformed and integrated into a less sexist surrounding culture.

Okin, like Barry, takes issue with two alternative perspectives: the Kymlicka view, which favors group protection along with measures to foster some cultural conformity, and one prominent alternative, which rejects the idea of group protection but nonetheless eschews enforced conformity. (See in particular Kukathas 2003.) Against those who favor group protection, Okin maintains that multiculturalism and women's interests are fundamentally in tension, since "most cultures are patriarchal … and many (though not all) of the cultural minorities that claim group rights are more patriarchal than the surrounding cultures" (Okin 1999: 17). Though many of her critics have tried to deny or play down the extent of the tension (Kymlicka 1999; Honig 1999; Parekh 2002), or criticized her for failing to appreciate the possibility of reconciliation (Shachar 2001), it seems clear that, at the very least, multiculturalism and feminism are in conflict as doctrines or theories (Kukathas 2001; Okin 2005). From the perspective of liberals such as Barry or feminists such as Okin, to protect illiberal or patriarchal cultures in the name of multiculturalism is to abandon the egalitarian ideals for which they have fought. The disagreement here is between competing understandings of equality. Multiculturalists like Kymlicka defend this idea as a way of securing equality by addressing the disadvantages suffered by cultural minorities, while critics such as Barry cannot see how this effort to cleanse egalitarianism might be accomplished without throwing the liberal baby out with the bathwater.

This contest between competing liberal egalitarianisms is in evidence in a number of issues that have dominated debates about multiculturalism. The broadest of these is the general question of whether advocates of multiculturalism have failed to appreciate the virtues (and perhaps the inescapability in the modern world) of cosmopolitanism (Waldron 1997). Some have suggested that group attachment and loyalty is positively dangerous, since a preoccupation with group identity is often the source of violence (Sen 2006). Others have looked for ways of upholding cosmopolitan ideals while giving some recognition to the concerns of minority cultures (Appiah 2005, 2007), but the issue has manifested itself in debates about a number of more particular questions. The most important of these have been about the claims of religion in society, the education and rearing of children, and attire. These questions are interrelated, most obviously because the claims of religious communities or traditions have become contentious when schooling or apparel have been at issue.

Among those who defend liberalism, most theorists view these issues in the way that Barry and Okin have, offering religion limited protection and unwilling to compromise principles of gender equality. Diversity might be an important good, and the protection of religious freedom a substantial consideration for any liberal society, but this does not justify permitting religious conviction to trump liberal principle. So, for example, Stephen Macedo argues that liberal states should not shy away from confronting religion and transforming their own citizens in order to inculcate the values needed for modern democracy to function effectively—even if it is important to be mindful of the dangers of suppressing difference (Macedo 2000). Education, ultimately, must be education for citizenship (Callan 1997; Reich 2002; Arneson and Shapiro 1996).

The important tension here is between the cosmopolitan and secularist outlook of modern liberalism on the one hand, and the hopes of traditional communities or cultures striving to hold on to practices or way of life that are out of step, or simply out of favor, in contemporary society. Thus, while ordinarily matters of individual attire would

not be regarded as a matter of anything more than purely private concern in a liberal society, in culturally diverse societies they have become the focus of intense debate. Whether Muslim women should wear headscarves in public is an issue because secular liberals regard this form of dress as a way of preserving inequality between the sexes in minority communities (and, so, in society more broadly). In France it is further viewed as an affront to the state's efforts to preserve a broadly secular public culture (Bowen 2007). For liberal political theorists the question is whether to view this simply as evidence of the on-going and ever-present conflict between tradition and modernity, or suggest a deeper development in modernity: the emergence of a cosmopolitan ethic that is set to dominate moral thinking across the globe (Benhabib 2008; Waldron 2008).

The cosmopolitan stance has, however, to confront the two critical groups whose concerns lie at the heart of the problem of multiculturalism: indigenous peoples and immigrants. While indigenous groups have had to deal with the problem of adjusting to the modern world, and the cosmopolitan ethic that is otherwise so unfamiliar to them, acquiring this new moral language has given them an opportunity to assert new moral claims. While Plenty Coups's concern was primarily to try to find a way of enabling the Crow to adapt to modernity, modern Native Americans have a vocabulary that enables them to make demands for justice. The cosmopolitan ethic of universal rights and equality made it evident to many indigenous peoples that they had not received just treatment, and that they had legitimate claims for redress. One cosmopolitan response has been to say that, whatever the wrongs suffered by conquered peoples, the past cannot be undone and historical injustice has been superseded by current concerns (Waldron 1992). And indeed if one looks at the matter from the perspective of individuals' rights, it may be difficult to recognize claims by current generations for compensation for wrongs committed by earlier ones (Sher 1981; Kukathas 2006b). From the indigenous people's perspective, however, this may mean there's all the more reason to incorporate group rights into cosmopolitan moral theory (Thompson 2002; Ivison *et al.* 2000; Tully 1995).

If the cosmopolitan rejects group rights, or group thinking more generally, however, immigration becomes a pressing moral issue. Liberals might well insist that immigrant minorities should be required to conform to the universal, cosmopolitan norms of society. But it then becomes necessary to explain why the liberal state is justified in restricting immigration and trying to preserve the integrity of the nation-state. Will Kymlicka, as a defender of group rights, has a ready answer: because cultural groups and national groups are alike in being types of group, and group protection is a worthy aim, as well as one that is entirely consistent with liberalism (Kymlicka 1995). It's not clear that Brian Barry or Stephen Macedo could so readily defend immigration restrictions if their cosmopolitan perspective rejects group-differentiated rights (Macedo 2007; Barry 1992). It might be possible to defend the idea of cosmopolitanism in one country (or separate countries) as a coherent ethical ideal, but the ambition has a distinctly illiberal feel about it, much as the idea of separate but equal did in the matter of racial segregation in American education (Dumitru 2011).

Kymlicka identified a significant difficulty for modern liberals when he realized that the focus on individual rights brought with it difficulties when trying to cope with cultural diversity that had yet to be fully appreciated. A part of the source of this problem was the aspiration to marry the individualist outlook with at least a modest nationalism. Some critics of liberalism, however, have suggested that it is the individualism that is given too much weight. This brings us to the third conception of multiculturalism.

4. Multiculturalism with a Difference

Although liberalism, as a product of the European Enlightenment, has much to commend it as a way of thinking that brought the ideals of freedom and equality to the fore, it is nonetheless important, according to Charles Taylor, to understand its limitations. Such an understanding would reveal, among other things, why Kymlicka's attempt to rescue liberalism remains inadequate. Taylor offers his critique in his extended essay, *Multiculturalism and the Politics of Recognition* (Taylor 1994).

According to Taylor's (much discussed) paper, liberalism's triumph lay in its establishment of the idea of human dignity as expressed in the form of universal recognition. Each person was to be viewed as a human being no different from any other, bearing a claim to recognition as an equal and as a possessor of the dignity appropriate to an autonomous individual. Yet while this politics of equal recognition was an improvement on the earlier politics that tied each person to a particular station in life, its limitations eventually became apparent. It was all very well to view people as equals and to insist on equal treatment on the grounds that all human beings were cut from the same cloth and the bearers of the same, universally valid, individual rights. The reality, however, is that people are not the same, but different. Moreover, they care about those differences, which stem ultimately from their attachments, their memberships in particular groups or cultural communities or traditions that give them their identities (Taylor 1994; Baumeister 2000: 136–47).

From this perspective, the problem with liberalism, even as modified by Kymlicka, is that it does not appreciate sufficiently the depth and importance of cultural attachments or, consequently, how inadequate is the theory of liberal multiculturalism. It is indeed important that groups be given cultural protection, but this alone is not enough for it fails to consider the significance of the longing of groups for cultural survival. Society, it turns out, is made up of all kinds of groups whose integrity is worth protecting, beginning with the group that is the nation-state. But the integrity of the state cannot be a reason to sacrifice, or even neglect, those other groups whose survival is integral to the well-being of their members: culturally distinct provinces, such as Quebec, or peoples such as the Inuit. The politics of recognition must give way to a politics of difference.

This line of thought shares much with Iris Young's analysis in *Justice and the Politics of Difference*, which similarly argues for a rethinking of political institutions to accommodate the plurality of identities found in modern liberal democracies (Young 1990). Like Taylor, Young is a critic of liberalism. What is crucial from her point of view is that the liberal model of group membership which emphasizes the nature of groups as associations of individuals. Identity is the product of complex forms of interaction, and the parties to this interaction are not merely individuals but persons who are parts of social groups that are themselves constituted by their practices and also their relations with other groups. To sustain a pluralist democratic culture, she argues, it is necessary not only to preserve a system of general rights but also to develop a set of group-conscious policies (Young 1990: 174). This is necessary both to nurture the groups to enable them to take part in the democratic polity but also to address issues of justice that arise out of the oppression suffered by groups that would otherwise be marginalized or assimilated by the dominant, mainstream culture.

This perspective shares a surprising amount with the thought of Joseph Raz, who argues from the standpoint of a liberal perfectionist that it is important that the state act to protect cultural identity and membership. This is the only way that liberalism can in

fact protect moral pluralism (Raz 1994). This is a version of liberalism that denies the primary importance of consent or associative obligations as the foundation of morality or political order but looks for the source of the ethical in an account of the good. In this account the justification of group protection lies not in the wishes of members but in the values that groups preserve (Raz and Margalit 1990).

While this broad approach to multiculturalism emphasizing the importance of groups and the need to appreciate why difference matters has played a substantial role in the development of a democratic theory of multiculturalism (Deveaux 2000, 2007), some questions remain to be answered. The most important of these is how the problem of group conflict is to be addressed when groups are properly understood as fluid and shaped as much by political interaction as by other factors such as ethnicity, or culture or common traditions. It may be very well to be concerned about groups and their well-being, but the clamor for group protection does nothing to resolve the problem of which groups to give succor to, and which groups to let wither. The recognition of difference is constantly sought, but the claimants demand not merely that they be identified as individuals but that they be identified with others—who may or may not be ready to be so labeled or categorized.

This problem is exacerbated by the existence of straightforward political conflicts between groups, notably in some societies between immigrant groups and indigenous peoples. All too often, indigenous communities want recognition as first peoples because of their concern to preserve their own traditions and identities—because of their longing for cultural survival. Immigrant minorities, however, not unusually, are preoccupied with maintaining their customs to some degree, but ensuring that they do not long remain second-class citizens. How this particular dilemma is to be resolved is a troubling problem, but not one that is adequately addressed by the politics of difference.

5. A Multiculturalism of Indifference

This brings us to the fourth alternative among theories of multiculturalism, which is also the one favored by the present author. This is a view that neither embraces multiculturalism nor condemns it. While it shares with Barry and Okin a repudiation of group-differentiated rights, it is unwilling nonetheless to intervene in the traditions or cultural practices of groups. While it might be difficult for states to remove groups from the legal and general institutional structures of political society, groups themselves should be viewed as having no special normative significance, but be regarded merely as associations of individuals whose legitimacy rests on the willingness of its members to remain associated. Such associations should neither be encouraged nor undermined by public authorities, which should not be in the business of trying to determine which groups should prosper and which should be let die. Cultural assimilation of minorities is inescapable, as is the cultural transformation of the majority. The state should not be an active participant in this process, even if its influence is not something that can be altogether eliminated. Perfect neutrality might be an unattainable ideal, but that is no justification for an actively interventionist form of liberal perfectionism (Kukathas 1998, 2003).

This perspective shares with Barry the disdain for groups; but as a consequence it attaches no special significance to the state as a group whose interests deserve special protection or nurturing. It does not favor intervention in the affairs of groups to uphold the state's interests, or to secure egalitarian or libertarian goals more generally. It conceives of political society largely as a realm of mutual toleration, and of liberal society

as a political order that comprises a diversity of groups, many of which might not themselves be liberal in their ways. Political society is thus not a form of association which necessarily amounts to an enduring form of social unity, for it may well be little more than a political settlement that reflects the reality of historical circumstances and the facts of the distribution of political power.

The main attraction of such a view is that it accommodates diversity without becoming hostage to the demands of particular minorities. But to its critics, this position confronts a number of difficulties. From one liberal perspective, the problem is that this view is too willing to tolerate forms of life or practices which are, if not immoral in themselves, at least too far removed from the core commitments of liberalism. If groups within a liberal society can, without penalty, compromise the freedom and equality of their members, then liberalism has succumbed to multiculturalism and failed to live up to its own fundamental convictions.

A deeper problem, however, is that this view seems not to take sufficiently seriously the importance of the political construction of the public sphere in which the groups in question associate. That sphere, the argument goes, cannot be merely an accidental meeting ground of different ethical or cultural views or communities. It is the product of political thinking and negotiation, and the care of this structure is an important public philosophical concern.

It is at this juncture that theories of multiculturalism are forced to engage the deepest questions of political theory. Most prominent among these is the very abstract question of how can the many live as one. What are the institutions that make for a good polity? While multiculturalism has provoked many theorists to confront precisely these questions, it perhaps remains to be seen whether more satisfying answers will emerge out of these engagements.

Related Topics

Liberalism, Feminism, Pluralism, Libertarianism, Nationalism, Human Rights and Cosmopolitanism, Toleration, Education, Marriage, Sex and the Family, Religion in Public Life

References

Appiah, K. A. (2005) *The Ethics of Identity*, Princeton, NJ: Princeton University Press.

—— (2007) *Cosmopolitanism: Ethics in a World of Strangers*, New York: Norton.

Arneson, R. and Shapiro, I. (1996) "In Defense of Yoder: Parental Authority and the Public Schools", in I. Shapiro and R. Hardin (eds.), *Political Order*: NOMOS XXXVIII, New York: New York University Press.

Barry, B. (1992) "The Quest for Consistency: a Skeptical View," in B. Barry and R. Goodin (eds.), *Free Movement: Ethical Issues in the Transnational Migration of People and Money*, University Park: University of Pennsylvania Press.

—— (2001) *Culture and Equality: An Egalitarian Critique of Multiculturalism*, Oxford: Polity.

—— (2002) "Second Thoughts—And Some First Thoughts Revived," in Paul Kelly (ed.), *Multiculturalism Reconsidered*, Oxford: Polity, pp. 204–38.

Baumeister, A. (2000) *Liberalism and the "Politics of Difference,"* Edinburgh: Edinburgh University Press.

Benhabib, S. (2002) *The Claims of Culture: Equality and Diversity in the Global Era*, Princeton, NJ: Princeton University Press.

—— (2008) *Another Cosmopolitanism*, ed. Robert Post, Oxford: Oxford University Press.

Bowen, J. (2007) *Why the French Don't Like Headscarves: Islam, the State, and Public Space*, Princeton, NJ: Princeton University Press.

Callan, E. (1997) *Educating Citizens: Political Education and Liberal Democracy*, Oxford: Oxford University Press.

Deveaux, M. (2000) *Cultural Pluralism and Dilemmas of Justice*, Cornell, NY: Cornell University Press.
—— (2007) *Gender and Justice in Multicultural States*, New York: Oxford University Press.
Dumitru, S. (2011) "Migration and Equality: Should Citizenship Levy Be a Tax or a Fine?" *Ateliers de L'Ethique* 6(1).
Festenstein, M. (2005) *Negotiating Diversity: Culture, Deliberation, Trust*, Oxford: Polity.
Hanke, L. (1974) *All Mankind is One: A Study of the Disputation Between Bartholomé de Las Casas and Juan Ginés de Sepúlveda on the Religious and Intellectual Capacity of the American Indians*, De Kalb: University of Illinois Press.
—— (2002) *The Spanish Struggle for Justice in the Conquest of America*, Dallas, TX: Southern Methodist University Press.
Honig, B. (1999) "'My Culture Made Me Do It'," in J. Cohen, M. Howard and M. Nussbaum (eds.), *Is Multiculturalism Bad for Women*, Princeton, NJ: Princeton University Press, pp. 35–40.
Ivison, D., Patton, P., and Saunders, W. eds. (2000) *Political Theory and the Rights of Indigenous Peoples*, Cambridge: Cambridge University Press.
Jones, P. (1994) "Bearing the Consequences of Belief," *Journal of Political Philosophy* 2(1): 24–43.
Kukathas, C. (1992a) "Are There any Cultural Rights?" *Political Theory* 20(1): 105–29.
—— (1992b) "Cultural Rights Again: A Rejoinder to Kymlicka," *Political Theory* 20(4): 674–80.
—— (1998) "Liberalism and Multiculturalism: The Politics of Indifference," *Political Theory* 26(5): 686–99.
—— (2001) "Is Feminism Bad for Multiculturalism?" *Public Affairs Quarterly* 5(2): 83–98.
—— (2002) "The Life of Brian, or Now for Something Completely Difference-blind," in Paul Kelly (ed.), *Multiculturalism Reconsidered: Culture and Equality and Its Critics*, London: Polity, pp. 184–203.
—— (2003) *The Liberal Archipelago: A Theory of Diversity and Freedom*, Oxford: Oxford University Press.
—— (2006a) "Moral Universalism and Cultural Difference," in J. Dryzek, B. Honig, and A. Phillips (eds.), *The Oxford Handbook of Political Theory*, Oxford: Oxford University Press, pp. 581–98.
—— (2006b) "Whom? Whom? Reparations and the Problem of Agency," *Journal of Social Philosophy* 37(3): 330–41.
Kymlicka, W. (1989) *Liberalism, Community and Culture*, Oxford: Oxford University Press.
—— (1995) *Multicultural Citizenship: A Liberal Theory of Minority Rights*, Oxford: Oxford University Press.
—— (1999) "Liberal Complacencies," in J. Cohen, M. Howard and M. Nussbaum (eds.), *Is Multiculturalism Bad for Women*, Princeton, NJ: Princeton University Press, pp. 31–4.
—— (2001) *Politics in the Vernacular*, Oxford: Oxford University Press.
—— (2007) *Multicultural Odysseys: Navigating the New International Politics of Diversity*, Oxford: Oxford University Press.
Las Casas, B. (1999) *In Defense of the Indians*, edited and translated by Stafford Poole, De Kalb: Northern Illinois University Press.
—— (2003) *An Account, Much Abbreviated, of the Destruction of the Indies*, ed. F. W. Knight, trans. A. Hurley, Indianapolis: Hackett.
Lear, J. (2006) *Radical Hope: Ethics in the Face of Cultural Devastation*, Cambridge, MA: Harvard University Press.
Levy, J. (2000) *The Multiculturalism of Fear*, Oxford: Oxford University Press.
Linderman, F. (2002) *Plenty Coups, Chief of the Crows*, Lincoln: University of Nebraska Press.
Macedo, S. (2000) *Diversity and Distrust: Civic Education in a Multicultural Society*, Cambridge, MA: Harvard University Press.
—— (2007) "The Moral Dilemma of U.S. Immigration Policy: Open Borders Versus Social Justice," in C. Swain (ed.), *Debating Immigration*, Cambridge: Cambridge University Press.
Nathan, G. (2010) *Social Freedom in a Multicultural State: Towards a Theory of Intercultural Justice*, Basingstoke: Palgrave Macmillan.
Okin, S. (1998) "Feminism and Multiculturalism: Some Tensions," *Ethics* 108: 661–84.
—— (1999) "Is Multiculturalism Bad for Women?" in J. Cohen, M. Howard and M. Nussbaum (eds.), *Is Multiculturalism Bad for Women*, Princeton, NJ: Princeton University Press, pp. 9–24.
—— (2002) "'Mistresses of their own Destiny': Group Rights, Gender, and Realistic Rights of Exit," *Ethics* 112(2): 205–30.
—— (2005) "Multiculturalism and Feminism: No Simple Question, No Simple Answers," in A. Eisenberg and J. Spinner-Halev (eds.), *Minorities within Minorities: Equality, Rights and Diversity*, Cambridge: Cambridge University Press, pp. 67–89.
Parekh, B. (2002) *Rethinking Multiculturalism: Cultural Diversity and Political Theory*, Cambridge, MA: Harvard University Press.

Phillips, A. (2007) *Multiculturalism without Culture*, Princeton, NJ: Princeton University Press.
Quong, J. (2006) "Cultural Exemptions, Expensive Tastes, and Equal Opportunities," *Journal of Applied Philosophy* 23(1): 53–71.
Raz, J. (1994) "Multiculturalism: a Liberal Perspective," in his *Ethics in the Public Domain: Essays in the Morality of Law and Politics*, Oxford: Oxford University Press.
—— and Margalit, A. (1990) "National Self-determination," *Journal of Philosophy* 87(9): 439–61.
Reich, R. (2002) *Bridging Liberalism and Multiculturalism in American Education*, Chicago, IL: University of Chicago Press.
Sen, A. (2006) *Identity and Violence: The Illusion of Destiny*, New York: Norton.
Shachar, A. (2001), *Multicultural Jurisdictions: Cultural Differences and Women's Rights*, Cambridge: Cambridge University Press.
Sher, G. (1981) "Ancient Wrongs and Modern Rights," *Philosophy and Public Affairs* 10: 3–17.
Spinner, J. (1994) *The Boundaries of Citizenship: Race, Ethnicity and Nationality in the Liberal State*, Baltimore, MD and London: Johns Hopkins University Press.
Tamir, Y. (1990), *Liberal Nationalism*, Princeton, NJ: Princeton University Press.
Taylor, C. (1994) *Multiculturalism and the "Politics of Recognition,"* ed. Amy Gutmann, Princeton, NJ: Princeton University Press.
Thompson, J. (2002) *Taking Responsibility for the Past: Reparation and Historic Injustice*, Cambridge: Polity.
Tully, J. (1995) *Strange Multiplicity: Constitutionalism in an Age of Diversity*, Cambridge: Cambridge University Press.
Vitoria, F. (1991a) "On Dietary Laws," in Anthony Pagden and Jeremy Lawrance (eds.), *Political Writings*, Cambridge: Cambridge University Press, pp. 205–30.
—— (1991b) "On The American Indians," in Anthony Pagden and Jeremy Lawrance (eds.), *Political Writings*, Cambridge: Cambridge University Press, pp. 231–92.
Waldron, J. (1992) "Superseding historic injustice," *Ethics* 103(1): 4–28.
—— (1997) "The Cosmopolitan Alternative," in W. Kymlicka (ed.), *The Rights of Minority Cultures*, Oxford: Oxford University Press, pp. 93–119.
—— (2008) "Cosmopolitan Norms," in S. Benhabib, *Another Cosmopolitanism*, ed. R. Post, Oxford: Oxford University Press, pp. 83–101.
Young, I. M. (1990) *Justice and the Politics of Difference*, Princeton, NJ: Princeton University Press.

Further Reading

Laden, A., and Owen, D. eds. (2007) *Multiculturalism and Political Theory*, Cambridge: Cambridge University Press. An important collection of papers by theorists of multiculturalism.

Modood, T. (2007) *Multiculturalism*, Cambridge: Polity. A brief but broad-ranging account of the problem of multiculturalism viewed through the lens of contemporary issues in public policy, which offers a critique of some forms of liberalism and a defense of multiculturalism.

Renteln, A. (2004) *The Cultural Defense*, Oxford: Oxford University Press. An examination of the use of cultural arguments in the courtroom, analyzing claims for recognition of cultural difference with respect to homicide, children, marriage, animals, drug use, attire, and the treatment of the dead.

Sterba, J. (2001) *Three Challenges to Ethics: Environmentalism, Feminism and Multiculturalism*, Oxford: Oxford University Press. Contains a useful analysis of the way in which western ethics has neglected the concerns of non-western cultures.

Swain, L. (2006) *The Liberal Conscience: Politics and Principle in a World of Religious Pluralism*, New York: Columbia University Press. An important study of the tension between liberalism and theocracy which goes to the heart to many philosophical problems of multiculturalism.

45
GLOBAL JUSTICE AND POLITICS

Thom Brooks

1. Introduction

The history of political philosophy has been largely focused on the problem of justice within borders. Contemporary political philosophers have only begun more recently to draw greater attention to problems of global justice rather than to domestic justice alone: they are concerned about identifying a just *international* distributive justice (see Brooks 2008: sect. 7). The most important issue has been how best to address severe poverty. Are there duties to provide support for those in severe poverty and, if so, who has these duties? What support may be justified? These are the most pressing and challenging questions confronting political philosophers today.

This chapter examines three different approaches to how global justice and politics might address the problem of severe poverty. The first approach argues that we have positive duties to assist those in need. Its advocates argue that we have a duty to assist where there are others in need irrespective of whether or not we contributed to their situation. A second approach claims that we have negative duties to those in need that arise because we have contributed to their severe poverty. Finally, a third approach argues that our responsibilities to those in need are not a matter of choosing between our positive or negative duties, but that these duties should be understood within a wider context of our remedial responsibilities. Our focus should be on identifying who has a responsibility to remedy suffering elsewhere and this requires a wider perspective to cover all cases. Each approach is considered in turn in a sympathetic analysis where the focus is on presenting each in its best light and allowing readers to judge for themselves which is most worth defending.

2. The First Approach: Our Positive Duties

The first approach to severe poverty is to argue that we have a *positive* duty to assist those in severe poverty because they are in need. It does not matter if we have contributed to their being in need: it only matters that others have such needs. There are two different ways to understand our positive duties. The first is developed by Peter Singer and involves weighing the relevant potential moral costs of action versus inaction. The second is the capabilities approach developed by Amartya Sen and Martha Nussbaum, which argues that all human beings should have the ability to exercise capabilities

irrespective of where they live. Severe poverty should be eradicated because it prevents our ability to freely exercise our capabilities. I will take each approach in turn.

2.1 Peter Singer on Positive Duties

Singer offers one influential defense of positive duties in a famous essay published in 1972. He argues that the problem of global justice and severe poverty is that often the most pressing problems seem too distant to relate to us. It might be uncontroversial that people elsewhere are suffering great hardship, but it could be unclear what duties we might have to others far away and perhaps in unfamiliar parts of the world. Singer powerfully criticizes the supposition that the distance between you and me has moral significance. Singer argues that it is not the space between us that has significance, but rather the relative moral weights of our possible actions. What matters is the balance of moral costs of rescuer and rescued and not the space between them. Singer says (1972: 231): "if it is in our power to prevent something bad from happening, without thereby sacrificing anything of comparable moral importance, we ought, morally, to do it." If we can prevent a bad to others, then we should where the moral costs to ourselves in acting are outweighed by the moral costs to another if we failed to act. The idea is that we should aim to prevent what is bad where we are able to prevent it, not that we should promote what is good. Therefore, the positive duty is a duty to prevent what is bad where possible because we morally ought to do so. The distance between us has significance only if it prevents me from preventing what is bad. Otherwise, distance lacks moral significance.

Singer provides an illustration of how this principle applies to a case of a drowning child. He says (1972: 231):

> If I am walking past a shallow pond and see a child drowning in it, I ought to wade in and pull the child out. This will mean getting my clothes muddy, but this is insignificant, while the death of the child would presumably be a very bad thing.

Where I am able to act, the moral weight of choosing to act is trumped by the moral weight of refusing to act. This example is simplified and assumes that I am able to act, the child can be saved, and the child is unable to save herself. Where these assumptions hold, I morally ought to act and save the drowning child.

The idea of moral weighting receives special importance. Singer's example of the drowning child weighs up moral costs on two sides: the first is the moral cost to me in providing rescue and the second is the moral cost to the child in my failure to provide rescue. We should note that the *material* cost to me in providing rescue has little significance. While Singer claims I morally ought to muddy my clothes to save the drowning child, it is of little moral significance whether my muddy clothes are, in fact, an old shirt and jeans or an expensive tuxedo instead. What matters is the *moral* cost to me and this cost may only be the inconvenience involved. If we altered the example so that I am running to help free people from a burning house when I notice the drowning child, we might believe that perhaps my positive duties to more people could outweigh my positive duty to the one drowning child. Such situations are thankfully rare. The main idea is that the moral cost to the drowning child is high if I fail to act: the child will drown. Where this cost to the potential victim is higher than the moral cost to the potential

rescuer, then the rescuer morally ought to save the victim. This may come at significant material cost to the rescuer if this does not much impact on the relevant moral cost.

Singer's illustration is meant to show that we morally ought to save the drowning child where the moral cost to the child in our refusing to act is greater than the moral cost to ourselves if we choose to act. His purpose is to offer a compelling argument about severe poverty: if the moral cost of providing relief for us is less than the moral cost to others if we failed to do so, then we morally ought to provide relief. It does not matter morally that we have not contributed to the suffering of others. Likewise, we should save the drowning child and it does not matter morally why the child is drowning in the first place nor who may have thrown the child in the pond. We should end severe poverty because it exists, it is a bad, and the relative moral cost to ourselves is easily outweighed by the moral cost to others in severe poverty should we fail to act, as this failure could bring widespread, and preventable, suffering. Nor does it matter, morally speaking, that many others are similarly situated: if anything, the moral cost may be relatively trivial to citizens of affluent states to assist those in severe poverty elsewhere. This is one view of positive duties and severe poverty on offer.

2.2 The Capabilities Approach and Positive Duties

A second understanding of positive duties for assisting others in severe poverty is the capabilities approach. This approach was first developed by the Nobel Prize winning economist Amartya Sen and given greater philosophical sophistication by Martha Nussbaum. While the approach has exploded in popular interest over recent years with many different formulations, my focus will be limited to the accounts by Sen and Nussbaum because they are the most widely accepted and discussed.

The capabilities approach originally aimed to address a particular problem. How to best understand development? A common approach had been to understand development through improving GDP per capita. The idea was that a country was making progress in terms of its development as its average wealth rose. Sen (1999) raises many important criticisms. One criticism is that an increase in average wealth does not always translate to improved material conditions for most in a population. This is because a rise in per capita wealth may mask large inequalities of wealth: the average wealth might be increasing, but the wealth of most in the population could be decreasing. The GDP per capita approach fails to recognize this problem. Sen has a second criticism: the GDP per capita approach compares the development of all societies along a shared continuum, their relative wealth per capita. The problem is that different societies will have different needs at different costs. Our focus should turn to why we believe GDP per capita a good indicator: that it might offer a useful approximation of a society's improvement on measures such as life expectancy, low infant mortality, education, and other indicators. We should recognize that GDP per capita is too crude a device to make any meaningful comparisons or analysis. Instead, we should focus on the *capability* of people in a society rather than purely on their relative material wealth.

Our capabilities should be understood in the following way and we start with the distinction between actual functioning and capabilities. Consider the difference between a man who is fasting and a second man who is starving. Both men share the same actual functioning, that is, both lack food. There remains a crucial difference in terms of their capabilities, however: the man who is fasting possesses the capability to eat, but the starving man lacks this capability. The capability approach is focused on our freedom

to exercise different kinds of opportunities relating to our well-being. The idea is that this focus draws our attention to the best moral considerations of whether well-being is improving or not. The question is not simply whether a society has certain material conditions, but whether its people may have freedom of choice in relation to their capabilities. Sen's classic illustration (1983) is that modern famines have never taken place in a democracy. Famine is a problem about political failure as every famine has happened in countries where there was sufficient food that was not best distributed to avert a famine taking place. It is not that a state has certain material conditions that matters most, but the freedom of its people to exercise their capabilities.

Martha Nussbaum (2003) has developed Sen's account in at least three important respects. The first is that Sen often speaks of our "capability to *x*" whereas Nussbaum refers to our "capabilities to *x*" (where *x* means to do or be). The idea is that we have multiple capabilities that demand recognition and protection and, therefore, we should use capabilities in the plural rather than the singular. Each capability is understood in a non-competitive way: we ought not endorse any trade-offs between them and all should be satisfied. This is an ideal and trade-offs might not be avoidable. A second development is that Nussbaum defends the capabilities approach as setting out a bottom threshold, not entirely unlike the idea of the "social minimum" that we find in John Rawls but in terms of capabilities and not resources (Rawls 1999: 251–2). The capabilities approach sets a threshold below which people are unable to enjoy satisfactory well-being. There is also a link made between our capabilities and human rights as both demarcate fundamental areas of great importance for us and each demands protection.

The third development that Nussbaum offers in defense of the capabilities approach is perhaps the most controversial. Many proponents refer only to our "capability" to *x* and they actively avoid offering any particular list, although some proponents, such as Sen, reject the use of lists. Their view is that our understanding of capabilities must be alive to present conditions and, thus, open to revision and change. Nussbaum accepts the view that capabilities are open to revision and reinterpretation, but rejects the view that a list lacks any useful function. Nussbaum (2011: 33–4) argues that there are ten capabilities: life; bodily health; bodily integrity; senses, imagination and thought; emotions; practical reason; affiliation: other species play; control over one's environment (political and material). Nussbaum says (2011: 36) that "[t]he basic claim of my account of social justice is this: respect for human dignity requires that citizens be placed above an ample (specified) threshold of capability, in all ten of those areas." The ten capabilities demarcate the social minimum that all persons should surpass in order to possess their human dignity. Capabilities are goals that people are "entitled" to and governments must ensure this standard is met to pass as "minimally just" (Nussbaum 2011: 169).

The capabilities approach perspective on severe poverty states that all governments must ensure their population may choose to satisfy their capabilities. Moreover, these capabilities address fundamental human rights and freedoms and, as such, they are universally valid. Severe poverty is best addressed as a problem of failure to guarantee capabilities and not merely a failure to satisfy bare material needs. Human dignity requires more than food, but choice: recall that the capabilities approach focuses on capability opportunities and not actual functioning. The problem of severe poverty requires action to ensure the human dignity of all is recognized and protected. It does not matter that we bear no responsibility for the misfortune of others, but it does matter that all governments respect human dignity in order to be minimally just. There is a positive duty on all to ensure that governments respect the capabilities of all.

3. The Second Approach: Our Negative Duties

A second approach to thinking about severe poverty is to say that we have *negative* duties to those in severe poverty in relation to our contributing to their situation. The view is that this is a more stringent understanding of duties arising from the harm to others that we have contributed toward and not any promotion of the good. Negative duties ground our addressing harms because we have contributed to them rather than because there is need. Proponents argue that this view of duties is also more compelling: there may be greater intuitive weight to the view that I have a duty because I have contributed to some harm rather than I have a duty because others have a need, but where I have not contributed to the problem.

The best-known formulation for the negative duty to address severe poverty is offered by Thomas Pogge (2002), who argues that our duties are not positive because they do not arise principally because of the bare fact that others are in severe poverty. Instead, Pogge argues our duties to those in severe poverty are grounded in our responsibility for their condition. Thus, severe poverty should be eradicated because we are responsible for its continuation and not merely because we can bring it to an end (2002: 201–3). Furthermore, severe poverty should be further understood as a violation of human rights. For example, the Universal Declaration of Human Rights states in Article III that all persons have a right to the secure access of "minimally adequate shares of all basic goods" (quoted in Pogge 2002: 38). The existence of severe poverty is evidence that such minimally adequate shares are not provided; this is evidence of a human rights violation because this can be attributed to political institutions. Pogge says (2002: 172; see Brooks 2007): "We are asked to be concerned about avoidably unfulfilled human rights not simply as they exist at all, but only insofar as they are produced by coercive social institutions in whose imposition we are involved."

The global institutional order does not merely permit severe poverty, but it engenders it. Pogge (2002: 18–19) offers several examples. The first example is international economic bodies, such as the World Trade Organization. These bodies have supported the global institutional order by enabling the exacerbation of death from global poverty through monetary agreements that favor affluent states at the cost of poor states. A second example is protectionist exemptions insisted upon by affluent states that damage poor states in their ability to trade on more equal terms. These are not the only ways in which the global institutional order is maintained and promoted, but they make clear that affluent states support a global institutional order that engenders severe poverty. This order is foreseeable and avoidable. The effects of protectionist measures are well documented and known in advance. Nor need such measures be in place. Affluent states might not bear all responsibility for severe poverty elsewhere, but they do bear some significant share of responsibility for the global institutional order in which severe poverty is found.

Furthermore, Pogge argues that "we" in affluent states have a negative duty to eradicate severe poverty elsewhere. The argument is that affluent states are predominantly democracies: we have a say in who our political leaders are and the system we uphold. We might not vote for who is selected to serve on different trading bodies, but we do vote for those who make these decisions and to some degree we each have responsibility for the decisions our political leaders make on our behalf as our elected representatives. Therefore, it is not merely true that our state is responsible for severe poverty elsewhere, but rather that we each bear some responsibility as the electorate of these states. Pogge

says (2002: 109): "We, the affluent countries and their citizens, continue to impose a global economic order under which millions avoidably die each year from poverty-related causes … We must regard our imposition of the present global order as a grave injustice." We have a negative duty to stop contributing to this harm and to provide relief to those harmed already.

Our responsibility for engendering severe poverty justifies our negative duty to provide relief. We should assist not simply because we can, but because we bear some responsibility for the plight of others. Pogge argues that the approach of negative duties offers a powerful alternative to the approach of positive duties because negative duties are more stringent and intuitively compelling. Negative duties are more stringent because our object is the removal of not any moral bad, but a bad we have responsibility for. Nor is our object the promotion of any particular good. Instead, our goal is to end the harm we cause others and to provide remedy. Negative duties are thought to be more intuitively compelling because the idea that we should assist *because we are responsible for harming* is thought to carry more weight than the idea that we should assist *because we can*. Whether or not negative duties have such intuitive force is left to readers to reflect upon.

Pogge (2002: 196–215) offers a specific proposal, the Global Resources Dividend, that is meant to show how we might fulfill our negative duties. He argues that (2002: 204) "those who make more extensive use of our planet's resources should compensate those who, involuntarily, use very little." The general idea is that affluent states use the greatest share of global crude oil, but the producers are largely poorer states. Pogge claims that crude oil should be subjected to a small tax of a few US dollars that is paid into a Global Resources Dividend (or GRD) and these resources could be used to eradicate severe poverty. Only a small share of the wealth of affluent states would be required to help eradicate severe poverty and the GRD is one justice-based account that may ground the negative duties of affluent states to compensate those in severe poverty.

The negative duties approach to addressing severe poverty is perhaps the dominant view in the literature and Pogge a major figure in all such debates. Criticisms tend to focus on the desirability and even possibility of the kinds of global institutions that might be required to enforce the negative duties that we have in the absence of any world government (as Pogge rejects the possibility of a world government). These criticisms extend to Pogge's illustration of the GRD: critics argue that it is unclear how the GRD might be collected and distributed, as well as how non-payment might be punishment if necessary (see Fuller 2005). Criticisms have also focused on whether we have all contributed to the global institutional order in the ways that Pogge identifies.

4. The Third Approach: Our Remedial Responsibilities

There is a third alternative to positive and negative duties on severe poverty. This alternative is the view that our focus is on addressing severe poverty and so we should center attention on determining the remedial responsibilities of those who can provide rescue. This view is championed by David Miller.

Miller's perspective is different from what we have examined thus far. Nussbaum, Pogge, Sen, and Singer each speak about the duties of individuals *as individuals* concerning assisting those in severe poverty. Miller differs in that he focuses (2007) on *the nation*, not individuals, as the primary agents of global justice. The nation is a group that must pass certain tests: our co-national relationship is intrinsically valuable, integral to our other relationships, and honoring this moral significance should not entail our

giving less than what is deserved to non-nationals. Our shared national identity may have moral significance, but only if these tests are met. We should recognize that not all forms of common identity will meet these tests. For example, white supremacists lack a relationship that is intrinsically valuable. Likewise, my being a supporter of Bayern Munich or Sheffield United football clubs fails to meet this test. Forms of identity that might meet this test would include our shared identity as American or British, as an illustration. These common bonds might have value integral to our relationships with one another.

The importance of national identity is that it grounds Miller's theory of global justice. His focus is to find some way to determine which *nation* should put right severe poverty elsewhere. He says (2007: 98): "it is morally unacceptable for people to be left in that deprived or needy condition, and there is no overriding justification … All that matters is that we find it morally unacceptable if the deprived person is simply left to suffer." Our goal is to identify who should help those in need.

Miller offers a connection theory of remedial responsibility. Its purpose is to help us decide which nation or nations have a responsibility to provide remedy elsewhere based upon the relevant weight of connections they share with those in need. There are six connections (Miller 2007: 100–4): causal responsibility; moral responsibility; capacity; community; outcome responsibility; benefit. We consider each nation in relation to these six connections and we select appropriate nations for remedial responsibilities based upon the relative weight of the connections they possess. A nation may have greater remedial responsibilities than another nation despite the second nation having more varieties of connections than the first. Miller denies there is any ordering or that any connection has greater weight than others, although it is clear that remedial responsibilities are only possible where a nation has the *capacity* to provide a remedy and, therefore, capacity has a lexical priority in fact. (See Brooks 2011.)

The remedial responsibilities approach focuses on the alleviation of severe poverty wherever it is found irrespective of how it is caused. It does not matter which nation contributed to severe poverty elsewhere, but identifying which nation or nations should remedy it does matter. The remedial responsibilities approach shares this perspective with the positive duties approach in that both claim there are duties to assist those in severe poverty because there is severe poverty. The difference is that the remedial responsibilities approach denies that each nation will have the same duties to assist. The remedial responsibilities approach agrees with proponents of negative duties that some may have greater duties than others. The strength of the remedial responsibilities approach is that it can address severe poverty that could be the result of natural catastrophes where there is no human responsibility for severe poverty. A further strength of the remedial responsibilities approach is that it can help us identify some nations as having greater remedial responsibilities than others based on more than mere moral responsibility for suffering, but on other connections such as their being part of a shared ethnic or linguistic community amongst other factors. The drawback is that this approach rests upon a view of group responsibility and the moral value of co-nationality that some may contest. Additionally, there is no clear guide on how more precisely we might weigh relative connections.

5. Conclusion

This chapter has examined three different approaches to the most pressing problem of global justice and politics, namely, the problem of severe poverty and who, if anyone,

has a duty to provide assistance. We have considered approaches from a wide range of perspectives. The first argues that we have positive duties to those in need. One perspective claims that we have a positive duty to assist others even if we bear no responsibility for their plight where providing assistance is at lower moral cost to ourselves than the moral cost to another should we not act. A second perspective argues that society should recognize, maintain, and preserve our capabilities. We each have a right to exercise certain freedoms, including a freedom from severe poverty. These two perspectives offer one powerful set of arguments in favor of positive duties.

A second approach claims our duties to those in severe poverty should be seen primarily as negative duties. "We" in affluent societies have foreseeably, knowingly, and avoidably supported an international structure that has contributed to severe poverty and its continuation. We should assist others not merely because they are in need, but because we are responsible for their need. Moreover, proponents of negative duties argue that negative duties are more compelling than positive duties: negative duties have greater intuitive weight.

Finally, a third approach argues that we should not choose between positive or negative duties. Both kinds of duties have importance and relevance. Instead, we should focus on our remedial responsibilities, or the responsibilities to assist those in severe poverty. The problem is being best able to identify who should provide assistance to address severe poverty. We can identify relevant groups through weighing their various possible connections to those in severe poverty.

Each approach offers a distinctive and novel perspective for thinking about our duties to those in severe poverty. The eradication of severe poverty is perhaps the most pressing problem in global justice and politics today. While some live in greater luxury than any past generation, others live on the margins of bare existence.

I leave it to readers to decide for themselves whether they find most defensible an approach grounded in positive duties, negative duties or remedial responsibilities—or perhaps some new fourth alternative. All proponents would accept the pressing problem of best addressing severe poverty and the need to determine how this view of global justice should be understood to inform our global politics. While the problem is widely acknowledged, it is far from settled which approach is best. My hope is that this chapter has shed some light on the leading arguments and figures in this important debate, and that the chapter spurs readers to think about these approaches anew. Severe poverty is everywhere in our global midst. If we want it eradicated, then we must rise to the challenge.

Related Topics

The Capability Approach (and Social Justice), Nationalism, Human Rights and Cosmopolitanism

References

Brooks, T. (2007) "Punishing States That Cause Global Poverty," *William Mitchell Law Review* 33: 519–32.
—— (ed.) (2008) *The Global Justice Reader*, Oxford: Blackwell.
—— (2011) "Rethinking Remedial Responsibilities," *Ethics and Global Politics* 4: 195–202.
Fuller, L. L. (2005) "Poverty Relief, Global Institutions, and the Problem of Compliance," *Journal of Moral Philosophy* 2: 285–97.
Miller, D. (2007) *National Responsibility and Global Justice*, Oxford: Oxford University Press.

Nussbaum, M. C. (2003) "Capabilities as Fundamental Entitlements: Sen and Social Justice," *Feminist Economics* 9: 33–59.
—— (2011) *Creating Capabilities: The Human Development Approach*, Cambridge, MA: Belknap/Harvard University Press.
Pogge, T. (2002) *World Poverty and Human Rights: Cosmopolitan Responsibilities and Reforms*, Cambridge: Polity.
Rawls, J. (1999) *A Theory of Justice*, rev. edn, Oxford: Oxford University Press.
Sen, A. (1983) *Poverty and Famines: An Essay on Entitlement and Deprivation*, Oxford: Oxford University Press.
—— (1999) *Development as Freedom*, Oxford: Oxford University Press.
Singer, P. (1972) "Famine, Affluence, and Morality," *Philosophy and Public Affairs* 1: 229–43.

46
JUSTICE AND BORDERS

David Miller

Before I built a wall I'd ask to know
What I was walling in or walling out …

(Frost 2001: 34)

The relationship between justice and borders is complex. When we make a border, we create a space in which certain things can happen, including perhaps practices of justice. But at the same time we exclude those who stand outside from participating in, or benefiting from, those practices. Is justice then on balance served or denied by the building of walls? Does it matter where the walls are built, who is protected within them and who is excluded? Does it matter whether the walls are easy or hard to climb across? These are the questions I shall try to address in this essay.

The borders I shall be concerned with are jurisdictional borders, of which the chief examples are state boundaries. We might otherwise think of the borders that surround private property, like the walls around Frost's farm. These also raise issues of justice, but perhaps simpler ones, because the wall does not in most cases create an internal practice of justice; the question, rather, is whether the property border prevents outsiders from getting access to something that they can justly claim. I shall organize my enquiry under four headings. First, why might we think that we cannot have justice without borders? Second, why might we think that borders are always repugnant to justice? Third, has justice anything to say about where borders should be placed? Fourth, has it anything to say about whether border crossing should be allowed, and under what conditions?

1. Justice Needs Borders

It would be too strong to say that there cannot be justice in a borderless world. In some of its guises, justice comes into play whenever humans interact with each other. It is unjust to kill, maim or rob another human being without provocation; it is unjust to break a promise, or to fail to help somebody in urgent need when one can do so at little cost to oneself. This holds regardless of any boundary questions. But other forms of justice have a more limited scope. For example, justice sometimes concerns the consistent application of rules: it is unfair if I get punished more severely than Peter for committing the same offence, or unfair if I am denied a promotion when I have performed as well as Jane who gets one. But this depends on the condition that the relevant rules apply to both Peter and me in the first case and to both Jane and me in the second. Such rules normally have a determinate scope: they do not apply to everyone. Another scope-restricted form of justice is the justice that emerges in co-operative practices,

where everyone contributes to the success of the practice, and benefits, in turn, from the contribution of others. (Think of a group of neighbors who decorate their street with plant pots and other items of furniture so that traffic is slowed and it becomes safe for children to play.) It is widely judged to be unfair, given certain conditions, for someone who benefits not to contribute. But this depends on being able to tell who is included in the practice and who isn't.

The boundaries that are at stake here are, in the first place, boundaries between people—between members and non-members of the relevant scheme. They are not yet physical borders of the kind that we are mainly interested in. But one can see why justice communities are very often spatially bounded as well. It is much easier to know who falls under a rule if the rule is applied uniformly throughout a given territory. In cases where this is not so—for example, in parts of medieval Europe, where law was personal in the sense that each person took his own law with him as he travelled from place to place—resolving disputes over justice was difficult, since it had first to be established whose law to apply to the issue in dispute (see Bloch 1962: vol. 1, 111–12). Similar considerations apply when we have to decide who forms part of a given co-operative scheme and who doesn't. If everyone living within determinate spatial borders is included, that question becomes much easier to answer.

It is no surprise, then, that some prominent recent theories of distributive justice should have been explicit about their bounded nature. Most famously of all, John Rawls said that his theory of justice was meant to apply to

> an ongoing society, a self-sufficient association of human beings which, like a nation-state, controls a connected territory … a closed system; there are no significant relations to other societies, and no one enters from without, for all are born into it to lead a complete life.
>
> (Rawls 1999a: 323)

Rawls said this because he believed that the kind of reciprocity between citizens that was needed in order for them to uphold his two principles of justice could not obtain unless each conceived of his or her life being led entirely within such a society. Somewhat different considerations underlie Michael Walzer's claim that "the idea of distributive justice presupposes a bounded world within which distributions take place: a group of people committed to dividing, exchanging and sharing goods, first of all among themselves" (Walzer 1983: 31). Walzer says that this bounded world must be a "political community," which brings together men and women with shared understandings of the meanings of the goods they have to distribute, and therefore (according to Walzer) of how these goods should be distributed, and political institutions that can police the resulting pattern of distribution. Political institutions also feature prominently in Thomas Nagel's critique of the idea of global justice. According to Nagel, "justice is something we owe through our shared institutions only to those with whom we stand in a strong political relation" (Nagel 2005: 121). Nagel argues that it is only when we are subject to a sovereign authority that both coerces us to obey its decisions and claims to act in our name that we owe each other obligations of distributive justice. For the legitimacy of such a system depends on the substantive justice of the laws that govern it.

Rawls, Walzer and Nagel agree that justice needs borders, but they disagree about why, and as a result they are also liable to disagree about which borders should count (see further Miller 2009). For Rawls, the relevant boundaries are those that enclose a

"co-operative venture for mutual advantage"; for Walzer, they are those that surround "a group of people committed to dividing, exchanging and sharing goods, first of all among themselves"; and for Nagel they are the borders determined by the scope of a system of coercively enforced law. Although all three authors assume in practice that the boundaries of justice coincide with the jurisdictional borders of independent states, only Nagel argues for this directly. So we might accept the general idea of bounded justice, and yet challenge the conclusion that principles of distributive justice apply only within nation-state borders as we find them today—for example, if we thought that the international trading system qualified as "a co-operative venture for mutual advantage." (This was the argument made by "Global Rawlsians" such as Beitz 1979: Part III and Pogge 1989: Part III, but rejected by Rawls himself (Rawls 1999b).)

2. Justice Abhors Borders

But should we accept that general idea? Quite apart from worries we might have about the reasons given by these three authors, and others, for restricting the scope of justice, we might think that there is something in the very idea of justice itself that tells against such restriction. For justice is opposed to arbitrariness. It is about treating like cases alike, about the consistent application of relevant criteria regardless of the identity of the person to whom they are being applied. And since the borders we have so far considered are conventional and man-made, it may be hard to see how they can make a difference to what justice requires. Pascal remarked that "it is a funny sort of justice whose limits are marked by a river; true on this side of the Pyrenees, false on the other" (Pascal 1966: 46). Extending Pascal's thought, it might seem strange not only that justice should require French people to treat one another differently from how Spaniards treat other Spaniards, but also that it should require French people to treat their compatriots differently from Spaniards, individual qualities such as need or desert aside. Consider a concrete example. Some countries, and some American states, allow the death penalty to be imposed for murder; many other countries do not. Although at one level justice might be satisfied provided that a particular murderer in Alabama, say, is treated in the same way as other homicides in that state, at a deeper level we might think that either a murderer deserves to be put to death or he does not. If he does, then the laws of all those states that make no provision for the death penalty are unjust; if he does not, then the laws of Alabama and similar jurisdictions that allow it are unjust. Merely treating all those *within* the jurisdiction in the same way is not sufficient for justice proper.

If this example carries weight, it is because we think that the kind of justice at stake here is *non-comparative* justice (for this idea see Feinberg 1980). Justice is non-comparative when the treatment that is due to someone as a matter of justice can be determined without regard to what other people are receiving, or should receive. To think that each individual murderer deserves to be hanged whatever is being done to other murderers is to use a non-comparative notion of justice. Whether such notions are defensible is moot. Our intuitions seem to be stronger about what justice does *not* allow than about what it requires positively. We are pretty confident, for example, that justice does not permit people to be executed for rustling cattle or poaching: that everyone within a particular jurisdiction who is caught taking part in these activities is routinely hanged hardly mitigates the injustice of the law. Where murder is concerned, we think that it should be punished severely—and certainly more severely than crimes such as theft—but what "severely" means can legitimately vary from place to place. Perhaps in a soci-

ety in which violence is endemic and a strong deterrent is required, the death penalty, used consistently, would be just; elsewhere it would not be.

Justice abhors borders, then, to the extent that it takes a non-comparative form. I have raised doubts about how determinate our sense of non-comparative justice really is. But there is another possibility here. Justice might be understood comparatively—the just treatment that is due to A can only be determined by looking at how B, C, D and so forth are being treated—and yet we might think that it is wrong to take particular borders into account when such comparative principles are being applied. Again, an example might help to explain this. When students come to me for advice about their work, I should try to respond to them fairly. Those who need more help should get more of my time, but otherwise I should treat them equally. But this applies only to students who attend my own university. When students from elsewhere approach me, I think I'm entitled to say that my time is limited and I give precedence to those who are students at Oxford. But someone might challenge this (indeed occasionally they have). They might say "I'm coming to you because I'm struggling with this conceptual problem about desert, and you are the person best placed to help me. Why should it matter whether I'm enrolled at Oxford or not?" This person thinks that the relevant border should be drawn in a different place—not around an institution, but around a set of persons identified by an individual property, in this case to be standing in need of a particular form of academic advice.

Those who favor principles of justice that are global in scope often present arguments that take this form. Two people are, say, equally in need—they have a life-threatening condition that requires specific medical treatment. Why should it matter, at the bar of justice, that one is a British citizen and the other is a Somali, since the ground of their claim to receive treatment is the same in each case? The rhetorical power of this question is obvious. But we need to tread carefully. Is the claim of justice that is being made here non-comparative or comparative? In other words, is the claim that the Somali citizen should receive the treatment he needs because that is something he is entitled to as a human being no matter what is happening to anybody else, or is the claim that he is entitled to receive the same treatment as a British citizen because the border that separates the two countries is irrelevant so far as claims to medical treatment are concerned?

If the claim is a comparative one, we need to inspect it more closely. Why should we assume that someone's citizenship is an irrelevance? Return for a moment to the student from outside Oxford who is seeking my advice. It is tempting to think that the only thing that distinguishes this person from an Oxford student is his formal status—he doesn't have the piece of paper in his file that says "Congratulations, you have been accepted to study at the University of Oxford" or words to that effect. But that picture isn't quite right. The Oxford student is engaged in an ongoing practice that involves responsibilities as well as rights. She can approach me for advice in the expectation that I will respond positively, but equally I can go to her with a request (for example, that she should help to organize a seminar I want to put on) in the expectation that she will respond positively unless she has competing commitments. There is unspoken but understood mutuality in our relationship. The same applies on a much larger and less personal scale among fellow citizens. So when the government raises taxes and uses the proceeds in part to finance medical services that only citizens are entitled to use, it is not merely singling out certain persons for help arbitrarily. It is acting as the conduit whereby those who already stand in a certain kind of relationship to one another can

practice mutual aid. We saw in the previous section that there is some disagreement about which feature of the relationship among citizens is relevant when we have to explain why they owe one another special duties of justice. So the anti-borders advocate can try to show that the jurisdictional boundaries that we can see around us in the world today do not correspond to the contours of human relationships that are relevant to justice. But that is not the same as showing that the very idea of bounded justice makes no sense.

Justice does not, then, abhor borders as such. What it abhors are *arbitrary* borders, borders that result in people being treated differently on grounds that are morally irrelevant. Indeed, as the first section suggested, there are important forms of justice that can only exist within well-defined spatial boundaries. So the next question we must address is whether justice has anything to say about where the borders should be placed.

3. Justice and the Placing of Borders

If we look around the world today and contemplate the boundaries of states and lesser jurisdictions, we are likely to conclude that their current positioning owes a great deal to historical accident. Although there might be good practical reasons for leaving most existing borders where they now are, it is natural to ask what a just set of borders might look like. What we are looking at is essentially a triangular relationship between an area of land, an authority that has jurisdiction over that land (makes the laws and the other policies that apply there) and a population who occupy the land and over whom the authority is exercised. Justice would then seem to require two things: first that the land should belong to the people who live there, and second that the authority that exercises jurisdiction should be legitimate, from the point of view of those same people. We can see intuitively the relevance of these two conditions by considering different ways in which territory can be unjustly annexed. The first occurs when a state simply invades a neighboring territory and subjects the people there to its authority without their consent—as Germany did in 1939 when it captured the western half of Poland by military force and incorporated it into the German Reich. The second occurs when a state implants large numbers of its own people on foreign soil, again without the consent of the incumbents, in such a way that a puppet regime can claim to be representative of (at least a considerable proportion of) the current inhabitants. The Soviet Union pursued such a policy in the Baltic states in the period after 1945, and China has done so more recently in relation to Tibet.

Our intuitions about these cases might be clear, but it is much harder to explain positively what it means for the land that lies within jurisdictional borders to belong, rightfully, to the inhabitants, and for authority to be legitimate. Broadly speaking there are two routes we might follow in addressing the first question. One sees land as belonging in the first place to individual people, the other as belonging collectively to "peoples"—a term popularized by Rawls to denote groups of people held together by social ties and common sympathies, without necessarily being "nations" in the modern sense (Rawls 1999b: esp. sect. 2).

If we choose the first route, we are likely to follow John Locke (and later disciples such as Robert Nozick) in holding that individuals can acquire property in land by "mixing their labour" with it, prior to the establishment of government and the jurisdictional authority that it entails (for Locke's account of property acquisition, see Locke 1988: ch. 5). Such acquisition is legitimate provided that it does not deprive others of adequate

opportunities to appropriate land. (The meaning of this so-called "Lockean proviso" has been widely debated—see, among others, Waldron 1979; Simmons 1992: ch. 5, sect. 4; Tully 1980: ch. 5; for Nozick's version, see Nozick 1974: 175–82.) Having secured property rights in land, the owners may then, by consent, come together to establish a government, which they authorize to exercise jurisdiction over the territory formed by amalgamating all of their individual property holdings. Since each rightfully owns his patch of land, the border that is created as a result of the amalgamation is also justly situated (for a contemporary account of this kind, see Steiner 1996 or Steiner 2005).

There are, however, significant problems with this account of just borders (see my longer discussion in Miller 2011). It would imply, for example, that many state borders that we are inclined to think of as justly placed are problematic, since they do not correspond to individual property holdings. Few people would want to suggest that there is anything wrong with the Icelandic government exercising jurisdiction over the whole of Iceland, despite the fact that only a small part of that windswept and largely barren island is held by individual Icelanders as private property. It might be possible to circumvent this problem since in this case at least there are no rival jurisdictional claims to consider. Other difficulties, however, are more intractable. Locke simply takes it for granted that all of the property owners in a given geographical area will agree to incorporate themselves into a single political community. But suppose instead they were to attempt to form several separate communities that overlapped in space. Since authority in the Lockean story depends on consent, there would be no reason to disallow this. But clearly no stable jurisdictional border would emerge. Instead we would have something closer to the feudal arrangement referred to earlier where two people living side by side might each be subject to a different body of law. A further problem with Locke's account is that he assumes that once a property owner has agreed to incorporate his property into a political jurisdiction, that agreement is irreversible. He cannot later change his mind and decide that he would prefer to become a member of a neighboring political community. But Locke does not explain why a theory that places so much weight on the importance of consent requires people to bind not only themselves but their successors to accept the authority of a particular government in perpetuity. (For criticism of Locke on this point, see Steiner 1996: 144 or Steiner 2005: 34.)

In one respect these critical remarks are unfair to Locke, since his concern was not to justify jurisdictional borders but to set down the conditions under which governments could enjoy legitimate authority. In the story that he tells, the boundaries of the political community are simply taken as given. In contrast, more recent individualist theories that see just borders as arising from individual consent—a just border is one that surrounds a group of people who agree to associate with one another to form a jurisdiction—have to allow that such borders are always provisional (see, for example, Beran 1984; Altman and Wellman 2009: ch. 3). Moreover since governments rarely allow their subjects to decide whether they wish to maintain the borders that they currently have, or to shift them by, for example, allowing a subset of the population to secede and establish an independent jurisdiction, we cannot tell on individualist grounds whether existing boundaries are just or not.

Consider, then, the second possibility, that just borders depend upon a people's *collective* right to the territory that they encompass. How might such a right be established? One answer is that rights to territory stem directly from the exercise of legitimate authority over the territory in question, at least so long as the authority in question has not unjustly displaced the previous occupants. This yields a *statist* view of just borders, if

we understand "state" loosely (and improperly) to refer to any geographically based system of political authority. The claim is that where a state is able to exercise stable and effective governance over an area of land, while serving as the legitimate representative of the people who occupy that land, it has the right to set and police the borders that surround that territory. This was the view of Kant, and also of utilitarians such as Sidgwick, according to whom "the strict right of a State to exclude other States from land can only be maintained by its exercising a tolerably effective and continuous governmental control over the territory in question" (Sidgwick 1897: 254). If the latter condition were sufficient, the statist view would confront the objection that an annexing state, such as Germany in 1939, will gain territorial rights in other places, *provided* it succeeds in imposing "tolerably effective and continuous governmental control" there. So statists need to add the condition that the state must legitimately represent the people who occupy the territory, which Germany did not in the case of the great majority of the inhabitants of western Poland. But this still leaves the problem of implantation, referred to earlier, or indeed the problem that arises when ethnic cleansing takes place. If a group living in a territory succeeds in expelling the members of a second group, and then proceeds to establish a form of government that legitimately represents the cleansers, while exercising sufficient political control to satisfy Sidgwick's condition, does that group now have a just claim to the territory in question, or does the expelled group have a just claim to be reinstated? It might be possible to refine the statist view so as to avoid having to answer the last question in the negative (see Stilz 2009, 2011). But the maneuvering involved brings out an underlying tension in the statist view of borders, namely between the emphasis it lays on the present performance of the state in justifying borders, and the need to recognize that the people who live in a place over time develop an historical entitlement to be there.

This suggests a second collectivist answer to the just borders question, one that points not to states but to *peoples* as the primary holders of territorial rights ("peoples" here might be nations in the modern sense, but they could also be indigenous groups such as the aboriginal peoples of North America or Australia). This view attaches great importance to the intimate relationship that usually develops over time between a people and the land that it occupies, so that it becomes seriously unjust if they are expelled or denied a right to control the territory in question. One strand of argument takes a quasi-Lockean form, referring to the material value that is added to land when successive generations improve it, for example by cultivating agricultural land, or erecting buildings. (The argument is *quasi*-Lockean because it focuses on the co-operative activity of a people in transforming land, not on improvements made by particular individuals; this also allows it to include features such as infrastructure (roads, waterways, etc.) that is not privately owned.) Another strand emphasizes the symbolic value that the land might have for the people who occupy it over time, as some places acquire religious significance, for example, and others may come to represent major episodes in the nation's history. Taken separately or together, these arguments are said to show that the people in question has a just claim to establish borders around the territory in question, and to exercise continuing control, through having their own state, of what goes on inside them (see Meisels 2009; Nine 2008; Miller 2012).

But this view also faces objections. One is that it might fail to provide guidance in cases where groups make rival claims to territory, each group being able to invoke justifying arguments of the kind sketched above. It cannot easily resolve, for example, disputes such as that between Israelis and Palestinians, where geographical intermingling

over time means that each group can claim to have added material value to the disputed land, and where each can also justifiably point to the great symbolic importance of particular sites, given its own culture and beliefs. If neither partition nor joint control of the territory provides a feasible solution, the best that this collectivist approach can offer is a judgment about which group has the stronger claim to occupy and control it. This will leave the rival group with a residue of injustice for which it might be impossible to provide adequate compensation, since the symbolic value that is lost (for example, by being denied access to sacred sites that are central to the group's culture) might be incommensurable with material resources that might be offered in return.

A second problem is that even in cases where there are no conflicting historical claims to the specific territory that a people occupies, arguments of justice might be advanced on the part of outsiders who do not have access to territory of comparable value or, in the extreme case, to any territory at all. (Consider what happens when land becomes uninhabitable, for instance when rising sea levels due to climate change begin to inundate low-lying islands.) Here we need to return to the distinction between comparative and non-comparative justice introduced above. A comparative claim of justice, made in relation to territorial borders, would hold that each person must have access to territory of similar value, or else be compensated in some other way if the territory she inhabits is less conducive to human well-being than that of others. As argued earlier, however, it has first to be shown why such comparative claims should be thought to hold *across* borders, rather than simply between those who already stand within them. Non-comparative claims do not beg the question in the same way. One such claim is that every human being is entitled to have the opportunity to lead a minimally decent life, which might not be possible with borders being drawn as they now are. This will most likely translate into an individual right to cross borders, to be discussed in the following section. However, one can envisage cases in which some redrawing of borders might be called for—for example, where a brutal civil war creates a mass exodus of refugees, justice might demand that a neighboring state permits the creation of a safe haven within its borders, over which it only exercises minimal control, primarily to protect the refugees against incursions by their persecutors.

In a case where land becomes uninhabitable, and where its previous occupants have constituted themselves as a distinct political community, there may be a *collective* claim of justice for access to new territory on which the community can be reconstituted (see Nine 2010). A case in point is the Maldive islands, whose government is currently negotiating with several countries to purchase land in case the islands should soon be submerged. The difficulty here is to identify which states hold the corresponding duty. Yet if claims to land are going to be justified according to the group entitlement view sketched in this section, then a group that, by no fault of its own, finds itself deprived of usable land is entitled as a matter of justice to some other group's "surplus"—land that can be surrendered without significant loss of either material or symbolic value.

4. Justice and Border Crossings

All states currently claim the right to regulate the movement of people and goods across their borders, unless, as in the case of the European Union signatories to the Schengen Agreement, they have agreed to waive that right, in whole or in part. The most contentious aspect of the right is the right to exclude would-be immigrants on grounds chosen by the receiving state. What does justice have to say about this? The distinction

between comparative and non-comparative justice is again helpful here. Non-comparative justice is widely agreed to include the universal duty to protect human rights—universal in the sense that it both falls upon and applies to everyone. So if the closing of borders violates this duty, it must be treated as an injustice. Much then depends on how widely or narrowly we understand human rights. Kant, who paradoxically is often regarded as a prophet of present-day cosmopolitanism, said that a stranger arriving in a foreign land "can indeed be turned away, if this can be done without causing his death" (Kant 1971: 105–6). This construes the stranger's rights very narrowly, as comprising only the right to life. Moreover, Kant seems to be thinking of cases in which refusing entry would cause death immediately, by, for example, forcing someone back on to a shipwrecked boat. It is not clear what he would say about instances in which the stranger's only option is to return to a place where his life will be put in danger by others. By contrast, contemporary international law reflects the 1951 UN Convention, which grants refugee status—and therefore a right of admission—to anyone impelled to leave her country out of a "well-founded fear of being persecuted" on grounds such as her race or religion.

This still, however, limits the grounds on which a person can demand the right to cross a border to threats to her basic interests that are actively imposed by others. What about rights to subsistence, health care or elementary education, which inhabitants of poor societies might be unable to claim unless they travel abroad? Allowing these to count would extend the right to cross borders very considerably. But it also highlights an issue that arises even with respect to existing refugee law, namely that it is indeterminate against whom the right can be invoked. Why must one particular state assume the burden of admitting refugees, since the human rights on which their claim for admission is based are held against everyone world-wide? So to show that a state that closes its borders is guilty of injustice, it would be necessary to show that it breached a particular duty to those it refused to admit. Perhaps the only clear-cut cases are the ones that Kant appears to have had in mind, where the immigrant's physical situation is such that he has become dependent on a particular political community to carry out a duty of rescue. Beyond that, it can be argued that states have a secondary duty of justice to devise a *system* that would allow refugees to be distributed fairly between receiving states.

If we turn our attention to comparative justice, it has been argued that open borders are required by the principle of global equality of opportunity (see Carens 1987; Caney 2005: ch. 4). The argument is straightforward: since each society provides its members with a different set of opportunities—and in general the opportunity sets available in rich societies are much fuller than those available in poor societies—the only feasible way to ensure that each person, no matter where they are born, has equivalent opportunities is to allow for free migration between them. The issue, then, is whether global equality of opportunity is a valid principle in the first place (see my longer discussion in Miller 2007: ch. 3). In Section 2, I explained why comparative principles of justice are often limited in scope, and in the present case there are good reasons to think that equality of opportunity is a principle that holds among the citizens of a democratic state, but not among people world-wide. (This was Rawls' view, sketched briefly in Rawls 1999b: sect. 16 and expanded upon in Freeman 2006.) If that is so, it cannot be used to show that the closing of borders is always unjust.

Comparative principles of justice might, however, apply among those who seek to immigrate to a particular state. Immigration, or at least *legal* immigration, is after all a rule-governed process. Perhaps, then, a state that controls immigration by setting

quotas, issuing visas, and so forth creates a relevant constituency among those who are trying to enter by legal means. What kind of justice applies to this case? One view is that the state is only required to observe norms of procedural justice: having decided what entry criteria it is going to use, it must apply these criteria consistently, take steps to ensure that the information it receives about potential immigrants is accurate, etc. But what would we think of a state whose (consistently applied) rules discriminated among applicants on grounds of race or sex? It seems that we would regard this as unjust, in which case justice across borders also places some limits on the criteria that may be used to determine entry. Using work skills, or language competence, as criteria seems acceptable, for example, when these can be shown to be relevant to the needs of the receiving society, and are not being used in an underhand way to select candidates who fulfill other conditions such as having the "right" skin color.

5. Conclusion

The relationship between justice and borders, we have seen, is complex. It cannot be said that a world with fewer borders, or softer borders, will necessarily be more just than its opposite. It will depend on where the borders are placed, what forms of justice are created inside them, and how border crossings are regulated. As they mend the wall that runs between their properties, Frost's counterpart repeats the old proverb "Good fences make good neighbours." But why do they? Frost wants to know, in a spirit of mischief—yet he piles the fallen stones back as he asks.

Related Topics

Nationalism, Human Rights and Cosmopolitanism, Global Justice and Politics, Equality

References

Altman, A. and Wellman, C. (2009) *A Liberal Theory of International Justice*, Oxford: Oxford University Press.

Beitz, C. (1979) *Political Theory and International Relations*, Princeton, NJ: Princeton University Press.

Beran, H. (1984) "A Liberal Theory of Secession," *Political Studies* 32: 21–31.

Bloch, M. (1962) *Feudal Society*, trans. L. Manyon, London: Routledge and Kegan Paul.

Caney, S. (2005) *Justice Beyond Borders*, Oxford: Oxford University Press.

Carens, J. (1987) "Aliens and Citizens: The Case for Open Borders," *Review of Politics* 49: 251–73.

Feinberg, J. (1980) "Noncomparative Justice," in J. Feinberg, *Rights, Justice, and the Bounds of Liberty*, Princeton, NJ: Princeton University Press.

Freeman, S. (2006) "Distributive Justice and *The Law of Peoples*," in R. Martin and D. Reidy (eds.), *Rawls's Law of Peoples: A Realistic Utopia?*, Oxford: Blackwell.

Frost, R. (2001) "Mending Wall," in E. Lathem (ed.), *The Poetry of Robert Frost*, London: Vintage.

Kant, I. (1971) *Perpetual Peace: A Philosophical Sketch*, in H. Reiss (ed.), *Kant's Political Writings*, Cambridge: Cambridge University Press.

Locke, J. (1988) *Two Treatises of Government*, ed. P. Laslett, Cambridge: Cambridge University Press.

Meisels, T. (2009) *Territorial Rights*, 2nd edn, Dordrecht: Springer.

Miller, D. (2007) *National Responsibility and Global Justice*, Oxford: Oxford University Press.

Miller, D. (2009) "Justice and Boundaries," *Politics, Philosophy and Economics* 8: 291–309.

Miller, D. (2011) "Property and Territory: Locke, Kant, and Steiner," *Journal of Political Philosophy* 19: 90–109.

Miller, D. (2012) "Territorial Rights: Concept and Justification," *Political Studies* 60: 252–68.

Nagel, T. (2005) "The Problem of Global Justice," *Philosophy and Public Affairs* 33: 113–47.
Nine, C. (2008) "A Lockean Theory of Territory," *Political Studies* 56: 148–65.
Nine, C. (2010) "Ecological Refugees, States Borders, and the Lockean Proviso," *Journal of Applied Philosophy* 27: 359–75.
Nozick, R. (1974) *Anarchy, State and Utopia*, Oxford: Blackwell.
Pascal, B. (1966) *Pensées*, ed. A. J. Krailsheimer, Harmondsworth: Penguin.
Pogge, T. (1989) *Realizing Rawls*, Ithaca, NY: Cornell University Press.
Rawls, J. (1999a) "Kantian Constructivism in Moral Theory," in J. Rawls, *Collected Papers*, ed. S. Freeman, Cambridge, MA: Harvard University Press.
Rawls, J. (1999b) *The Law of Peoples*, Cambridge, MA: Harvard University Press.
Sidgwick, H. (1897) *The Elements of Politics*, 2nd edn, London: Macmillan.
Simmons, A. J. (1992) *The Lockean Theory of Rights*, Princeton, NJ: Princeton University Press.
Steiner, H. (2005) "Territorial Justice and Global Redistribution," in G. Brock and H. Brighouse (eds.), *The Political Philosophy of Cosmopolitanism*, Cambridge: Cambridge University Press.
Steiner, H. (1996) "Territorial Justice," in S. Caney, D. George and P. Jones (eds.), *National Rights, International Obligations*, Boulder, CO and Oxford: Westview Press.
Stilz, A. (2009) "Why do States have Territorial Rights?" *International Theory* 1: 185–213.
Stilz, A. (2011) "Nations, States and Territory," *Ethics* 121: 572–601.
Tully, J. (1980) *A Discourse on Property: John Locke and his Adversaries*, Cambridge: Cambridge University Press.
Waldron, J. (1979) "Enough and as Good Left for Others," *Philosophical Quarterly* 29: 319–28.
Walzer, M. (1983) *Spheres of Justice: A Defence of Pluralism and Equality*, Oxford: Martin Robertson.

47
WAR

Fernando R. Tesón

1. The Place of War in International Society

War is the most terrifying and destructive form of violence. Any philosophical analysis of war, therefore, must start with a strong presumption against its permissibility. Yet many people share the intuition that some wars are justified. In spite of the recurrent suspicion that war somehow escapes moral evaluation, that when nations fight all bets are off, philosophers have for centuries attempted to explain what can possibly justify this extreme form of human violence. The morality of war is subdivided into two areas. *Ius ad bellum* consists of the principles under which states are authorized to wage war. *Ius in bello* consists of the principles that govern lawful combat. This chapter treats them separately, although the two sets of principles are related in complex ways. But we must first see why war has such a special place in international relations.

Each sovereign state establishes the monopoly of force within that state. States have police, courts, and armies, and individual violence is narrowly confined. The modern state outlaws interpersonal violence by monopolizing force. It prohibits both opportunistic plundering and private retribution, while providing means of redress to those wronged by the aggressive behavior of others. The international society, however, is not a state. Its central feature is precisely that it does not *in fact* establish a global monopoly of force on any higher sovereign. At its inception, the United Nations Charter was meant to create a system of collective security that would replace unilateral wars, but its weaknesses are well known. Suffice it to say that international law does not provide a satisfactory remedy for nations, groups, or persons wronged by other nations. To be sure, the international society has made great strides in developing *norms* to curb war. International law has codified in detail the *ius in bello*. The *ius ad bellum*, however, is a different story. The United Nations Charter reads *as if* it prohibits most interstate violence and *as if* it empowers the Security Council to respond to aggression on behalf of the international community (U. N. Charter, art. 2.4 and ch. 7). Because of this *discursive* consensus, if today a state invades another for no good reason observers are more ready to condemn the invasion than they would have been only 70 years ago. Yet the mechanisms available to *actually* restrain war are woefully weak. Before the United Nations Security Council authorizes the use of force against an aggressor *and* a government is willing and able to act on that authorization, many improbable things have to happen. For one thing, the five permanent members of the Security Council have the right to veto. This means that the Security Council will not authorize force against permanent members or their allies. But more importantly, while domestic law *obligates* the government to act in response to private violence, international law (in the best

scenario, when the United Nations Security Council functions properly) merely *authorizes* governments to respond to aggression (and genocide or other major humanitarian crises, as we shall see.) Assuming, against the facts, that the Security Council acts expeditiously in every case of aggression or genocide and authorizes the use of force to address those, still governments are not *bound* to react. To begin with, only a powerful state (usually but not always the United States) can act. Unless weaker states can forge a sufficiently powerful alliance (another implausible scenario), they will free-ride on the powerful state. But the powerful state will not act unless its government believes the national interest to be at stake (this, again, is a rosy scenario: politicians will go to war if *their* own interests are thereby advanced, especially their electoral interests). States in a position to respond to aggression can simply decline the invitation.

States do not have an obligation to protect other states against aggressors, nor can the United Nations Security Council force them to do so. Faced with threats, governments continue to rely on self-help, and war, of course, is the extreme form of self-help. Notwithstanding the protestations of international lawyers, there is no effective global mechanism to prevent war. Self-help is omnipresent in international relations. Given the sobering fact that states and others will continue to commit aggression and other serious wrongs, war, unfortunately, is here to stay.

2. Ius ad Bellum

Writers have adopted a range of positions on the question of the right of states to wage war. They range from Realism, the view that states are in a state of nature and have, for that reason, ample discretion to wage war, to pacifism, the view that no war is ever justified. Most contemporary theorists fall somewhere in the middle, however: war is generally wrong, but there are some justified wars. This position, called just war theory, is also the position in modern international law. This chapter will address the more widely accepted theories of *ius ad bellum*: Realism, just war theory and pacifism.

2.1 Realism

According to this doctrine, nations find themselves in a state of nature where morality has no place. Because nations are compelled to fight for survival, no war can be morally condemned from a universal standpoint, since statesmen owe moral considerations only to their own citizens (Hobbes 1962: ch. 18; Kennan 1986). Realists need not be moral skeptics: they just think that the lack of meaningful global institutions deprives international relations from the commonality of values that alone justifies a common moral language. Morality makes sense only within states. This doctrine can be called *normative Realism*. It must be carefully distinguished from *descriptive Realism*, the view that *in fact* states always act in the national interest. The upshot of normative Realism is that governments have an unlimited discretion to wage war. Of course, Realists can criticize governments for starting reckless, hopeless, or counterproductive wars, or for violating the *ius in bello*, but not for lacking a just cause.

This doctrine has been often criticized. The criticism is *not* that, contrary to what Realists claim, international institutions are adequate (this would be self-delusion). The criticism is that moral judgments are possible even in a situation of semi-anarchy like the international arena (Cohen 1985). Unless one is a moral skeptic, it seems odd to withhold moral judgment precisely about the one event that has caused such a hor-

rific destruction in history. If morality is universal, then there is every reason to subject the act of war to moral appraisal.

Realism, however, contains a grain of truth. The lack of global institutions has *in fact* freed political leaders to wage war with impunity. The absence of global government creates incentives for rulers to seek their objectives through war. As Kant observed, these leaders often externalize the costs of war while internalizing military success (Kant 1991). But this is an insight of descriptive Realism. If the structure of incentives makes war easier, that is more of a reason to subject war to ethical scrutiny.

2.2 Just War Theory

Reduced to its minimal expression, just war theory simply holds that some wars are justified and some are not. Just war doctrine thus rejects both Realism and pacifism. More specifically, just war theorists claim that wars are justified when they meet a number of conditions: they must be waged for a just cause; they must be waged with the right intent; they must be waged by the government (no private wars); and they must be proportionate (Aquinas 2002: 240–2). In the classical formulation of the doctrine, a state has a just cause whenever it vindicates a *right* violated by another state. Thus, any violation of international law gives just cause for war (Vitoria 1991: 303–4). Modern writers, however, have rejected such a broad view of just cause. From a moral standpoint, waging war for any violation of a state right is simply disproportionate any time the moral costs of war can be averted by diplomacy. For that reason, modern versions of just war theory (in line with international law developments) accept only two just causes for war: national self-defense and (more controversially) humanitarian intervention.

National self-defense (called "national" to distinguish it from individual self-defense in the criminal law) is the military action by a state against another state's unjustified military attack. The argument is straightforward: national self-defense is justified as the response to aggression, which is an unjustified attack. National self-defense relies on the *domestic analogy* (Walzer 2006: 58–63). Sovereign states are analogized to persons. Just as individuals live in civil society, so states live in the international society. This allows us to transpose the categories of the criminal law (trespass, aggression, self-defense) into international relations. However, because international institutions are weak and imperfect, the analogy cannot be carried too far. For example, whereas in domestic law self-defense is limited to repelling someone's attack, in international law the victim of aggression can pursue and punish the aggressor. But the analogy is good enough: national self-defense is the defense of *states*. Citizens owe allegiance to their state and are bound to one another to defend it.

An alternative view rejects the domestic analogy and holds that self-defense is justified as a defense of *persons*, not of states. War is justified when it is waged in defense of persons, their rights, and, if applicable, their free institutions (Tesón 2005: 150–4). This view does not require that citizens of the state should owe allegiance to their government. War in self-defense is simply an extension of the permissibility of individual self-defense (since the aggressor threatens the agent) and defense of others (since the aggressor threatens them too). The government performs, on this view, a coordination role: national self-defense would not be successful unless a central authority is prepared to take charge and organize the defense efforts. The government simply coordinates the efforts of persons who are willing to defend themselves against the aggressor. This view can explain why even illegitimate governments have the power to organize the defense

against aggression. The holders of the right of national self-defense are the persons who inhabit the state, not the government. It is merely an instrument.

The justification of war as defense of persons must address the question of whether resisting aggression is, for each citizen, obligatory or merely permitted. One view is that resisting is permitted but not mandatory. Persons have a liberty to resist or to submit, and if they choose to submit they should be left alone. Conscription, the forcible enlisting of citizens, is a grave violation of individual rights. The alternative view is that resisting aggression is obligatory. The duty may rest on social contract considerations or, more plausibly, on considerations of collective action. National self-defense is a public good: because fighting is risky, everyone will be tempted to free-ride on the courage of others. National defense will then be under-produced and the state will succumb to the aggressor. Many people believe that the government may permissibly force people to fight to avoid this free-rider problem. This argument, however, fails to address the problem of "honest holdouts." Some persons who refuse to resist are not free riders: they genuinely prefer to surrender than to fight. A solution to the problem of conscription is to organize a voluntary army or hire mercenaries.

Strict self-defense must be distinguished from *extended* self-defense. Both are military actions in response to an unjustified attack. But strict self-defense is confined to those military actions necessary to repel the aggressor. Extended self-defense is military action originally launched to respond to aggression but pursued beyond the action that was strictly necessary to repel the aggressor. Virtually everyone agrees that strict self-defense is morally justified. David Rodin, however, has argued that national self-defense as currently understood in international law is unjustified precisely because it authorizes extended self-defense (Rodin 2003). He claims that soldiers fighting defensive wars typically do things that are not morally justified under a plausible philosophical account of self-defense, such as the strict self-defense rationale found in the criminal law. For example, under the current law of war, soldiers fighting defensive wars can kill enemy soldiers who are sleeping or marching. Also, under the current *ius ad bello*, states can adopt proportionate military actions well beyond those needed to repel the aggressor. Yet the rationale for permitting an agent to use deadly force in self-defense is that the aggressor is threatening him. Because international law allows a belligerent using defensive force to kill enemy soldiers who are not threatening it, national self-defense as conventionally understood is morally untenable. Killing non-threatening persons is unjustified; therefore, national self-defense (other than strict physical response to aggression) is unjustified. The same objection applies to other extensions of self-defense, such as punitive wars or wars pursued to disarm the aggressor or replace the regime.

There are two answers to this objection. First, national self-defense can be reduced to a collection of individual rights to self-defense. National self-defense lies in a continuum between self-defense and defense of others (McMahan 2004: 76–7). If we assume that an individual can legitimately use force in self-defense, then we can expand that case to an increased number of persons that coordinate their defensive efforts against a common threat. These persons are justified in killing sleeping or marching enemy soldiers because those soldiers have already attacked and are preparing to attack again. They *are* threatening their victims. The threat is part of a complex action developing at different periods of time starting with the initial acts of aggression. This conception explains why it is morally *prohibited* to kill enemy soldiers who are retreating. Unlike those who are sleeping in order to restore their strength to continue their aggression, retreating soldiers presumably pose no future threat. Second, if international law is morally defective, then

it should be reformed to prohibit killing persons who do not threaten us. This does not mean that a properly revised national self-defense is morally impermissible, only that current international law wrongly allows for instances of morally impermissible killings.

Some writers claim that national self-defense is the *only* just cause for war. This position makes sense, perhaps, if the rationale for war is the defense of *states*, but not if the rationale for war is the defense of *persons*. The latter rationale can justify self-defense *and* humanitarian intervention. National self-defense is a war to defend compatriots; humanitarian intervention is a war to defend foreigners, usually against their own government. While there are important differences between self-defense and humanitarian intervention, it is misleading to contrast *defensive* with *offensive* wars, and consider national self-defense as an instance of the former and humanitarian intervention as an instance of the latter. Humanitarian intervention is *also* a defensive war in an important sense: it is, like self-defense, a war in defense of persons who have been wrongfully attacked. If war has only one generic justification, the defense of persons, then both wars in self-defense and wars in defense of others (humanitarian intervention) can be just.

Of course, humanitarian interventions will often be wrong because of their bad consequences: they might be counterproductive, or they could encourage abuse by the intervener, or they might cause disproportionate harm. (An interesting question: why aren't these considerations as relevant in cases of national self-defense? If a state is attacked, couldn't the consequences of resisting be also disproportionate?) The interesting case is, then, a military intervention that rescues many persons from rulers who seriously violate their rights. One issue is the threshold of permissibility. Some writers require the threat to be particularly serious, such as genocide or its equivalent; others require serious violations but not necessarily reaching genocidal levels. Other issues are especially pressing: does a humanitarian intervention require authorization by an international body such as the United Nations Security Council? What are the duties of a successful intervener (matters of post-intervention)? At any rate, it seems plausible to require that a humanitarian intervention, to be justified, must be invited or at least welcomed by the victims. This requirement is met when the *victims themselves* welcome the intervention. Some writers have argued that, to be justified, a *majority* of citizens of the state must welcome the intervention. But this does not seem right. If the government, aided and abetted by the ethnic majority, is perpetrating genocide against an ethnic minority, it is absurd to require the consent of the majority to foreign intervention to save the victims.

A further requirement of just war theory is right intention. The question is not as important in national self-defense, since the state has been attacked and there is presumably no question that the military response intends to repel the attack. But in humanitarian intervention intention is important. Just war theorists require that the intervener *act out* of the desire to rescue, and not for self-regarding reasons. It is not enough that the war liberates the victims. The warriors must *intend* to rescue the victims, and nothing else. This condition, however, is too demanding, for two reasons. First, governments do not have a separate moral standing: they owe a fiduciary duty to their citizens. They cannot freely engage the resources of the state to save others without regard to the interests of their own citizens. Therefore, they should not have (and actually will *never* have) *solely* the interests of the victims at heart. Second, this requirement confuses intention with motive. Rescuing victims of genocide is the right thing to do, even if performed as *means* to achieving a further, non-altruistic, goal. Say a state wishes to achieve geopolitical dominance in the continent. In order to do so, it invades a neighbor whose government happens to be carrying out genocide against

its own people. The intervener overthrows the genocidal regime and saves the people from genocide. This was, plausibly, a justified humanitarian intervention, even if the intervener had a further (and perhaps dishonorable) motive.

2.3 Pacifism

There are many varieties of pacifism. *Absolute* pacifism holds that war is always prohibited because killing (or violence) is always prohibited. A variation of absolute pacifism makes a "clean hands" argument: every ethical agent must lead a virtuous life. This means that the agent must avoid the state of affairs where *he* kills someone. This is true even if his refraining to kill someone means that many other people will be killed (usually by that someone). Most people find absolute pacifism implausible, for the same reasons that they find the banning of self-defense and defense of others implausible. If personal self-defense is acceptable, then at least some wars are acceptable as well. Absolute pacifism, then, stands or falls with the right of individual self-defense. For the absolute pacifist there is nothing specific about war, as it is just another form of impermissible violence. The "clean hands" argument is not very persuasive either, especially as applied to war. While failing to defend *oneself* (turning the other cheek) might, perhaps, evince noble character, failing to defend *others* against unjustified attacks is more suspicious, as it imposes the costs of our (supposed) virtuous behavior on those others.

Contingent pacifism concedes that in theory some wars might be justified, but claims that the technological reality of war is such that no *actual* war can be justified. Because the just warrior cannot confine his destructive action to morally permitted targets, all war is impermissible. If he could confine his military actions to purely military targets, then the just war would be surgical, as it were: no innocent people would die. Alas, this is not possible. Noncombatants have not waived their right to life, so those killings are wrong (Holmes 1989). One who upholds an absolute prohibition of killing the innocent must, it seems, reject the permissibility of all wars. This pacifist argument, then, amounts to the charge that any war, even an initially justified one, necessarily violates the *ius in bello*, more specifically the rule that establishes the immunity of noncombatants. The objection is powerful because it detects an inconsistency in just war theory.

A possible reply may resort to reflective equilibrium. If collateral deaths are absolutely prohibited, then we should indeed be pacifists. But pacifism clashes with our intuitions that some wars and revolutions were rightly fought. We *know* that some wars (and some violent revolutions) are morally justified. Therefore, the absolute prohibition on killing innocents and its corollary, pacifism, must be rejected. This, however, begs the question: maybe the pacifist is right and those wars that we thought justified were objectionable. Perhaps we should revise our intuitions and become pacifists. (A more plausible reply to this argument appeals to the Doctrine of Double Effect, addressed below.)

3. *Ius in Bello*

Ius in bello is the system of rules that govern the conduct of hostilities. These rules have a long history. They were customary in origin and were later codified by the Geneva Conventions. It is an open question whether or not the laws of war (as these rules are called) track the *morality* of war. Some of the rules are clearly morally justified: the prohibition of killing those who are *hors de combat* and the obligation to discriminate between combatants and noncombatants. But the possible dissonance is between morality and

what the laws of war *permit*. Notably, the laws of war authorize the *collateral* killing of innocent persons such as civilians. This chapter will address three controversial issues in *ius in bello*: the requirement of proportionality; the permissibility of collateral killing of innocent persons; and the right to kill enemy soldiers.

3.1 Proportionality

What counts as proportionate military action is obscure. The traditional view is that an act of war is proportionate when it meets a roughly utilitarian test. Henry Sigdwick's formulation is typical: the aim of the moral combatant must be to disable his enemy and force him into submission, "but not do him (1) any mischief which does not tend materially to that end, nor (2) any mischief of which the conduciveness to the end is slight in comparison to the amount of mischief" (Sigdwick 1908: 267). The proportionality rule applies to all combatants, and to both sides in a war. Traditionally, it has been held to be independent of the justice of the cause. For example, the Allied bombing in Dresden was wrong because it was disproportionate, even if the Allies were engaged in a just fight. Usually, the principle of aiming at legitimate military targets and not at civilians is explained in terms of proportionality. There is, however, a different sense of proportionality: the (act of) war must be proportionate to the *moral urgency* of the cause. Stopping genocide is more urgent (morally) than, say, restoring democracy. The threshold of permissibility of destruction will rise for the just warrior in proportion to the importance of the wrong that the warrior attempts to redress. This is a normatively qualified calculation of proportionality and does not collapse into a calculation of costs and benefits. The Libyan revolutionaries might think that starting a revolutionary war to overthrow Qaddafi is justified even if expected costs are high, because of the importance of the objective—ending Qaddafi's tyranny. A belief in the justice of this war is consistent with an expectation that the war will be costlier in terms of blood and treasure than a decision not to fight.

3.2 The Permission to Kill Enemy Combatants

The traditional view, endorsed by the laws of war, is that soldiers on any side of a conflict have a moral permission to kill enemies. This view has been dubbed "the moral equality of soldiers" (Walzer 2006: 34–7). The idea is to sever the *ius in bello* from the *ius ad bellum*. In many wars, we simply do not know who is right or wrong and, more importantly, the soldiers themselves do not know this. They sincerely believe they are fighting for their homeland. Because of this moral equality, where both sides believe to be in the right, we must at least insist that the belligerents abide by the *ius in bello*, the war convention. The only way to do this is to grant all soldiers the right to kill enemy soldiers. As long as they abide by the laws of war, these soldiers will be protected from prosecution and retaliation. A soldier who complies with the *ius ad bello* is an honorable soldier, even if his government did not have a just cause to fight.

Jeff McMahan has vigorously challenged the traditional view (McMahan 2009). His revisionist argument is straightforward. Every killing by someone who does *not* have a just cause is an unjustified killing, even if the killer complies with the laws of war. On this revisionist view, combatants are not morally equal. The henchmen that defended Saddam Hussein or Qaddafi loyalists should have simply deposed their arms and gone home. It may well be that the war convention is necessary, but that does not absolve

philosophers from observing that many acts permitted by the laws of war are, morally, murder. On this view, then, the only justified killing in war is the killing by a just warrior. Of course, all killings must *also* respect the laws of war and be proportionate. Someone who initiates a war will be morally justified only if he *both* has a just cause (complies with *ius ad bellum*) and fights lawfully (respects *ius in bello*). A soldier who kills for an unjust cause in observance of *ius in bello* may well be excused, but his killing will be *objectively* wrong. (The revisionist view does not entail a prohibition of collateral killing of civilians and is compatible with Doctrine of Double Effect discussed below.)

Traditionalists have a couple of responses to the revisionist position. First, more often than not the justice of the cause is uncertain, especially for the foot soldier. This uncertainty is compounded by the absence of a neutral authority to decide the justice of the cause. Given this epistemic barrier, it is best to bracket the issue of just cause and make sure that armies fight morally. Second, if soldiers who fight for an unjust cause will be held morally or legally accountable, then they will have no incentives to comply with *ius in bello*. More generally, some believe that the revisionist position misdescribes just war theory because, according to the latter, the doctrine of the moral equality of combatants is simply temporary and pragmatic, not a central feature of the theory: just war theorists (other than Walzer) agree that a morally justified killing in war must comply not only with *ius in bello* but with *ius ad bellum* as well (Kaufman 2010).

Another important subject of debate is what exactly makes persons morally liable to be killed in war. Let us take a war in national self-defense. One possibility is that the aggressor is liable because he is guilty and deserves punishment. This view is not widely accepted, as most combatants do not meet this condition. An alternative is to say that the aggressor is liable because he poses a threat. This criterion is more plausible, but still seems over-inclusive: a scientist who does weapons research in his laboratory poses a great threat, but, under the common understanding of the laws of war, he is (arguably) not liable to defensive force. For this reason, some authors have attempted intermediate standards, such as the "moral responsibility for objectively unjustified threat" (McMahan 2009: 208). The many complexities of this issue cannot be addressed here, but these criteria should also account for the liability to be killed in wars in defense of *others*, such as humanitarian intervention. Suppose the government in another country is killing thousands, and the neighboring country invades to stop the killing. If the right criterion for liability is that the target should pose a threat, then the government that is "attacked" (the murderous regime) has a right to fight defensively. If the right criterion is moral responsibility for an unjustified threat, then the "attacked" regime does not have defensive rights because the attack against it by the invading army is justified as an attempt to foil the regime's unjustified attack against its subjects. Unjust attackers do not have defensive rights against third parties who assist those attackers' victims.

4. The Doctrine of Double Effect

The Doctrine of Double Effect (DDE) distinguishes between *intended* killings and *merely foreseen* killings (Boyle 2001). It is morally wrong to deliberately target innocent persons. This prohibition grounds one of the central principles of *ius in bello*: the immunity of noncombatants. The commander who targets innocent civilians to demoralize the enemy is guilty of murder because he *wills* their deaths. (We bypass the issue of what counts as innocence in this context.) But the commander who aims the guns at the enemy while merely *foreseeing* that civilians will die is on a different moral footing. The

deaths of the civilians are not essential to his destruction of the enemy; he would spare them if he could. Some authors say that in the first case the commander treats the civilians as *means* to his end of winning, whereas in the second case the commander does not treat the civilians as means. This intent-based distinction between the two cases is essential to the DDE.

The formulation of the doctrine varies considerably in the literature. We can provisionally state it thus:

An act of war that brings about the deaths of innocent civilians is permissible if, and only if:

1. the means used are permitted by the laws of wars (thus, no poisonous or biological weapons);
2. the agent *targets* only the enemy, even if he *foresees* that civilians will die;
3. the good effect (the defeat of the enemy) does not come through the deaths of civilians; and
4. the good result (the defeat of the enemy) is significant enough to permit the death of civilians to come to pass.

Condition (1) stipulates that the means used by the warrior have to be themselves acceptable. If the commander uses poisonous or other prohibited weapons the act will be unjustified even if he complies with the other conditions. Most of the literature has focused on condition (2): the warrior should not intend the deaths of civilians, even though he may foresee those deaths. According to DDE scholars, the collateral deaths of innocents are not *intrinsic* to the just warrior's action, and that means that in some important sense the agent does not intend those deaths. But the language of intention is misleading, because the evil effect is certainly *imputable* to the agent. He willingly and knowingly brings about the deaths of those civilians. In what sense, then, can we plausibly say that he did not *intend* their deaths? The commander was informed that if he fired his guns the innocent civilians would die. He went ahead and fired. On a plausible meaning of "intent" he intended those deaths: he knew that if he fired those persons would die. In view of this difficulty, proponents of the doctrine have conceded that in these cases the agent intends the evil, and have distinguished between *direct* intention and *oblique* intention. The commander intends all the predictable causal consequences of his behavior. He *directly* intends the killing of the enemy soldier, while he *obliquely* intends the death of the noncombatants. An agent aims at achieving X (a good consequence) but knows that by achieving X he also brings about Y (a bad consequence). We ask the agent the following question: "Would you have proceeded in achieving X if you had been told that Y would not happen?" If the answer is yes, then the agent intended Y obliquely. The evil effect is irrelevant to the agent's goal, which is to prevail in a just war.

Yet this is a just a terminological move. The question is not conceptual (what counts as intent?) but normative: why exactly are collateral deaths in a just war morally permissible? The most promising candidate is this: action aimed intentionally at an evil is *guided* by that evil (Nagel 2001). Direct harmful agency is, therefore, more *disrespectful* than collateral harm, and therein lies the moral difference between both actions (or between both effects of the same action). The commander who does not aim at noncombatants does not really wish to cause the deaths of innocent persons; he would spare them if he could. The bad effect is nothing to his intent; his action is not guided by evil. The commander who aims at noncombatants wishes their deaths; he does not

want to spare them. Their deaths are essential to his intent; his action is guided by evil. If the distinction is plausible it might be a reply to the contingent pacifist, because it establishes a *moral* difference between direct and oblique harmful agency.

If we incorporate the suggestions that killing for an unjust cause is wrong even if done in compliance with *ius in bello*, and that the just cause must be urgent enough to allow for collateral killings, a revised DDE might read as follows:

An act of war that brings about the deaths of civilians is morally justified if, only if:

1. the warrior uses permissible means of fighting, i.e., permissible weapons;
2. the warrior does not directly *intend* the deaths of civilians;
3. he has a *just cause*, namely defeating an unjust enemy;
4. the act of war is materially conducive to the realization of the cause;
5. the degree to which the act of war is materially conducive to the realization of the just cause is great enough to compensate for the collateral deaths of civilians. (This rule establishes a requirement empirically to weigh harms and benefits.)

The just cause mentioned in (3) must be *compelling enough* to compensate for the collateral harm. This condition recognizes that there are degrees of moral urgency, so that not all just causes will justify collateral harm. The more compelling the cause, the lower the threshold for collateral harm.

A further matter of controversy is how a just warrior should weigh the distribution of harm among all persons (including his own troops) put at risk by the hostilities. The traditional view holds that commanders owe a primary duty to their own soldiers, and only secondarily to the innocent civilians collaterally threatened. Modern writers have suggested, in contrast, that commanders have a moral obligation to transfer some risk to themselves in order to protect civilians (including citizens of the enemy) (Walzer 2006: 155). The question is complex, for many reasons. First, it is just not true that all civilians are innocent. Second, surely a distinction must be drawn *among* those civilians who are mere bystanders and those who expect to benefit from the military action (McMahan 2010: 363). Third, many of the deaths of civilians are caused by the unjust enemy in their reaction against the just warrior. For example, most of the deaths in Afghanistan are caused by the Taliban, but those deaths perhaps would not have occurred had not the NATO coalition initiated the war. It is unclear what is the responsibility of the just warrior for these deaths.

5. Conclusion

Wars have tremendous costs for everyone. Moreover, even the best-motivated governments will have serious epistemic difficulties in assessing the conditions that make war morally permissible. The kinds of questions that a moral warrior asks himself are hard to answer: Is the just cause *compelling* enough to justify the destruction this war will cause? How much will the war (or a particular act of war) contribute *causally* to the achievement of the good result? The fact that governmental decisions about war, even in modern democracies, are shrouded in secrecy increases the public's suspicions and anxieties. Unfortunately, successful wars usually contribute to the careers of the politicians that initiate them. This mere possibility of war creates perverse incentives, so citizens in a democracy should be ready to challenge their governments' decisions to go to war. Still, wars are sometimes (rarely, to be sure) worth fighting. As in many other areas of human

action, there is no mechanical procedure available to identify the right thing to do. But at least we can try to identify the *kinds* of factors that responsible leaders and commanders should weigh in making these fateful decisions.

Related Topics

Aquinas, Utilitarianism and Consequentialism, Nationalism, Global Justice and Politics, Terrorism

References

Aquinas, T. (2002) *Aquinas: Political Writings*, R.W. Dyson (ed.), Cambridge: Cambridge University Press.

Boyle, J. Jr. (2001) "Toward Understanding the Principle of Double Effect," in P.A. Woodward (ed.), *The Doctrine of Double Effect: Philosophers Debate a Controversial Moral Principle*, Notre Dame, IN: Notre Dame University Press.

Cohen, M. (1985) "Moral Skepticism and International Relations," in C. Beitz, M. Cohen, T. Scanlon, and A.J. Simmons (ed.), *International Ethics*, Princeton, NJ: Princeton University Press.

Hobbes, T. (1962) *Leviathan*, New York: Collier Books.

Holmes, R. (1989) *On War and Morality*, Princeton, NJ: Princeton University Press.

Kant, I. (1991) "Perpetual Peace: A Philosophical Sketch," in H. Reiss (ed.), *Kant: Political Writings*, Cambridge: Cambridge University Press.

Kaufman, Whitley (2010) "Review of McMahan, *Killing in War*," *Ethics* 120: 399–404.

Kennan, G. (1986) "Morality and Foreign Policy," *Foreign Affairs* 64: 203–18.

McMahan, J. (2004) "War as Self Defense," *Ethics and International Affairs* 18.

—— (2009) *Killing in War*, Oxford: Oxford University Press.

—— (2010) "The Just Distribution of Harms between Combatants and Noncombatants," *Philosophy and Public Affairs* 38: 342–79.

Nagel, T. (2001) "Agent-Relative Morality," in P.A. Woodward (ed.), *The Doctrine of Double Effect: Philosophers Debate a Controversial Moral Principle*, Notre Dame, IN: Notre Dame University Press.

Rodin, D. (2003) *War and Self-Defense*, Oxford: Oxford University Press.

Sigdwick, H. (1908) *The Elements of Politics*, London: McMillan.

Tesón, F. (2005) *Humanitarian Intervention: An Inquiry into Law and Morality*, Ardsley, NY: Transnational Publishers.

United Nations Charter, U.N. Conference on International Organization, vol. XV.

Vitoria, T. (1991) *Vitoria: Political Writings*, A. Pagden and J. Lawrence (eds.), Cambridge: Cambridge University Press.

Walzer, M. (2006) *Just and Unjust Wars: A Moral Argument with Historical Illustrations*, 4th edn, New York: Basic Books.

Further Reading

G.E.M. Anscombe, "War and Murder," in *Ethics, Religions and Politics: The Collected Philosophical Papers of G.E.M. Anscombe*, vol. 3 (Minneapolis: University of Minnesota Press, 1981); M. Ceadel, *Thinking about Peace and War* (Oxford: Oxford University Press, 1987); C. von Clausewitz, *On War*, trans. by A. Rapoport (Harmondsworth, UK: Penguin, 1995); M. Doyle, "Kant, Liberal Legacies and Foreign Affairs," *Philosophy and Public Affairs* (1984): 204–35 and 323–53; J.B. Elshtain, *Just War Theory* (Oxford: Basil Blackwell, 1992); R. Fullinwinder, "War and Innocence," *Philosophy and Public Affairs* (1975): 90–7; D. Fisher, *Morality and War: Can War be Just in the 21st Century?* (Oxford: Oxford University Press, 2011); W.B. Gallie, *Understanding War* (London: Routledge, 1991); J.T. Johnston, *The Just War Tradition and the Restraint of War* (Princeton, NJ: Princeton University Press, 1981); S.P. Lee, *Ethics and War: An Introduction* (Cambridge: Cambridge University Press, 2012); R.B. Miller, *Interpretations of Conflict: Ethics, Pacifism and the Just War Tradition* (Chicago, IL: University of Chicago Press, 1991); T. Nagel, "War and Massacre," *Philosophy and Public Affairs*, vol. 1 (1971/72): 123–43. W. Quinn, *Morality and Action* (Cambridge: Cambridge University Press, 1994); K. Waltz, *Man, The State and War* (Princeton, NJ: Princeton University Press, 1978); M. Walzer, *Arguing About War* (New Haven, CT: Yale University Press, 2004).

Part VI

POLITICAL CONCEPTS

48
EQUALITY

Thomas Christiano

The idea that persons are equals is at the foundation of modern political thought and practice. Much of the impetus behind the abolition of slavery and colonialism, the rejection of the practice of the master race, of sexism, and of racism, as well as ethnic or national chauvinism is grounded in the thought that all persons are equals. The fundamental ideas underpinning democracy and the practices of liberal rights as well as the contemporary practice of international human rights are explicitly based on the idea of the equality of persons. A dominant idea in contemporary moral and political philosophy is that all persons should be treated as equals.

What is meant when it is said that all persons are equals or are to be treated as equals? Why is it so important? What is the justification for this equality? Whether one holds to a rights-based, a consequentialist or a contractualist view, that persons are equals in some fundamental way is always assumed and nearly never argued for or even fully explained. To be sure, the idea of equality functions in different ways in these different approaches. But there is a core of similarity and they all fail to account for the idea.

1. Equal Status: Conceptual Aspects

First, what does it mean to say that persons are equals? In contemporary thought, the idea of moral status is generally used to refer to the qualification in virtue of which moral principles assign advantages and disadvantages to things, "advantages" being attributions of rights, deserts and goods, "disadvantages" duties, liabilities and negative deserts. That something has a moral status implies that, at a minimum, moral principles in some way assert that it has a right to something, may be owed something, may deserve something, or that it ought to receive a certain good.

One set of problems of equality of status arises for those who think that persons have a distinct moral status, to which distinct advantages apply. Two questions arise: why do different kinds of entities have different moral statuses and why is there an equality of moral status within the set of persons (particularly since some of the persons appear to be very similar to those of the group of animals).

Before we discuss these questions we must be clear on the distinction between equal status and equal treatment. They have a deep moral connection: it is in virtue of the equal moral status of persons that persons are owed equal treatment is at the foundations of contemporary moral and political thinking.

There are really two main ways in which one can think of the connection of equal status to equal treatment. The first is that equal status is a ground of equal treatment. This grounding relation can take different forms. On one account the equality of persons

grounds moral rules as a kind of strategically relevant fact to be taken into account. I am thinking here of the role of equality in the state of nature to the laws of nature in Hobbes. On another more prominent account, the relationship between these two seems to be a kind of relation of fittingness or appropriateness. It is fitting that persons be given some kind of equal treatment because they have equal status. This account asserts that their equal status provides a moral explanation for why one ought to treat persons as equals.

The second main account says that equal status is a kind of global assertion of all or at least the most basic requirements of equal treatment (Dworkin 2011: ch. 9). In this account there is no independent fact of equal status, to which it is fitting to make treatment correspond. The sum of the requirements (or the most basic requirements) of equal treatment constitutes equal status.

"Dignity" is often used to refer to the particular status of persons. There is disagreement about the proper interpretation of Immanuel Kant's view that persons have dignity. Some argue that it is a distinct status that grounds moral principles; others argue that it is merely a summation of basic moral principles (Kerstein 2009). At least on the first kind of account, equal status is quite different in structure from the idea of equal treatment. Equal status is not a practical notion in the sense that it characterizes how one ought to treat persons. It is a morally important fact about persons in virtue of which certain propositions hold about how one ought to treat them. As a consequence, equal status is not something that one can fail at or fail to perform while equal treatment is something one can fail to do.

Derivative and Restrictive Conceptions of Equality

Within the grounding conception of moral status two very different notions of status are in play in modern moral thought. The first is a highly derivative notion employed by consequentialist theories. Indeed, the notion of moral status hardly makes any appearance in consequentialist thought. Because consequentialism usually simply implies that certain goods ought to be brought about, the status of a thing is usually just the result of its having the capacity to have the good in question. The moral status of a thing is derivative here because it just falls out of the moral principle and does not explain any part of the moral principle. For example, hedonistic utilitarianism implies that some function of pleasure and pain ought to be maximized. Any thing that can experience pleasure or pain thereby has moral status and, in a sense, an equal moral status. Some forms of consequentialism can yield a distinction between persons and nonpersons to the extent that the consequentialist theory is based on a conception of the good that differentiates between different kinds of goods, which are had by different entities. Consider the good of rational activity and the good of pleasure. And suppose only persons can have the good of rational activity. If rational activity is a superior kind of good over other kinds, as Mill thought, then personhood will in effect constitute a distinct moral status, albeit a derivative one (Mill 1979: ch. 2). This feature of consequentialism enables it to avoid the difficult problems of status that arise with more restrictive conceptions of status, which we will explore next.

In contrast to the derivative view, a restrictive conception of moral status takes moral status as a kind of ground of appropriate treatment that includes a ground of the moral principles regulating the treatment of the thing. By virtue of possessing certain morally salient features, an entity is due certain treatment. Furthermore, it asserts that because of their statuses different things are due different kinds of treatment depending on their statuses. Hence, restrictive conceptions of status imply that some entities are

to be accorded equal treatment whereas other entities are to receive different treatment because of their distinct statuses. A restrictive conception of status will imply a conception of equality of status and inequality of status. For example, many think that persons are equals while animals and persons are unequals, though both have moral status. The idea is that if an entity has a certain status then it ought to be treated in a certain way, and so those with the same status ought to be treated in the same or equal way.

One way to see the distinction between the derivative and the restrictive conceptions of status is that the former seem often to imply that one ought to bring about more entities of the sort that have the relevant status because that is necessary to bring about more of the morally desirable good. Restrictive views tend to make treatment depend on status while status itself is not something that is to be promoted. Status, on the restrictive conception, is often denoted by the term "dignity" and is often correlated with the attitude of respect.

2. Basic Equality

The equal status of persons is a basic moral equality. It does not imply equal status in all the different contexts in which status matters in complex societies. For most contemporary views, basic equality among persons is a background requirement that is compatible with a number of different kinds of inequality in everyday life. Let us distinguish some main kinds of inequality of status among persons that play a role in moral and political thought and that are compatible with basic moral equality.

First, different kinds of unequal status can be justified for different aspects of legal and political systems. For instance the idea of political equality involves equality in political status and the requirement that all sane adult members of a society have an equal right to participate in politics. This involves at least some variant of the principle of one-person/one-vote, equal rights to association, conscience and expression as well as equal rights to run for elected office. These democratic rights are extended to all, or nearly all, sane adult members of the political community, but they do not extend to visitors or to small or adolescent children. In some societies convicted felons are excluded from the vote and other rights while they serve their terms and some societies permanently exclude convicted felons from the vote. Furthermore, the rights of property and contract are usually thought to be attributable to all or nearly all sane adult persons but not completely to children, the severely mentally challenged or convicted felons. Still few would argue from these unequal political statuses that these political societies are committed to the idea that children are not the equals of adults in some fundamental moral sense. Usually the idea is that the political equality among adults and exclusion for children from the political process is justified by the idea that children cannot make use of the political system in a way that is comparable to adults. They cannot protect their rights or their interests because they have a much less secure sense of these interests or rights or the means to protect or advance them. The same can be said of children's abilities to use property and contract to advance their interests and concerns. Here the differences in moral status between children and adults are relatively superficial and are related to differences in abilities to judge wisely in practical matters. Convicted felons may be excluded from the vote or from full rights of property on grounds of desert or sometimes it is claimed they are excluded on grounds of incapacity due to corruption. In contrast the sameness of basic moral status is grounded in sameness of basic capacities, it is usually thought.

Second, fundamental equality is normally thought to be consistent with a legitimate partiality of persons towards other persons. Persons are members of groups such as families, friendships, voluntary associations, and nations that may allow some partiality of the members towards each other. In these instances, co-members are treated with greater concern than nonmembers, often members will do more for fellow members, and might even have more extensive and weightier duties towards them. Within the family, for instance, many think that parents act wrongly if they treat other children with the same concern as their own. Some argue, more controversially, that we have more extensive and weightier duties to fellow citizens or fellow nationals. Here persons within the nation state might be thought to be owed equal treatment in terms of basic rights, political rights, and distributive justice, among other things whereas persons outside of that state are not. Yet, these purportedly justified differences in treatment are not thought to be inconsistent with recognition of the basic equality among all persons (Blake 2001).

A third form of inequality may be thought to be compatible with basic equality: inequalities may arise in ways that are acceptable from the standpoint of equality. For instance, differences in merit might justify some differences in distributions of goods relevant to merit. Or in the case of negative deserts, some might be thought to be deserving of harsh treatment because of their actions while those who are innocent may not be. Such inequality of treatment is thought also by many to be compatible with basic equality. Finally, inequalities that arise through the voluntary activities of persons have been thought to be legitimate under certain circumstances as well (Cohen 1989; Nozick 1974).

3. The Grounds of Equal Status

Equal status is often thought to be a ground of equal treatment, but what are the grounds of equal status?

In Hobbes's rendering, people are equal by nature because in a world without government people are threats to each other's security. The equality is only a very rough equality. This equality is a strictly non-moral, factual equality. And it is an equality that only obtains within the state of nature. For Hobbes, this mutual threat capacity is the basis for the laws of nature including the law that requires that one treat others as equals. The descriptive equality generates, given the circumstances of the world, a kind of moral equality, in the sense that everyone is subject to the same basic moral rules in the same way (Hobbes 1994: ch. 13). But the mainstream idea of equality derives, I think, from the scholastic idea that all persons have a status in virtue of which they are not to be merely used by one another (Vitoria 1991). This was the central consideration behind the Spanish Scholastics' questioning of the Spanish conquest and domination of the American Indians, which itself was often justified on the grounds that the American Indians were natural slaves. This equal moral basic status is grounded for the Scholastics in the idea that human beings are created in the image of God. What this amounts to is complex, but on this account the key feature of human beings is their possession of reason (Aquinas 2002: 131). In this tradition, Locke argues that human beings are moral equals because they are members of the same species with "all the same advantages of Nature and the use of the same faculties." Since they are members of the same species they are owed equal treatment (Locke 1988: 269). This is at the origin of the contemporary view that all human beings are equal.

This view is also at the origin of the idea that human beings have a higher status than other animals. The possession of the capacity for reason that involves moral and speculative thought is thought to give humans a higher status than animals and a right to use those animals for their own benefit. One might wonder why. Here modern thought faces a perplexity that the Scholastics did not. The Scholastics argue, within a teleological conception of nature and living things, "less perfect things exist for the sake of more perfect things" (Aquinas 2002: 131). Perfection is understood in terms of the extent to which a being is capable of realizing the end for the sake of which all things are created. Since the end of creation is God and human beings are capable of contemplating God as well as acting from a conception of law while other animals are not, human beings are more perfect than other animals and thus have a higher status.

Without the teleological conception of nature, modern thought has had a harder time showing why some kinds of beings have a higher status than others. Some have relied on intuition to establish this thesis. While this is highly intuitive for many, the persistence of controversy about the relative status of human persons and other animals suggests the need for intellectual reinforcement of this intuition. In the most profound modern treatment of this question, Kant argued that persons must view themselves as having a moral status as distinct from all other kinds of things. His argument was that persons must see themselves as having the authority to impose the moral law on their selves and thus must see themselves as objects of respect and as having absolute worth. Nothing else in nature has this kind of authority, he argued, and so nothing else can have this status (Kant 1997: 81–5). One worry with Kant's view is that it seems to imply no moral status for animals that fall short of personhood, which is quite counterintuitive.

The possession of the capacity for reason (understood in very different ways for different views) is the most widely accepted ground for the distinct status of persons. Kant married this basis with the capacity for self-determination as the basis for the status of personhood. Some have also included the capacity for self-consciousness (Williams 1976). The possession of the capacity for consciousness and in particular the experience of pleasure and pain are the most common bases of the moral status of animals. And life is the basis of the distinct status of living things. But once we acknowledge these distinct moral statuses and we reject the scholastic basis of hierarchy, it is unclear how we can justify the hierarchy of status that is implicit in much modern moral thinking.

4. A Contemporary Puzzle about Equal Status

A related and very vexing problem in contemporary moral and political philosophy concerns the nature and grounds of equality and inequality of basic moral status on the restrictive understanding of status. The idea that persons have a distinct and equal moral status is accepted by many different views in some form or other. And it seems necessary to many accounts of basic moral principles. But a problem arises with different conceptions of the ground of the moral status of personhood. The basic concern is that if status is grounded in certain valued traits that persons have, and these valued traits are present to different degrees in different persons, then it is hard to see how persons generally can have the same status. One might call this the "status trilemma." One, the status of persons is grounded in the extent to which they have certain distinctive traits. Two, persons have the status-conferring traits to relevantly different degrees. Three, persons have equal status.

The first proposition in the trilemma asserts that the statuses of persons are grounded in the degree to which they possess certain distinctive traits. Let me explain the ele-

ments of this proposition. First, the idea is that the status of persons is grounded in distinctive traits. This is necessary to ground the distinctive moral status of persons. The thought is that persons have a distinctive status, higher than that of animals, that calls for distinctive moral norms. So the status of persons is not thought by many to be grounded in the mere capacity to experience pain, because the experience of pain is common to many animals. As we have seen (Section 3), many have thought that the distinctive status of persons is grounded in some way in their rationality. Rationality is understood in different ways on these accounts. One may be just the capacity to make coherent choices on the basis of given ends, another the capacity to appreciate value and to act on that appreciation. Although there may be significant differences between these grounds, some are at least partly shared by animals. The capacity to make coherent choices, the capacity to make moral choices, and capacities for rational thought in relation to action and understanding are all, to some extent, characteristic of animals. So if these traits ground the distinctive status of persons then it must be the high degree to which they are present in persons that grounds the status of persons.

But just as humans may possess a higher degree of such characteristics than animals, some humans possess them in a higher degree than others. I am thinking here of basic capacities of persons. Though these are difficult to measure we have some sense that some persons are more capable of rational thought in relation to thought and action than others are. We have some sense that some persons are better capable of appreciating moral or aesthetic value than others. Of course it is very hard to pull apart the contributions of capacity from the contributions of environment but some differences of capacity are likely to be present.

From these two premises, it would appear to follow that persons do not have equal moral status. The thought is that the status of persons is grounded in the possession of a certain capacity, because of the high value of the capacity. But if that capacity can come in degrees, and some persons have greater capacities than others in respect of the relevant features, then those persons have something of even greater value than those who have a lesser capacity. If we value the capacity then we must value the greater capacity even more. So those who have the capacity to a greater degree have higher status.

The standard way to defeat this inference is a move that denies either the first or the second premise. It says, abstractly speaking, that the differences of capacity beyond the threshold are negligible. But this can be true in two different ways.

The first way is to accept that there are differences but assert that they do not matter. The main effort has been to assert that extent of possession of the key traits determines status only in part. The thesis, asserted by Rawls and others, is that as long as a person possesses traits such as rationality and morality beyond a threshold level, then they possess the status of personhood. Beyond that threshold level persons can be different in extents to which they possess the traits but still have equal status. This move denies the first proposition of the trilemma (Rawls 1971: 508).

The second move denies the second proposition in the trilemma by asserting that the differences between different persons above the threshold are very small or nonexistent. The difference between Einstein or Gandhi and most others is not very great on this account.

Conceptually these are neat resolutions of the problem. And they are intuitively attractive. But they do not provide a satisfying rationale. One problem is that no rationale is offered for thinking that there is a threshold or for determining where the threshold is. Normally, a threshold is thought to be present when some remarkable new prop-

erties appear at the threshold point. The threshold point must present some kind of very strong discontinuity in order to qualify as a genuine threshold. If we are thinking of a simple graph with degree of capacity on one axis and type of entity on the other, for example, the slope of the graph must change quite significantly around the threshold. The threshold idea has two parts: one, there is a huge difference between the capacities before the threshold is reached and the capacities at the threshold and, two, there is a negligible difference between the capacity at the threshold and those significantly beyond the threshold. To be sure, the threshold can be quite vague so that we don't know exactly when we have crossed it. But the thresholds suggested by Rawls and others do not have these properties, at least no such discontinuity is defended. Hence the idea that the status of personhood appears at a particular threshold and then remains unchanged after that threshold appears ad hoc. There just aren't any reasons offered for why this should be thought of as threshold concept.

A second reason for doubt is that we have a clear idea, I think, of relevant increases in capacity beyond the threshold of personhood though we may know of no instances of this type. Presumably God and the angels would be of a higher status than human persons, and this would be because of their vastly greater rational capacities. Certainly this has been the traditional view of these entities' moral statuses. But this defeats the simple claims that any difference of traits beyond the threshold makes no difference or that there are no such differences.

A second effort to reject the first proposition is to say that equal status is grounded in pragmatic considerations. Here the idea is grounded in a deeper idea that need not involve equal status of the sort in question. For instance, one might be a utilitarian and argue that because there is so much uncertainty about who is more rational and because the inclination to bias tends to distort judgment on these matters in counterproductive ways, and finally because people are likely to resent being treated as inferiors, it is better simply to treat all persons as basic equals. This, in effect, abandons the idea that equal moral status is a fundamental moral idea (Arneson 1999).

A third effort to resolve the problem is to argue from the thesis that we owe respect to rational beings to the idea that they have equal status. This argument proceeds through the premise that we owe rational beings a duty of epistemic abstinence, which asserts that we owe it to them not to measure their relative capacities. This is because we owe it to them not to look too deeply inside them. The idea is that one owes a duty of respect to all those who qualify minimally as rational beings, not to inquire further into their differential capacities (Carter 2011).

The first worry about this approach is that it may seem to beg the question. The view allows us to inspect the capacities for the purpose of determining minimal qualifications but then disallows us from searching for further differences. But if a change of status could arise as a consequence of increased capacity, then one should be allowed to look to see whether that new threshold has been crossed. As long as we have a clear idea that there can be higher moral status, it is unclear why this view does not foreclose a morally important possibility, and if we have the idea that there cannot be a higher moral status, then the case is made by defense of that assertion alone. The second worry is that it is not clear that the view answers the problem. For the consequence of the duty of "opacity respect" is not that persons are of equal moral status but that at most we must treat them as if they had equal moral standing or perhaps that we should ignore their relative moral status in our dealings with them. It does not actually answer the question of relative moral status, rather, it merely makes relative moral status forbidden from view.

A fourth possible response to the difficulty might be to reject the idea that the moral status is a ground of the moral duties. One might think instead, with some interpreters of Kant's doctrine of dignity, that the notion merely summarizes and stands for the duties we owe persons. This deflationary account of dignity provides an answer but at a cost. The cost is that it deprives us of one very powerful kind of explanation of the duties we owe to others.

5. Equal Treatment

Comparative and Non-comparative Equality

Probably the most fundamental distinction regarding equal treatment concerns the distinction between non-comparative conceptions of equality and comparative conceptions. Vitoria seems to suggest a kind of absolute requirement that derives from the status of each person: that each person "does not exist for another's use, like an animal, but for himself" (Vitoria 1991: 249). Nozick's historical entitlement conception of justice is a case of a non-comparative approach to all the basic norms of justice (Nozick 1974: ch. 7). The idea here is that each person has rights and they all have the same rights. But justice is not done, on this account, by cutting down on the possession or fulfillment of the rights of one person in order to increase the possession or fulfillment of the rights of another. Though each person has the same rights, each person is wronged separately by any action that violates or impinges the right, regardless of the extent to which others' rights are respected. It is an affront to any person's dignity that they do not enjoy the exercise of the same rights as others, assuming none of the inequality legitimating processes above have come into play. But that affront occurs because the person has not had his rights respected. The inequality can play at best a secondary role. For example, if respect for my rights is unequal to those of some other person and I am not relevantly different, there is reason to think there is injustice, but not because the inequality itself is unjust but only because at least one person must be mistreated. One of the two levels must be the correct level associated with a being of the sort we are talking about. And both persons should be at that level. Here the idea is that the right is directly grounded in the person's status.

One might wonder here whether rights are grounded merely in status of the right-holder, since beings of superior status may have different duties to beings of inferior status than they have to each other. For instance, one might think that God could have rights that might include the right to curtail the rights of persons. God may be permitted to coerce human beings in contexts in which human beings are not permitted to coerce each other. This might be best expressed by saying that human beings do not have rights against God not to be coerced though they do have such rights against other human beings. There is a structural analogy here to the idea that there are certain things that parents may do to children that children are not permitted to do to each other. If this observation is correct, the rights of persons may be grounded not only in their own statuses but also partly in the status of the duty holder. Something like this view seems to be held by Locke in his explication of equality among persons in the state of nature. Locke argues that "all the power and jurisdiction [among persons] is reciprocal, none having more than another." But he accepts the possibility that God could place one person above others (Locke 1988: 269).

A fairly clear-cut instance of comparative equality that is widely accepted is the principle of one-person/one-vote. The idea is not, as with the above case, that the

possession of a vote is suited directly to personhood. What one person rightfully possesses is a function of what others rightfully possess. And the proper number of votes each person has is a function of the total number of votes available to the community. On the egalitarian distribution, if one person has four votes and two others have one apiece, then the one should redistribute one vote to each of the others. To be sure, inclusion itself is of great importance here, but it is of the essence that each has the same vote as the others whatever that number of votes is.

Other instances of the comparative/non-comparative distinction are the differences between sufficientarian and egalitarian principles of distributive justice and those between prioritarians and egalitarians. A sufficiency view requires that each person has enough goods for an adequate or a decent human life. Sufficientarian and egalitarian principles are compatible with each other so that one could be an egalitarian while recognizing the special urgency of achieving some minimal level of goods for each person. The prioritarian principle defended by Derek Parfit asserts that persons ought to be made better off and that the moral urgency of making them better off increases as people are worse off. Hence there is more goodness in increasing the well-being of a worse-off person than the very same increase in the well-being of a better-off person (Parfit 1991). These principles apply to individuals whether in relation to others or not. They imply that there will be some tendency to equal distribution of well-being when the prioritarian principle is satisfied, but it is a weak tendency and the value of equality is entirely derivative on this view. It is not of fundamental importance that two people are close to each other in level of goods, it only matters what the total goodness is.

Some have worried that any comparative conception of equality of treatment is subject to a difficult objection called the "leveling down objection" (Frankfurt 1987; Parfit 1991). The worry is that if the distribution of some good must be equal in a comparative sense so that what a person ought to have depends on and is meant to be equal to whatever others have, then a distribution in which everyone has very little is better in this respect than a distribution in which some have more than others but everyone has more than in the egalitarian distribution.

There have been a number of responses to this worry among those who defend a comparative conception of equality. One answer is to accept the implication that a comparative conception of equality could recommend on its own leveling down. But it is argued that since equality is only one among a number of moral concerns, one need not think that leveling down is required all things considered. The idea is that it would be worse in some respect that some persons are better off than others even if they are all better off than they would be were they equally well off. But this carries no implication that one must endorse the worse-off state of affairs (Temkin 2009).

Another response is that the principle of equality as a comparative principle does not imply leveling down; indeed, one might argue that it prohibits leveling down. This answer is based on a conception of equality that functions a bit differently from the above notion. The idea is that the principle of equality implies that for each situation there is an ideal egalitarian distribution that is Pareto optimal (where the rankings are articulated in terms of whatever things are to be distributed). This distribution might not be feasible in the situation though it is connected to the situation in that it takes all the good that can feasibly be produced in the situation and distributes it equally. In the relevant situation, all feasible distributions are evaluated in terms of how well they approximate the ideal distribution. And all distributions in the situation that are not the same as the ideal distribution are unjust. Whether the principle so defined implies

leveling down will depend on the nature of the approximation principle. But it seems plausible to assert an approximation principle that affirms that a distribution of goods in a situation that makes everyone worse off is not as close an approximation to the ideal distribution as a distribution in which everyone is better off. The distribution in which everyone is better off is such that each person is closer to what they would have in the ideal egalitarian distribution than they would be were they to be in the Pareto suboptimal distribution (Christiano and Braynen 2008).

Some might protest against this solution that it looks like the above purported egalitarian principle is a mix of egalitarianism and Paretianism. But the view on offer starts from the assertion that there is an internal connection between the principle of equality and the desirability of the things that are distributed equally. If the things, be they welfare, opportunities for welfare, rights or primary goods, were not themselves desirable, there would be no demand for equalization. Given this internal connection, we have a reason to think that an appropriate principle of comparative equality will strongly favor Pareto optimal distributions.

One might worry that the principle of equality does suggest that sometimes complete justice will not be feasible as when the ideal egalitarian distribution cannot be achieved. But we do commonly accept principles of justice that affirm ideals that we cannot fully achieve. One such principle is that the innocent must not be punished. But we are quite sure given the imperfections of our cognitive abilities that a criminal trial procedure will, no matter how conscientiously used, sometimes produce the wrong verdict. This is an injustice. The ideal of never punishing the innocent is an important and widely accepted one despite the infeasibility of fully achieving it. It has practical significance for us to the extent that we think that we can approximate the ideal. I think the same can be said for the principle of equality as I have articulated it.

Related Topics

Aquinas, Hobbes, Locke, Kant, Mill, Contractualism and Political Liberalism, Utilitarianism and Consequentialism, Natural Law and Rights Theory, Luck Egalitarianism, Libertarianism, Desert, Nationalism, Human Rights and Cosmopolitanism, Rights

References

Aquinas, Thomas (2002) *On Law, Morality and Politics*, 2nd edn., trans. Richard J. Regan, Indianapolis, IN: Hackett Publishers.

Arneson, Richard (1999) "What, If Anything, Renders All Humans Morally Equal?" in Dale Jamieson, ed., *Singer and His Critics*, Oxford: Blackwell Publishers.

Blake, Michael (2001) "Distributive Justice, State Coercion and Autonomy," *Philosophy and Public Affairs* 30(3): 257–96.

Carter, Ian (2011) "Respect and the Basis of Equality," *Ethics* 121(3): 538–71.

Christiano, Thomas and Will Braynen (2008) "Inequality, Injustice and Leveling Down," *Ratio* 21(4): 392–420.

Cohen, Gerald (1989) "The Currency of Egalitarian Justice," *Ethics* 99(4): 906–44.

Dworkin, Ronald (2011) *Justice for Hedgehogs*, Cambridge, MA: Harvard University Press.

Frankfurt, Harry (1987) "Equality as a Moral Ideal," *Ethics* 98(1): 21–43.

Hobbes, Thomas (1994) *Leviathan*, trans. Edwin Curley, Indianapolis, IN: Hackett Publishers.

Kant, Immanuel (1997) *Groundwork of the Metaphysics of Morals*, in *Practical Philosophy*, trans. Mary J. Gregor, Cambridge: Cambridge University Press.

Kerstein, Samuel (2009) "Death, Dignity and Respect," *Social Theory and Practice* 35: 505–530.

Locke, John (1988) *Two Treatises of Government*, ed. Peter Laslett, Cambridge: Cambridge University Press.

Mill, John Stuart (1979) *Utilitarianism*, ed. George Sher, Indianapolis, IN: Hackett Publishers.
Nozick, Robert (1974) *Anarchy, State and Utopia*, New York: Basic Books.
Parfit, Derek (1991) "Equality or Priority?" University of Kansas, Lindley Lecture.
Rawls, John (1971) *A Theory of Justice*, Cambridge, MA: Harvard University Press.
Temkin, Larry (2009) "Illuminating Egalitarianism," in Thomas Christiano and John Christman, eds., *Contemporary Debates in Political Philosophy*, Malden, MA: Wiley-Blackwell.
Vitoria, Francisco (1991) *Political Writings*, ed. Anthony Pagden and Jeremy Lawrence, Cambridge: Cambridge University Press.
Waldron, Jeremy (2000) *God, Locke and Equality*, Cambridge: Cambridge University Press.
Williams, Bernard (1976) "The Idea of Equality," in *Problems of the Self*, Cambridge: Cambridge University Press.

49
FREEDOM

Katrin Flikschuh

The post-WWII philosophical history of liberal freedom can be summed up as a theoretical development from two, to one, and finally three concepts (or conceptions) of liberty, or freedom. The debate begins with Isaiah Berlin's 1959 publication of "Two Concepts of Liberty," comes to rest temporarily in Gerald MacCallum's analysis of freedom as a "triadic concept" (MacCallum 1967), and has more recently erupted into a dispute among descriptivists and republicans over the nature of negative freedom. This history is not old, nor is it especially original so far as contested substantive meanings go. One could read "Two Concepts" as offering little more than an (ideologically tainted) historical narrative of what we have come to mean by liberty given what first "the ancients" and then "the moderns" meant by it. However, the abiding appeal of "Two Concepts" lies less in its wealth of historical reference points and more in its philosophical thesis, which is a version of the more general thesis that concepts denote objects, or, yet more generally, that (all) concepts have determinate content. On this view it is possible to distinguish between correct and incorrect applications of the concept of liberty: in Berlin's words, "everything is what it is; liberty is liberty, not equality or fairness or justice or culture" (Berlin 2002: 172).

Berlin's view, both that the meaning of freedom is determinate and determinable and that the negative concept of liberty as non-interference offers the most promising point of departure for an accurate understanding of what freedom *really* means, continues to exert its influence on freedom theorists. It has not gone unchallenged. There have been numerous statements in support of positive freedom, including its account of freedom as an inner cognitive moral state associated with self-knowledge (Taylor 1985; Christman 1989). Within mainstream liberal political philosophy, however, the most formidable early challenge came in the form of MacCallum's contestation of Berlin's distinction between *two* freedom concepts and MacCallum's contrary claim that any complete freedom analysis must acknowledge negative and positive elements as constitutive aspects of *one and the same* concept. On this account, the positive element complements the negative: we are never merely free *from* interference but are always at the same time free *to* do (or become) particular things (or persons).

Although MacCallum endeavored to restore the unity of the concept, his analysis did little to rehabilitate the substantive commitments associated by Berlin with the positive tradition. MacCallum's legacy lies in the rigor his analysis brings to bear on what amounts, in retrospect, to a first move towards a descriptivist approach to negative freedom. Recent descriptivism about negative freedom reached its zenith with Hillel Steiner's defense of "pure negative freedom" (Steiner 1994). "Purity" refers to the absence of any evaluative dimensions that would undermine Steiner's endeavor to arrive at a

quantitative freedom theory designed to enable us to determine the exact extent of one person's freedom relative to that of others. Steiner is interested not in how free a person feels, but in how free they in fact are. Similar freedom conceptions have been offered by Ian Carter (1999) and Matthew Kramer (2005), though Kramer's view that some evaluative commitments must enter into our assessment of a person's overall freedom renders his account less "pure" than Steiner's original analysis.

The most recent challenge to descriptivist conceptions of negative freedom comes from proponents of republican freedom (Skinner1998; Pettit 1997). According to them republican freedom offers a third alternative conception that is reducible to neither liberal negative freedom nor positive freedom. Republican freedom is the quest for non-domination. Non-domination depends on others' recognition of a person as their moral equal. This re-introduces a normative dimension into the negative freedom debate, yet in order to delimit themselves from the positive tradition, republicans disavow commitment to any further claims about the rationality or goodness of persons' choices.

Clearly, the contested terrain, since Berlin's "Two Concepts," has been almost entirely that of negative freedom. Before examining the sketched line of development in more detail below it is worth asking, briefly, what has become of positive freedom. Berlin's association of positive freedom with the rise of totalitarianism has proven a powerful deterrent to serious engagement with that tradition. Yet despite having been officially unmasked, disgraced, and repudiated, positive freedom has quietly flourished within the parameters of mainstream liberalism—albeit under the guise of another name. It now goes under the title of "personal autonomy" and in that reincarnation enjoys a moral reputation unequalled by any other liberal value or ideal, including that of negative liberty.

Personal autonomy, understood as each person's capacity to form, pursue and revise her own conception of the good, shares little with the Berlinian association between positive freedom and a politics of truth; however, the conceptual link between freedom and reason, as well as that between reason and morality resurfaces persistently. Personal autonomy has become a mainstay of liberal political justification—its value, promotion, and protection underpin arguments about distributive justice, human rights, and myriad public policy provisions.

1. Freedom: Negative and Positive

In the political context, Berlin says, the idea of freedom is usually considered in relation to questions about coercion and obedience. The proponent of negative freedom asks, "what is the area within which the subject—a person or a group of persons—is or should be left to do or be what he is able to do or be, without interference by other persons?" The advocate of the positive tradition asks, "What, or who, is the source of control or interference that can determine someone to do, or be, this rather than that?" Though Berlin says that "the two questions are clearly different" (Berlin 2000: 169), the relevant sense in which they differ is not immediately obvious. The notion of interference figures in both. However, the first question emphasizes the absence of interference; the second assumes the legitimacy of a certain type or source of interference. While for the proponent of negative freedom,, interference and freedom are antithetical, advocates of positive freedom view some forms of interference as freedom enhancing. By the same token, while for the former interference entails unfree-

dom, for the latter the absence of a certain kind of interference could be indicative of unfreedom.

These contrasting views on interference are explicable partly in terms of the underlying metaphysical commitments of proponents of negative and positive freedom, respectively. The negative theorist as Berlin conceives him thinks of freedom as the capacity for choice and action, and views it as constitutive of personhood: persons *are* free naturally. Assuming a view of persons as naturally free, others' interference with a person's choices places him or her, *prima facie*, in a state of unfreedom. By contrast, positive theorists think of freedom as either a moral or a social achievement (or both): persons have an antecedent *capacity* for freedom but that capacity must be nurtured and developed, by oneself or others, into a *state* of freedom.

Berlin favors freedom as non-interference. He rejects the thought of a person's freedom as requiring nurture and development. His hostility in this regard stems from a suspicion towards prescriptivism about freedom, i.e., the view according to which for a given action to count as free it must be a morally good, or rational action. However, although Berlin rejects prescriptivism about freedom, he is not a descriptivist or physicalist about freedom. A physicalist follows Thomas Hobbes in thinking of freedom as consisting essentially in unimpeded movement. According to Hillel Steiner, the most prominent contemporary freedom physicalist, persons are free to do whatever they actually do (Steiner 1994). Young children are free to do what they actually do. Physically unprevented by his parent, the toddler who falls down the stairs is free to do so. Berlin disagrees: to be free, the toddler would have to will himself to fall down the stairs. An agent's will must guide his action by its choice for the action to count as freely performed. Absent the link between will and action, free actions would not be imputable. The imputability of their actions to agents centrally motivates Berlin's resistance against governmental interference (Berlin 2002: 4–15). For one to be free is for one to take responsibility for one's actions. Where governments prescribe how he ought to act as a free person he is not being treated by them as a responsible and accountable agent. Thus, even though Berlin is not a prescriptivist about freedom, he assumes some link between freedom and responsibility for choice and accountability for action. Nonetheless, Berlin says nothing about the maturation of the capacity for choice-directed action. In "Two Concepts," the idea of freedom as choice-directed, imputable agency is a suppressed premise, eclipsed by Berlin's exclusive focus on non-interference as an external freedom condition. Berlin's suppression of this premise places his negative freedom conception in problematic proximity to ideas of self-control and self-development he associates with the positive tradition. As we shall see, the overlap at the level of self-directed willing comes to haunt his original account of negative freedom.

The fact that even in "Two Concepts" Berlin thinks of freedom of action as emanating from the will's capacity for choice is evident from his view of coercion as the quintessential form of freedom-interference.

> I am normally said to be free to the degree to which no man or body of men interferes with my activity. Political liberty in this sense is simply the area within which a man can act unobstructed by others. If I am prevented by others from doing what I could otherwise do, I am to that degree unfree; and if this area is contracted by other men beyond a certain minimum, I can be described as being coerced, or, it may be, enslaved.
>
> (Berlin 2002: 169)

Only other persons can render one unfree. Interferences that do not consist in the intentional disturbances of my proposed actions by others can render me unable but not unfree to act. "Coercion implies the deliberate interference of other human beings within the area in which I could otherwise act" (2002: 169). If freedom is choice-directed action, unfreedom is obstructed choice, not obstructed action: I can be coerced by another into doing what I would not otherwise do. I am then unfree in the sense of being directed to act by a will other than my own.

The view of coercion as deliberate interference and of deliberate interference as the distinguishing mark between inability and unfreedom implies that where another unintentionally prevents me from acting as I choose or intend, she renders me unable but not unfree to act. An inattentive person might unintentionally interfere with my walking down the road along my intended path. Although I must swerve in order to avoid her, her unintended obstruction has not rendered me unfree (Strawson 1974). Similarly, an established economic system might leave some people with fewer options to do things than other persons. Where no participant within that system intends to obstruct their choices, the system renders them unable but not unfree to exercise choice. If freedom is the capacity to bring about chosen states of affairs, and if unfreedom is a function of others' intentional obstruction of one's choices, not all physically unobstructed acts a person performs are free actions, and not all physical preventions by others of a person's actions are freedom-interferences.

Berlin views governments' primary relation with subject citizens as coercive; it follows that governments are always freedom-curtailing. This does not in itself render governments illegitimate; constraints on freedom are often justified. Nonetheless, left unchecked, the state has a propensity towards freedom curtailment. Hence the negative freedom theorist's pre-occupation with the question of *how much* freedom a person needs to possess in order to retain their status as a free being. The negative freedom theorist is concerned "with the area of control, not with the source" of freedom (Berlin 2002: 176). That we are free is taken as given: the question is how much of our natural freedom we can legitimately be asked to give up without loss of our status as persons. While there are myriad perfectly legitimate grounds for governments' imposing freedom restrictions—justice, equality, national security, etc.—no government may impose freedom constraints arbitrarily, or pretend that freedom constraints are anything other than what they are. Moreover, while different governments might impose different freedom restrictions depending on context and circumstance, "there ought to exist a certain minimum area of personal freedom which must on no account be violated" (Berlin 2002: 171).

The cardinal error of theorists of positive freedom is their failure to appreciate that governments are freedom-constraining, not freedom-enabling. Governments sometimes constrain freedom legitimately, sometimes illegitimately, but they never *make* persons free. The mistaken thought that they do is encouraged by the positive theorists' erroneous view of freedom as a condition to be developed. Berlin acknowledges that the connection between positive freedom and totalitarian government is contingent; however, he believes that the positive tradition has left itself vulnerable to political abuse. He draws a series of parallels between the positive tradition's penchant for thinking of freedom as an inner state achieved through self-discipline and totalitarian rulers' cajoling subject populations into the mistaken belief that the repressive measures imposed on them are liberating them from some state of unfreedom. The positive freedom theorist influenced by Stoic thinking conceives of freedom in terms of the will's suppression

of unwanted desires; analogously, a totalitarian ruler urges his subjects to embrace the hardships of nation building or the disciplines of communist justice, as the mark of "true" freedom.

It is also in connection with the positive tradition that Berlin raises the charge of conceptual confusion. Stoics and Kantians confuse freedom with self-control or moral goodness; revolutionaries-cum-totalitarian-leaders confuse it with justice or national self-determination. Berlin's ruminations over the positive tradition's supposed totalitarian proclivities are generally ignored today, which is not to say they haven't left a deep impression. Of more directly noticeable impact has been his rejection of an intrinsic connection between freedom and morality or freedom and reason. Both are associated with Kant, though Berlin traces them backwards to Stoic thought and forward to Hegel and Mill. The Kantian idea of freedom as the will's independence from determination by sensible inclination has its roots in Stoic acquired indifference to the pull of desires. Similarly, the idea that free action rests on reasoned deliberation and judgment is prominent in Kant as well as Hegel and Mill, both of the latter of whom make further claims between freedom and truth. The widespread rejection among liberal thinkers of so-called "moralised" freedom conceptions (Cohen 1995) could be due to Berlin's sustained attack on the positive tradition in these regards.

2. Freedom as a Triadic Concept

Among the most decisive early interventions in the debate generated by "Two Concepts" is that by Gerald MacCallum, who rejects Berlin's distinction between two concepts of freedom, arguing that the negative/positive distinction must be understood as internal the same general concept. While there might be different freedom traditions, and while these may differently specify and weight the contents of the three "term variables" stipulated by MacCallum, all operate with the same general concept and must acknowledge the constraints of the concept's "triadic structure." That structure has the following general form: "x is (is not) free from y to do (not do; become, not become) z," where, "x ranges over agents, y ranges over preventing conditions, and z ranges over actions or conditions of character or circumstance" (MacCallum 1967: 314). "Free from" and "free to" are meant to correspond with Berlin's "negative" and "positive" distinction. According to MacCallum, any apparently "negative" statement—"x is free from constraint y"—implicitly presupposes a "positive" enabling condition ("z") which x is therefore free to do. Conversely, any positive statement to the effect "x is free to z" contains an elliptical reference to a constraint ("y") x is free from. Anyone who does not implicitly or explicitly invoke both negative and positive term variables in a statement purportedly about freedom is applying the concept incorrectly.

On this stipulation, Berlin's negative freedom statements are elliptical freedom statements. When Berlin says things like "I am free to the degree to which no man or body of men interferes with my activity" (Berlin 2002: 169), "my activity" must be treated as an elliptical reference to what, specifically, I am free to do given another's non-interference. Defenders of Berlin have protested that he can intelligibly remain uncommitted to MacCallum's triadic formula. A person may be free from others' interference to pursue or not pursue any number of alternatively available activities or inactivities. The demand that we specify what particular constraint an agent is free from to do which particular corresponding action violates our basic freedom intuitions: a characteristic feature of freedom as non-interference is precisely that one is unprevented from doing any

number of things, including doing nothing. John Gray has insisted on the *dyadic* structure of Berlinian negative freedom, involving only two term variables—x and y—and omitting z on the grounds that its specification constrains x to do or not do something particular in consequence of y's non-interference (Gray 1989: 45–68).

One might view Gray's response as a good example of an elliptical reliance on "z": to insist that y's non-interference leaves x free to do or not do any number of things is not to deny that x is now free to do (or not do) *something*. One need not specify any particular content for z, but can hold fast to the "non-specific value" of z so long as one leaves logical space for z as necessary to an analysis of the concept of freedom (Carter 1999: 11–30). However, a defender of the dyadic structure of freedom might retort that term variable "z" is logically required only if one specifies term variable "y" in such a way as to allow preventing conditions other than other agents. MacCallum allows "constraints, restrictions, interferences, and barriers" (MacCallum 1967: 314) of various kinds—not just other persons can render one unfree. Another's coercive interference with my will is not a paradigm case of freedom prevention; the impersonal economic system is likewise a possible freedom-preventer. A defender of Berlinian negative freedom may worry that the distinction between "being unfree" and "being unable" has now been elided and with it the insight, preserved by the dyadic structure, of unfreedom as a function of the normative relation between the will of one person and that of another.

MacCallum's position is motivated by Berlin's failure to appreciate that circumstances and forces beyond particular agents' control can be sources of social unfreedom for others. MacCallum has a specific conception of the virtues of the positive freedom tradition. Although his triadic formula mentions the freedom "to become," his substantive focus is on socio-economic disadvantage and its effects on persons' capacity to act as they (would) choose (if they could). His restoration of the positive tradition appeals more to its socialist and related welfarist credentials than to Stoic or Kantian thoughts about freedom as the will's independence of determination by desire. It is also in this spirit that liberal-egalitarian thinkers such as John Rawls (1971: 201–5) or Ronald Dworkin (2000: 120–3) have incorporated the idea of freedom as a triadic concept into their theories of social justice.

From the perspective of freedom analysis itself the abiding influence of MacCallum's formula, apart from the rigor of its articulation, lies in its push towards descriptivism about freedom. The triadic formula tells us when an agent is free from a specified constraint to perform a specific (or specifiable) action. The idea of freedom as an inner state is barely captured: one might say that agent x is free from inner turmoil to become a self-knower—but this sounds strained. By contrast, saying that x is free from legal restrictions to attend the school of her choice is informative in the intended sense: here the formula tells us what the agent is free to do, given absence of a corresponding social constraint.

3. Pure Negative Freedom

I noted above that an aspect of Berlin's initial outline of negative freedom comes to haunt him, placing negative freedom in close proximity to the repudiated positive conception, and compelling Berlin to revise his initial account. The problem is that of the contented slave. Unlike the dissatisfied slave the contented slave is happy with his condition. While the dissatisfied slave does not wish to be in a state of confinement, the contented slave has adapted her desires to her situation. Though physically confined,

she is inwardly at peace, like the Stoic, with her situation. While a proponent of negative freedom might find the thought of a person's being inwardly free though outwardly constrained counterintuitive, the phenomenon of adaptive preferences is a problem for Berlin's negative freedom conception.

Berlin thinks of freedom as choice-directed action. A person is free when he is unhindered by others to do as he wishes and wills. Yet where others interfere, a person may simply change his mind. He may adapt his wants to whatever others want for him; others then no longer interfere with his doing as he pleases, hence are no longer rendering him unfree. Taking the charge seriously, Berlin's revised account omits any reference to agents' inner mental states:

> The sense of freedom in which I use his term entails not simply the absence of frustration, but the absence of obstacles to possible choices and activities—absence of obstruction on roads along which a man can decide to walk. Such freedom ultimately depends not on whether I wish to walk at all, or how far, but on how many doors are open, how open they are, upon the relative importance in my life.
>
> (Berlin 2002: 32)

Freedom is no longer the capacity to do as one wishes and wills unhindered by others, but is a function of having options available to one.

This shift from mixed—internal and external—freedom conditions to a purely external perspective has been perfected by Hillel Steiner in his account of pure negative freedom. Pure negative freedom contains no reference to inner mental states; it is conceived in wholly agent-external terms. Steiner acknowledges that his approach will conflict with many people's intuitions about freedom. He himself has the strong intuition that "persons are free to do what they actually do" (Steiner 1994: 8). Steiner's freedom conception rests on this single strong intuition.

Persons do what they actually do. But are they always free to do what they do? Intuitively, one might think persons free to do only what they could also not do. The fact that a person does x does not show that they are free to x unless they could, alternatively, not-x. This view invokes some notion of choice; yet this is to appeal to an agent-internal capacity of the kind Steiner is keen to avoid. According to pure negative freedom the fact that a person is free to x need not imply that she is also free to not-x. She might be free to x but not free to not-x; alternatively, she might be free to x and free to not-x. A person who is free to x and free to not-x is freer than one who is free to x but not free to not-x. A person who sits in a rudderless boat is free to drift down the rapids but not free not to do so; a person in a boat equipped with a rudder is free to drift or not-drift. On Steiner's account, we cannot compute the fact that the second person is freer than the first unless we individuate freedoms.

If we want to be able to make comparative freedom judgments about who is more or less free, we must find a way of quantifying freedom. How free a person feels is a qualitative question; Steiner rejects qualitative judgments as subjective. We can determine whether or not a person *is* free—and how free she is compared to others—only by computing the objective difference between the actions she is free to do and those she is not free to do, and similarly computing the actions she is free to do to the actions others are free to do. Steiner accordingly rejects Berlin's mixed conditions and related definition of freedom as non-interference.

Freedom as non-interference makes crucial reference, we saw, to persons' wills. A pure negative freedom conception eliminates reference to willing. Freedom becomes physical unpreventedness. A person is free to do what he is physically unprevented from doing: unprevented, the toddler freely falls down the stairs. Steiner rejects the view of coercion as unfreedom. Coercion works on the will: when another threatens you with sanctions unless you do as she wants you to do, she is trying to influence your will formation. On the physicalist account, your freedom remains unperturbed: at the point of the threat being issued, no physical prevention occurs. You are free either to comply or not to comply. If you comply, you are free to do as the coercer tells you. If you do not comply you are free to suffer the relevant sanction. It is also true that at the point at which the coercer becomes the preventer, she renders you unfree. If you decide neither to comply nor to suffer the sanction, the coercer who successfully imposes (physical) sanctions renders you unfree not to suffer the sanction.

One might conclude that under freedom as physical unpreventedness, unfreedom simply kicks in a little later than it does on the coercion-based account: while the latter diagnoses unfreedom at the point of issuance of the threat, the former diagnoses it at the point of physical sanctioning. Physicalists claim greater analytic accuracy: to say that a person's freedom is restricted at the point of issuance of a threat is a confused way of saying that a person's freedom is restricted at the point of the threat's being successfully carried out. This still leaves us with the case of the compliant person who acts unfreely because contrary to her preferred wishes or intentions. It is not clear that the physicalist has a way of computing the non-voluntariness of the compliant person's action: for the physicalist the compliant person is free to do what he is coerced into doing.

The basic concern of theorists of pure negative freedom is to ask, "How free am I compared to you?" This is continuous with Berlin's question, "how much freedom?" However, in contrast to Berlin, pure freedom theorists do not assume that freedom is a capacity of the will. Instead, freedom is a distributive good: to be free is to control the physical components of a proposed action. To be unfree is to have that control taken away from one by another. In that sense, the amount of freedom there is in the world is fixed: there is just as much freedom in the world as there are possible things to do. This amount can be variously distributed across persons. How much freedom I have is a function of how much freedom you have. Since freedom is control of the physical components of eligible actions, to be free is to have property. Any property distribution is at the same time a freedom distribution, and there can be more or less equal systems of freedom as property rights.

Pure freedom theorists are interested in equal freedom distribution: hence the need to devise a method for individuating—and counting—free actions and hence Steiner's physicalist approach which aims to render freedoms countable and computable. Yet even with a physicalist account of freedom, computing the extent of persons' freedom remains a fiendishly difficult task (Carter 1999: 169–287). Until pure negative theorists succeed in first individuating and then counting all of the actions all persons do (and could do) their aim of telling us exactly how much freedom each of us has will elude them.

4. Freedom as Non-Domination

The most recent development in the negative freedom debate is the republican attack on pure negative freedom. Republicans react against what they perceive to be the reduc-

tivism of Berlinian and pure negative freedom theories alike. Berlin's definition of negative freedom as non-interference is viewed by them as a prelude to the pure negative account, which republicans regard as an unadorned return to Hobbesian physicalism (Skinner 1998). For Hobbes, freedom is physical unpreventedness: a stone rolling down the hill is as free as a person walking down the road. Berlin restricts freedom to relations between persons; his approach is unHobbesian in that he predicates freedom only of beings with a capacity for will (as choice), resulting in his troublesome mixed account. Although pure negative theorists disavow appeal to mental capacities, they likewise restrict themselves to computing the freedom amounts of persons. The grounds on which pure negativists disqualify stones are unclear; they may simply plead abstraction from freedom issues to do with "mental states." Were we to ask how free persons feel, inner capacities might well be relevant—it just so happens that pure negative theorists are not interested in that question. They are interested in political freedom and according to them, political freedom is a distributive good in relation to which we can abstract from subjective freedom evaluations (but see Kramer 2005).

Republicans contest the assumption that the relation between freedom and inner capacities is politically irrelevant. Here republicans are in *de facto* agreement with Berlin: political freedom specifies a normative relation between the will of one and that of another. However, republicans disagree that persons are antecedently free; they make contact with the positive tradition in viewing freedom as a kind of political achievement. That achievement is of mutual non-domination (Pettit 1997, 2001).

If Berlin's initial sketch of negative freedom privileges the external freedom condition (non-interference) over the internal freedom condition (capacity for choice), and if pure negative freedom theorists reduce freedom conditions to external non-prevention, republicans retrieve the centrality of an *interpersonal* freedom condition which they articulate in normative but not strongly evaluative terms. Republicans are unconcerned with the *goodness* of individuals' choices; indeed, in a sense they are unconcerned about persons' choices. Although freedom pertains to the will, the will-related capacity in question is less that of individual choice and more the moral capacity to acknowledge others as moral and political equals.

This account borrows from the Roman tradition of the free man as citizen, not slave. To be free is not to dominate others and not to be dominated by them. Freedom as non-domination differs from freedom as non-interference. The latter construes interferences as episodic disturbances of a person's natural freedom state: others periodically interfere with one's capacity for choice and action. By contrast, domination is a socially instituted relation between two or more persons. Dominated persons are treated by others as incapable in principle of exercising their will. Standard examples of dominated persons include slaves, women, colonized peoples. A crucial feature of domination is others' denial of their status as equals (Pettit 1997: 51–79).

Advocates of republican freedom tend to focus on unfreedom as domination rather than on freedom as non-domination. The free (non-dominated) person is portrayed as exercising power of choice in relation to his affairs; he is to that extent difficult to tell apart from his liberal, negatively free neighbor. It seems reasonably clear that the republican's freedom should be theorized as consisting not in exercising choice for himself but in not exercising choice on behalf of others. On the republican account, we should regard that person as free who is not only not dominated by others, but who also refrains from dominating others. Freedom is not essentially about having an adequate range of options. A person might have very few options available to herself: so long as she refuses

to dominate others and is not dominated by them, she is freer than one who has many options from which to choose and who dominates his wife, his children, and his employees. It is an implication of the republican focus on freedom as non-domination that the domineering person is no more free than the people he dominates.

This is not an intrinsically incoherent freedom conception. Its most famous erstwhile proponent is Hegel (1977: §179), who offers a detailed account of the master–slave dialectic along these lines. Current advocates of republican freedom usually think this takes matters a step too far—most take the view that the dominated are less free than those who dominate them, not that both parties are locked together in a relation of unfreedom. In shrinking back from what would seem to be an implication of their analysis of freedom as non-domination, republicans open themselves up to the charge of being "closet liberals"—freedom theorists who for all their disavowal of Hobbesian reductivism cling fast, in the end, to the quintessentially liberal thought of freedom as being left unhindered to do as one pleases, or more "purely," freedom as having a range of options available to one. Domination may also be a freedom consideration, but if it is, it is an addendum to liberal freedom thinking, not an alternative to it (Kramer 2005: 125–43).

5. Conclusion

Berlin's seminal "Two Concepts" has had an abiding impact of liberal theorizing about freedom—an impact that is partly positive and partly negative. Positively, Berlin's essay has encouraged an analytical focus on the general concept in abstraction from the divergent philosophical views on freedom offered by various thinkers in the liberal canon. Negatively, his ideological stance towards the positive tradition has helped breed a climate of philosophical suspicion and intolerance towards that tradition. MacCallum's intervention—intended to re-affirm the unity of the concept—pushed the debate even more vigorously towards negative freedom; MacCallum's narrow analytic method failed to accommodate the positive tradition's interest in the metaphysics of freedom. The suspicion of what one could call the ineffable aspects of the idea of freedom—aspects that do not lend themselves to descriptive specification—is notable in pure negative theorizing. Yet even restorers of the republican tradition remain fearful of association with any but the broadly negative freedom tradition, as can be seen from their refusal to engage with the thoughts of Rousseau, Nietzsche, Foucault (Flathman 2003)—historical critics of domination. That the exclusive focus on negative freedom is unsatisfactory even within the framework of mainstream liberal thought is evident from the fact that alternative conceptual space has had to be found for much that was thematized by the positive tradition: notions of self-development, of freedom as independence from subjective inclination, and of freedom as reasoned (moral) judgment are now explored under the concept of autonomy. The high degree of division of labor in current analytic philosophy makes this possible; it makes it possible to avoid questions concerning the relation between personal autonomy and pure negative freedom, for example. Whether this assists or hinders our philosophical understanding of a central liberal idea is a moot point.

Related Topics

Hobbes, Rousseau, Kant, Hegel, Mill, Liberalism, Republicanism, Left Libertarianism, Autonomy, Power, Authority and Legitimacy

References

Berlin, I. (2002) *Liberty* (ed. H. Hardy), Oxford: Oxford University Press.
Carter, I. (1999) *A Measure of Freedom*, Oxford: Oxford University Press.
Christman, J. ed. (1989) *The Inner Citadel. Essays on Individual Autonomy*, Oxford: Oxford University Press.
Cohen, G.A. (1995) *Self-Ownership, Freedom, and Equality*, Cambridge: Cambridge University Press.
Dworkin, R. (2000) *Sovereign Virtue*, Cambridge, MA: Harvard University Press.
Flathman, R. (2003) *Freedom and its Conditions: Discipline, Autonomy, and Resistance*, London: Routledge Publishers.
Gray, J. (1989) *Liberalisms*, London: Routledge Publishers.
Hegel, G.W.F. (1977) *The Phenomenology of Spirit*, trans. A.W. Miller, Oxford: Clarendon Press.
Kramer, M. (2005) *The Quality of Freedom*, Oxford: Oxford University Press.
MacCallum, G. (1967) "Positive and Negative Freedom," *The Philosophical Review* 76: 312–34.
Pettit, P. (1997) *Republicanism: A Theory of Freedom and Government*, Oxford: Oxford University Press.
Pettit, P. (2001) *A Theory of Freedom*, Cambridge: Polity Press.
Rawls, J. (1971) *A Theory of Justice*, Oxford: Oxford University Press.
Skinner, Q. (1998) *Liberty before Liberalism*, Cambridge: Cambridge University Press.
Steiner, H. (1994) *An Essay on Rights*, Oxford: Blackwell Publishers.
Strawson, P.F. (1974) *Freedom and Resentment and Other Essays*, London: Methuen Press.
Taylor, C. (1985) *Philosophy and the Human Sciences*, Vol. 2, Cambridge: Cambridge University Press.

50
AUTONOMY

Horacio Spector

1. Introduction

Ever since Ancient Greece the ideal of leading an autonomous life, a life of one's own choosing and making, occupied a paramount position in the Western constellation of values. It was passed down, in the form of free will, from Greek philosophy to Christian thought, and it was put to political use in Renaissance humanism. *Political autonomy* was cherished by republican and liberal writers alike yet it was Rousseau who eventually crafted the contemporary notion. While Rousseau portrayed primitive man as a free creature in *A Discourse on the Origin of Inequality*, in *The Social Contract* he argued that primitive man was enslaved by his appetites and instincts and that civil liberty was the true kind of freedom:

> We might, over and above all this, add, to what man acquires in the civil state, moral liberty, which alone makes him truly master of himself; for the mere impulse of appetite is slavery, while *obedience to a law which we prescribe to ourselves is liberty*.
>
> (Rousseau 1993: 196; italics added)

Kant would borrow this notion to formulate his conception of *moral autonomy* as obedience to our common practical reason. Whereas Rousseau was concerned with political autonomy and Kant with moral autonomy, John Stuart Mill shifted to *personal autonomy*. Mill contrasted men with "automatons." The mark of men is to have a character:

> A person whose desires and impulses are his own—are the expression of his own nature, as it has been developed and modified by his own culture—is said to have a character. One whose desires and impulses are not his own, has no character, no more than a steam-engine has a character.
>
> (Mill 1962: 189)

Personal autonomy, political autonomy and moral autonomy are distinct and potentially conflictive ideals. For instance, a radical commitment to political autonomy might compromise personal autonomy. Similarly, conceptions of moral autonomy might affect personal autonomy if those conceptions rely on repressive moral codes.

Because political autonomy refers to collective self-government, it is generally studied under the heading of democracy, as democracy is the best known form of

self-government. On the other hand, moral autonomy is studied in moral philosophy. For these reasons, this chapter focuses on personal autonomy, which is a fundamental issue in liberal political philosophy. Basically, "the ruling idea behind the ideal of personal autonomy is that people should make their own lives" (Raz 1986: 369). Making a life implies having the power to actualize various possibilities. This imposes some constraints on the kind of creature that is capable of being autonomous. Basically, an autonomous person must be a conscientiously diachronic being who defines her own existence out of a compass of options by exercising capacities of judgment, evaluation and decision. Martin Heidegger used the concept of "Dasein" (i.e., the historically situated self) to discuss the essential attributes of an autonomous being. The basic state of the "Dasein" is one of being-in-the-world; it exists in a natural and historically given world. Its mode of being-in-the-world is not merely theoretical, but mainly practical. The "Dasein" cares about its own existential projection in the world as a temporally extended creature that is thrown upon its various possibilities. The autonomous person must have the capacity to project his life out of the various possibilities open to him, as well as the responsibility for doing so (Heidegger 1962). Heidegger's discussion brings out three main conditions for autonomous choice: (1) epistemic and practical capacities, (2) a range of options or possibilities, and (3) actual interaction with the natural and social world.

As regards the first condition, it is accepted that the exercise of personal autonomy requires the possession of rationality competencies. Stanley Benn has analyzed the *subjective conditions* of autonomy. He holds that any normal human being must possess these practical and cognitive competencies to a minimal degree, and he calls "autarchy" this set of normal rational capacities. An autarchic person has the ability to develop an ordered set of preferences (which, in turn, affects her cost-benefit analysis) and to form and revise beliefs on the grounds of available evidence (which, in turn, affects her practical deliberation) (Benn 1975–76: 111–12; 1988: 153–4). Among the defects of practical and epistemic rationality, Benn mentions paranoia, kleptomania, schizophrenia and other psychopathic conditions. Benn says that, whereas autarchy is "a condition of human normality," autonomy refers to a certain excellence of character (Benn 1975–76: 113; 1988: 155).

The availability of a plurality of options is often considered a condition of personal autonomy (Raz 1986: 373–7). Benn emphasizes condition (2) in holding that autonomous choice involves *objective conditions* such as a range of powers and capacities, a set of resources, a range of possible impediments and a set of opportunity costs (Benn 1975–76: 111; 1988: 152). If we regard human behavior as a function of two variables, the agent's nature and his environment, the following relation holds: the wider the range of options available, the more fully his decisions and actions express his own nature rather than the environment he happens to live in. But this relation is only a pro tanto one and cannot be brought to the limit because radically independent self-creation is incoherent (Berofsky 1995: 112–20). While a society that allows a rich repertoire of options certainly favors personal autonomy (Spector 1992: 97–8), no society is viable or even conceivable without a diversity of constraints on human choice. It seems clear that assessments of autonomy cannot be purely quantitative but must presuppose normative criteria. One says that an agent's act is nonautonomous if he would not have chosen it in the face of an *adequately larger* compass of options, where "adequately" refers to criteria picking out options that people should have for various reasons. For instance, Julien Sorel—the son of a carpenter in Stendhal's *The Red and the Black*—chose a religious career, but his choice was arguably nonautonomous because in the times of the

Restoration someone of his class unjustifiably lacked the option to become a successful military (Stendhal 2011).

Robert Nozick stresses condition (3). Though not writing explicitly on autonomy, Nozick offers an illuminating example to show that an autonomous life implies not only a repertoire of possibilities but also actual involvement with the external world. Nozick invites us to imagine an experience machine that could give us any experience we desired:

> Superduper neuropsychologists could stimulate your brain so that you would think and feel you were writing a great novel, or making a friend, or reading an interesting book. All the time you would be floating in a tank, with electrodes attached to your brain. Should you plug into this machine for life, preprogramming your life's experiences?
>
> (Nozick 1974: 42)

Nozick believes that "plugging into the machine is a kind of suicide." Why? Our actual life in the world matters to us more than experiences from the inside. We want to lead our lives, to do certain things, to be certain ways.

But when are our actions and projects truly ours? When does the life of a person spring from her own choice and creation? It is generally agreed that an agent does not act autonomously merely when his behavior is caused by his desires and preferences. Some desires and preferences may be nonautonomous in that they are alien to the agent's personhood, even if they exert motivational force on his behavior. Abraham wanted to sacrifice Isaac when God commanded him to do so, but this was not an autonomous choice. Othello wanted to strangle Desdemona out of pathological jealousy, but he passionately loved her.

Autonomy is widely associated with the *Real Self condition* and the *Independence condition* (Wolf 1993; Arneson 1994: 54). The Real Self condition imposes a degree of psychic order by awarding practical authority to some agency within the person. Various theories interpret this condition by supposing that the self constitutes itself as an "I" that rules the "me." The Independence condition says that an autonomous person is not ruled by anyone else. We will address four "pure" conceptions of autonomy: while the first two are related to the Real Self condition, the third gives an account of the Independence condition. The fourth conception narrows the separation between the self and the external social world. "Mixed" views combine two or more of these conceptions.

2. The Rationalist Conception

One classic conception of autonomy, which goes back at least to Plato and Aristotle, regards "self-direction as an expression of the *rational* nature of persons" (Young 1986: 10). For Aristotle the self is divided up into two main parts, desire and reason, which are not on a par. Man's rational nature implies that reason should govern the passions (Aristotle 1999: 1102a; 1102b; 1139a). By exerting practical reason, an agent knows what is good for her and rationally controls her desires. Though for Aristotle human goodness itself is not determined by the self, but is rather a necessary outcome of the self's inclination to fulfill its natural function, this conception of practical reason certainly encapsulates a basic notion of autonomy.

Contemporary rationalist ideals of personal autonomy were influenced by Kant's conception of moral autonomy, which emphasizes that rational behavior is governed by non-instrumental, categorical reasons. For instance, Gary Watson proposes a version of the rationalist conception by differentiating between valuing a state of affairs and desiring that it obtains, and claims that reason is an original spring of action. Yet Watson does not think that there is a radical separation between the appetites and the passions, on the one hand, and value judgments, on the other. In effect, he holds that some activities are valued because they are the object of certain appetites. For instance, most people value sexual activities because they are the object of erotic impulses. The difference between desiring and valuing is not related to the content of desires and wants but to their sources and roles in the agent's *motivational* and *valuational* systems. Indeed, autonomous agency involves a motivational system and a valuational system. Whereas a motivational system is a set of considerations that move people to action in a direct way, a valuational system is a set of considerations that yields normative judgments when combined in a reasoned way with factual beliefs. An agent's valuational system includes those principles and ends that he takes "in a cool and non-self-deceptive moment" to be "definitive of the good, fulfilling, and defensible life" (Watson 1989: 116). The valuational system constitutes the agent's standpoint or viewpoint, one that he cannot entirely dissociate from. Free or autonomous action is then action that the agent "really or most wants," that is, action that satisfies both the agent's motivational and valuational systems.

In a number of articles gathered in his *Structures of Agency*, Michael E. Bratman has proposed a variant of the rationalist conception based on the ideas of agency and self-governing policies (Bratman 2007). Basically, Bratman's analysis is an elaboration on the subjective conditions of autonomy (in Benn's sense). Bratman holds that autonomous agents have the capacity of valuing separable forms of their own motivations and of weighing practical considerations in cases of value conflict. Agents exercise this capacity by developing non-instrumental higher-order self-governing policies, with which they are satisfied, that establish which desires to treat as justifying considerations in their practical reasonings. Now these higher-order policies solve two problems at the same time: the problem of "reflective self-management" and the problem of the "under-determination by value judgments." The former is the problem of reflecting about our first-order desires and emotions. The latter is the problem of how to create a personal practical viewpoint within the bounds of socially convergent practical reasoning.

For Bratman autonomy involves *agential direction* and *agential governance*. Agential direction requires direction of one's own conduct by a sufficiently organized motivational system. Agential governance is a particular form of agential direction: it involves the treatment of certain considerations as justifying reasons for action (i.e., as "end-setting" for practical reasoning). Agential governance functions through self-governing policies that exclude certain forms of practical reasoning. They must be higher-order if the exclusion of certain desires as reasons for action is both transparent and reflexive. The *autonomy-transparency thesis* and the *reflexivity thesis* require that the agent know that her psychological functioning is due to higher-order policies that refer to the first-order desires that are practically excluded. Transparency means that the agent knows how her practical reasoning is structured. Reflexivity means that the policies are in part about their own function of practical guidance. Thus Bratman's self-governing policies are essential for autonomy because autonomous choice is essentially guided by forms of practical reasoning that are allowed or banned by those policies in a transparent and reflexive way.

Liberals often rely on the rationalist conception. For instance, Gaus defends a form of ultra-minimal personal autonomy, according to which an autonomous agent is someone who sees his actions as determined by his own deliberations (Gaus 2005).

3. The Structuralist/Hierarchical Conception

The structuralist conception, sometimes also called "hierarchical," maintains that the self splits itself into two or more "orders" or "tiers," the highest of which endorses one class of preferences and choices as the person's own. This conception emerged as a contemporary effort to craft a notion of free will compatible with a Humean model of explanation of action that basically relies on desires and preferences. Because this explanation revolves around passions and desires and leaves out practical reason as a self-sufficient motivating principle, self-governance must be defined in terms of structural properties of preferences. Structuralism says that autonomous preferences must hold a certain position within a hierarchical structure of motivations and preferences.

In the early 1970s Gerald Dworkin and Harry Frankfurt sought to explain various forms of unfree action against the background of a Humean view of motivation (Dworkin 1970; Frankfurt 1988: 11–25). Can we explain free will without resorting to practical reason? How can we split the acting self when the springs of action are circumscribed to desires and preferences? For solving this problem Dworkin and Frankfurt posited a hierarchy of preferences and desires. But the basic idea was not totally new: the notion of second-order desires in the broad sense of "desires about desires" had been used by Francis Hutcheson (Hutcheson 2002: 25). Though behavior is generally moved by first-order desires, free and autonomous choice results from practically efficacious first-order desires that we want to be moved by. By contrast, nonautonomous choice derives from first-order desires that we disapprove of. There are significant differences between Dworkin's and Frankfurt's articulations of this theory.

Dworkin predicates that a person does *X* freely if and only if he does *X* for reasons which he does not mind acting from; conversely, a person does not act freely in doing *X* when he does *X* for reasons he resents having to act for (Dworkin 1970). Dworkin's starting-point is a rejection of the claim that one always acts in accordance with one's own desires. In cases of coercion, says Dworkin, the coercer provides the victim with motives in the sense of supplying conditions in which the victim becomes motivated to act in the way the coercer wants. However, as Thalberg noted, it is not correct to say that an agent that acts under coercion (e.g., the victim of a highwayman) resents having to act from the prudential motives he acts from (Thalberg 1989: 126–7).

In later work Dworkin proposed the *authenticity condition* of autonomy: for a person to be autonomous his second-order identifications must be congruent with his first-order motivations (Dworkin 1976). This means that the autonomous person is identified with the desires moving his conduct, that he sees himself as the sort of person who wants to be moved by these desires. Yet Dworkin abandoned the authenticity condition in *The Theory and Practice of Autonomy* by contending that it is not identification that is really crucial for autonomy but, rather, the capacity of critical self-reflection. He supported his contention with four considerations. First, whereas autonomy is a property of extended portions of a person's life, identification is an episodic state of affairs. Second, certain manipulations of a person, such as keeping him ignorant or lobotomizing him, do not interfere with his identifications but with his capacity to make or reject such identifications. Third, a person might succeed in identifying himself with motivations that he

does not wish to have by changing his second-order desires. Fourth, a drug addict might want to be motivated by the addiction even though he may not be capable of modifying this first-order motivation. Accordingly, Dworkin conceived autonomy as "a second-order capacity of persons to reflect critically upon their first-order preferences, desires, wishes, and so forth and the capacity to accept or attempt to change these in light of higher-order preferences and values" (Dworkin 1988: 20). In terms of Young's distinction between an *occurrent* sense of autonomy (autonomy in particular situations) and a *dispositional* sense (autonomy in the course of a person's life), Dworkin defends the latter (Young 1986: 8).

Frankfurt holds in "Freedom of the Will and the Concept of a Person" that an essential quality of persons is their capacity to form second-order volitions. Second-order volitions are types of second-order desires, but only in a "truncated way." Thus an agent might want to want to X in a sense in which this does not entail that he wants to X; he just wants to have the desire. But in a stronger sense, an agent wants to want to X if he wants the desire to X to be *effective*. "[W]hen the statement that A wants to want to X is used in this way, it does entail that A already has a desire to X" (Frankfurt 1988: 15). Second-order desires can then be either second-order desires in a "truncated" sense or second-order volitions. Thus Frankfurt says: "Someone has a desire of the second order either when he wants simply to have a certain desire or when he wants a certain desire to be his will" (Frankfurt 1988: 16). Therefore, it is not the mere possession of second-order desires to have desires of the first order that characterizes a person: a person is a creature that has the capacity for second-order *volitions*, that is, the capacity to want a certain desire to become her own will. Creatures that can form second-order desires but not second-order volitions Frankfurt calls "wantons": "[t]he essential characteristic of a wanton is that he does not care about his will" (Frankfurt 1988: 16). For instance, an unwilling addict who has conflicting first-order desires is a person if he can have second-order volitions on the basis of which he can disown his first-order desire to take drugs and instead identify with his first-order desire to abstain from taking drugs. For Frankfurt, "[a] person's will is free only if he is free to have the will he wants" (Frankfurt 1988: 24).

In accordance with the structuralist/hierarchical view, one can define an autonomous desire of an order n as a desire that a person wants to be his will on the basis of a volition of order $n + 1$. For instance, an unwilling addict's first-order desire to take drugs is not really his own will, and he cannot identify with it, because he has no second-order volition that can constitute this desire as his own will. However, a fatal dilemma arises. Because higher-order volitions are also higher-order desires, whose autonomous character can in turn be contested, either the definition leads to an infinite regress, and then to the impossibility of autonomous desires, or the regress is cut off at one arbitrary order n by taking an n-order volition as the first "identifier" or "authenticator." John Christman calls the puzzle raised by the second horn the "*ab initio* problem" (Christman 1989: 9).

Frankfurt tries to avoid the first horn of the dilemma without getting trapped in the second. Thus he argues that it is possible to terminate the regress without cutting it off arbitrarily: "When a person identifies himself *decisively* with one of his first-order desires, this commitment 'resounds' throughout the potentially endless array of higher orders" (Frankfurt 1988: 21). That is, if a person takes a decisive commitment with some n-order desires at order $n + 1$, there is no need to continue the ascent to higher orders. Gary Watson nevertheless adds a dilemma to this purported solution. Either terminating the ascent at

$n + 1$ is, indeed, arbitrary because decisiveness of commitment does not guarantee lack of manipulation, or the notion of higher-order volition is not the fundamental one (Watson 1989: 119). In fact, if nonautonomous higher-order desires are empirically possible, then cutting off the ascent at one order is arbitrary. As Jon Elster observes:

> There is no reason […] to believe that second-order desires are always immune to irrelevant causal influences. If they were, the regress would be cut short, but as shown in the important work of George Ainslie, second-order desires may take on a compulsive character and become as heteronomous as the impulsive first-order desires they were supposed to protect us from.
>
> (Elster 1983: 21)

On the other hand, if stopping the regress is warranted by decisive commitment and identification, it is the latter notions that make the job, not higher-order volitions.

In "Identification and Wholeheartedness" and "The Faintest Passion," Frankfurt tries to escape from the first dilemma by introducing the ideas of "wholeheartedness" and "satisfaction." In the former article, he says that "ambivalence," a kind of psychological instability that affects second-order volitions, is resolved when the agent identifies wholeheartedly with one of the volitions, and consequently incorporates it into his constitution, excluding any competitive volition "as an outlaw" (Frankfurt 1988: 171). In the latter article, Frankfurt introduces the concept of satisfaction within his definition of autonomous desires:

> Hierarchical accounts of the identity of the self do not presume, however, that a person's identification with some desire consists simply in the fact that he has a higher-order desire by which the first desire is endorsed. The endorsing higher-order desire must be, in addition, a desire with which the person is *satisfied*.
>
> (Frankfurt 1999: 105; italics in the original)

Unlike endorsement or authentication, which involves mental action, satisfaction is rather a passive condition, that is, a non-occurrence:

> What satisfaction does entail is an absence of restlessness or resistance. […] There is nothing that he needs to think, or to adopt, or to accept; it is not necessary for him to do anything at all. This is important, because it explains why there is no danger here of a problematic regress.
>
> (Frankfurt 1999: 103–4)

Yet it is doubtful that appealing to wholeheartedness or satisfaction can appease the worries of those who press the infinite regress critique. Just as preferences holding at any given order n can result from psychic manipulation, indoctrination, and subliminal propaganda, so too an agent's *absence* of uneasiness about his will, that is, a conformist attitude toward the constitution of her personhood, can be achieved through heteronomous mechanisms of will formation. If, following Elster, we use "conformity" "as a technical term for a desire caused by a drive to be like other people," conformist wholeheartedness is caused by other people's behavior (Elster 1983: 23). This is not an acceptable method for autonomous formation of preferences. Suppose Jim wants to make a career in literature

and so pursues a degree in a demanding college, but his motivations are not stable because he also wishes for a high standard of living. His will is not totally formed. Now Jim's friends are succeeding in business and other profitable professions while he has to live on his parents' financial support. Just before getting his bachelor's degree, Jim takes the wholehearted decision to take a master's degree in publicity. Is his change of mind autonomous? All depends. Jim might have discovered that he does not really want to work in a literature department all his life, and this might have set the balance of motivations in favor of a publicity career. In this scenario Jim's new plan is autonomous. But the behavior of Jim's friends might have entered into the formation of his new motivational equilibrium. In the latter case, Jim attains a wholehearted will in a nonautonomous way. Just substituting "satisfaction" as a negative motivational state of affairs for a higher authenticating level of positive preferences does not render the break of the regress non-arbitrary. Stopping at the *absence* of uneasiness or conflict in the will is just as dogmatic as stopping at the *presence* of a higher-order preference. The dilemma seems inescapable.

Furthermore, the dilemma might really be a *trilemma*. If authentication or validation is not an *asymmetric* relation, the possibility of circularity might give rise to a third horn. In effect, the hierarchical view does not contemplate "circles" of preferences of different order because it assumes that the relation of endorsement or authentication that structures the hierarchy is asymmetric. On this assumption, a lower-order desire cannot endorse or authenticate a higher-order desire. But endorsement and authentication might be a two-way process. Let us see why. Endorsement and authentication are *intentional mental acts*. In Franz Brentano's sense, intentionality means that mental acts are about something, or directed toward something. Therefore, intentional acts have two correlative components: the act itself and its intentional object, which Edmund Husserl respectively called *noesis* and *noema* (Husserl 1998: 211–35). Since desires are intentional acts, they also exhibit intentionality, that is, they are directed toward a subject matter (i.e., the desired object). Now the intentional objects of reflective second-order desires are first-order desires. With this conceptual apparatus in place we can see that the hierarchical theory dogmatically assumes that for an intentional act I_2 to authenticate or validate another intentional act I_1, the latter must constitute by logical necessity the intentional object of the former. Since the theory assumes that it is impossible for a higher-order desire to be the intentional object of a lower-order desire, it rules out "circles." But it does not seem impossible for the intentional content of I_1 to validate I_2. For instance, if an agent desires that his society were a just society and, furthermore, desires that this desire constitute his will, it is not self-contradictory to say that the content of the first-order desire (i.e., his society's being just) validates the corresponding second-order desire.

The structuralist conception cannot dispense with considerations relative to the ways in which various psychological structures have arisen. Frankfurt acknowledges that a person is not a free agent if she is continually manipulated by an imaginary Devil/neurologist. But he thinks that if the same Devil/neurologist programs a person in a stable way, so that she can identify herself with some of her desires, she can be a free agent (Frankfurt 1988: 47–57). Yet it seems controversial that lack of autonomy can only be produced by alien and/or unstable interventions. Suppose a research-biomedical company discovers drug WH, an anxiolytic which, if regularly administered, induces stable wholeheartedness in patients affected by volitional ambivalence. Even if a person wanted to be administered WH, we would hardly say that the motivational structure caused by WH is an autonomous one. Just as WH, there are various causal mechanisms

that are not compatible with a person's seeing herself as the result of the autonomous development of her will in normal conditions of interaction with the environment. Broadly speaking, autonomy requires a *condition of normality:* an autonomous person's character must be developed under conditions of normal interaction with her surroundings (Spector 1992: 94).

4. The Proceduralist/Historical Conception

The proceduralist/historical conception relies on the relation of a person's preferences and values to the procedures that have caused them, or to the identity of the person as historically shaped throughout her life. Some first-order and higher-order preferences are nonautonomous because of the causal mechanisms that have formed them. Dworkin introduced the notion of "procedural independence" to dub the manifold ways of nonautonomous preference formation (e.g., hypnotic suggestion, manipulation, coercive persuasion, subliminal influence, etc.) (Dworkin 1988: 18–19). I examine three kinds of nonautonomous procedures of preference formation, but this list is not an exhaustive one.

Repressive rearing is our first kind of nonautonomous procedure. Autonomous human beings are the outcome of liberal education. A child cannot develop an autonomous character without living and being reared in an open environment that fosters the development of her innate propensities (Arneson 1994: 59–60). Naturally, those propensities should be blocked that are conflictive, antisocial or inimical to productivity, wellbeing and responsibility for leading a worthwhile life. However, repression of inclinations and talents that is not justified as a way of promoting the wellbeing and autonomy of the child is incompatible with the rearing of autonomous adults.

The second kind is coercion. This is a form of heteronomous control that occurs when a coercer reshapes in a compulsory way an agent's set of options in such a way that (i) one or more options are removed, and (ii) the reshaping is intentionally addressed to moving the agent to do or refrain from doing a certain action. It is debatable whether the concept of coercion is normatively laden (Spector 1992: 14–22). Though neutral concepts of coercion are possible, it seems indisputable that some normative criteria must be appealed to in order to pick out forms of option reshaping that are violative of autonomy. Both threats and coercive offers are unacceptable forms of compulsory reshaping of the coerced person's range of options. But, for instance, the parent who gives her teenage son less money as a way of encouraging him to search for a job does not coerce him.

Mill maintained that a society concerned with the development of human beings should cultivate their individuality and spontaneity: "To give any fair play to the nature of each, it is essential that different persons should be allowed to lead different lives" (Mill 1962: 192). On the contrary, personal preferences imposed by the community cannot be considered autonomous. Raz says that autonomy needs an autonomy-supporting social environment (Raz 1986: 390–5). For instance, in the film *Little Jerusalem* (2005) Laura, a philosophy student living with her Orthodox Sephardic family in a Parisian suburb, falls in love with Djamel, a Muslim co-worker. Laura converts to Kant's moral philosophy and ironically learns soon after that Djamel decides to break up with her for fear of his community's sanctions. Given Djamel's social mileu, his decision was nonautonomous.

Finally, Elster has studied the phenomenon of adaptive preference formation, that is, the formation of preferences as an adaptive reaction to a narrow range of options.

Elster argues that adaptive preferences are nonautonomous because adaptation is, in this context, a causal mechanism that takes place in an unconscious manner ("behind my back") (Elster 1983: 25, 174). His famous example is taken from the fable "The Fox and the Grapes." Because the grapes are inaccessible, the fox "sees" sour grapes and does not want them (Elster 1983: 22–5; Elster 1996). Arneson has argued that a person would be released from his adaptive preferences if she were to deliberate about them in a fully rational way. He contends that rational autonomy has instrumental value in maximizing rational preference satisfaction because the release from adaptive preferences does not lower the satisfaction of rational preferences (Arneson 1994: 62–7).

Autonomous preferences have a direct proportional relationship to the number of available or feasible options (the "feasible set of options") because the larger the feasible set, the smaller is the likelihood that preferences are formed adaptively. Conversely, if a society provides a small feasible set, it is likely that people's preferences are formed by adaptive mechanisms (Elster 1983: 130–1). (This relationship is empirical and independent from the pro tanto relation between number of available options and degree of expression of one's nature, mentioned earlier).

However, some authors challenge the view that abundance of options is valuable. Gerald Dworkin contends that it is not always better to have more choice than less. Dworkin points out that more often than not having choice implies increased costs, as, for example, the costs of looking for the relevant information, of being exposed to social and legal penalties, the symbolic costs resulting from the adding of new options, and the welfare costs imposed by unregulated prisoner's dilemmas and unresolved forms of weakness of the will (Dworkin 1988: 62–81). Frankfurt also points out that offering people "a beguiling variety of alternatives" makes no good if this enhances vacillation and ambivalence (Frankfurt 1999: 102).

Is there a unified theory of procedural/historical independence? Certainly, one could not define the proper processes as those that autonomous persons would accept or not reject without running into a vicious circularity. Elster understands autonomy in a negative way, as the absence of distorting mechanisms of preference formation. On this view, autonomy is a residual category (Elster 1983: 24). But Alfred Mele and John Christman have suggested positive *historical* analyses that establish the conditions for autonomous processes. These conditions may be subjectivist or non-subjectivist, according to whether they essentially include or exclude reference to the agent's own assessment of her processes of character formation (Christman 2009: 138). Mele's proposal is basically non-subjectivist: an autonomous agent's preferences and values over an interval *t* must not have been compelled, or the compulsion must have been arranged by the agent himself. (Note that the latter disjunct introduces a subjectivist element) (Mele 2001: 165–72).) Christman's current view is subjectivist and historical, but not procedural. A person's preferences or values are autonomous if she would not be alienated from them should she undergo a process of critical reflection that is (1) repeated over a variety of circumstances, and (2) sensitive to the historical processes that gave rise to those preferences or values (Christman 2009: 142–56). Now if the concepts of non-compulsion and non-alienation presuppose the notion of autonomy, we are back in a vicious circularity.

5. The Sociorelational Conception

A common critique of the previous conceptions is that selves cannot be conceived in an individualistic fashion. For instance, Christman grounds his historical conception

of autonomy on the social and, especially, narrative constitution of selves (Christman 2009). Feminists avail themselves of the social embeddedness thesis to criticize those liberal views that ignore the submissive status that various social relations and institutions impose on women in different societies (Meyers 2002; Friedman 2003, 2005). So-called "recognitional" accounts argue that skills such as self-respect, self-trust and self-esteem can only be supported externally, that is, by social practices of mutual recognition (Anderson and Honneth 2005).

Oshana defends an objectivist and paternalistic position with respect to autonomy (Oshana 2003, 2006). She examines three case studies in which certain characters (a voluntary slave, a homemaker and a Taliban woman) have psychological authenticity and procedural independence but lack nonetheless *de jur*e and *de facto* powers to control their lives. She maintains that counterfactual powers are essential for autonomy, that is, the powers to do certain actions if the agent wanted to. Her most controversial case study is that of a woman who is a self-supporting successful professional but is unhappy because she would like to surrender to a loving man. Oshana makes her objectivist paternalism explicit in asserting that, if the woman were to surrender, she would not really be autonomous (Oshana 2006: 64–8).

At this juncture, the frontier between autonomy and freedom gets blurred. In fact, in focusing on the subjection and vulnerability of women the sociorelational view comes close to the republican conception of freedom as non-domination (Oshana 2006: 91–3; Pettit 1997, 2001). Domination often operates not (mainly) through *actual* external constraints but via *potential* external constraints and their induced internal constraints, such as fear. If a gang of youths bullies a girl student without openly threatening or coercing her, it nonetheless can dominate her by creating a scary environment in which independent behavior is ruled out. Submissive fear is an internal constraint that creates personal dependence without necessarily implying a dominator's external action (Friedman 2005; Spector 2010). On this view, preferences formed in the context of dominance relations (e.g., male power) are nonautonomous.

Related Topics

Kant, Mill, Liberalism, Libertarianism, Freedom, Power, Democracy

References

Anderson, J. and Honneth, A. (2005) "Autonomy, Vulnerability, Recognition, and Justice," in J. Christman and J. Anderson (eds.), *Autonomy and the Challenges to Liberalism*, New York: Cambridge University Press, pp. 127–49.

Aristotle (1999) *Nicomachean Ethics*, trans. T. Irwin, Indianapolis/Cambridge: Hackett.

Arneson, R.J. (1994) "Autonomy and Preference Formation," in J.L. Coleman and A. Buchanan (eds.), *In Harm's Way, Essays in Honor of Joel Feinberg*, Cambridge: Cambridge University Press.

Benn, S.I. (1975–76) "Freedom, Autonomy and the Concept of a Person," *Proceedings of the Aristotelian Society*, New Series, Vol. 76.

—— (1988) *A Theory of Freedom*, Cambridge: Cambridge University Press.

Berofsky, B. (1995) *Liberation from Self. A Theory of Personal Autonomy*, Cambridge: Cambridge University Press.

Bratman, M.E. (2007) *Structures of Agency*, Oxford: Oxford University Press, pp. 68–88, 137–61, 162–87, 195–221, 222–56.

Christman, J. (1989) "Introduction," in *The Inner Citadel, Essays on Individual Autonomy*, New York: Oxford University Press, pp. 3–23.

—— (2009) *The Politics of Persons, Individual Autonomy and Socio-Historical Selves*, Cambridge: Cambridge University Press.
Dworkin, G. (1970) "Acting Freely," *Noûs* 4: 367–83.
—— (1976) "Autonomy and Behavior Control," *Hastings Center Report* 6: 23–8.
—— (1988) *The Theory and Practice of Autonomy*, Cambridge: Cambridge University Press.
Elster, J. (1983) *Sour Grapes, Studies in the Subversion of Rationality*, Cambridge: Cambridge University Press and Paris: Editions de la Maison des Sciences de l'Homme.
—— (1996) "Sour Grapes—utilitarianism and the genesis of wants," in A. Sen and B. Williams (eds.), *Utilitarianism and Beyond*, Cambridge: Cambridge University Press and Paris: Editions de la Maison des Sciences de l'Homme, pp. 219–38.
Frankfurt, H. (1988) *The Importance of What We Care About*, Cambridge: Cambridge University Press.
—— (1999) *Necessity, Volition, and Love*, Cambridge: Cambridge University Press.
Friedman, M. (2003) *Autonomy, Gender, Politics*, New York: Oxford University Press.
—— (2005) "Autonomy and Male Dominance," in J. Christman and J. Anderson (eds.), *Autonomy and the Challenges to Liberalism*, New York: Cambridge University Press, pp. 150–73.
Gaus, G. (2005) "The Place of Autonomy within Liberalism," in J. Christman and J. Anderson (eds.), *Autonomy and the Challenges to Liberalism*, New York: Cambridge University Press, pp. 272–306.
Heidegger, M. (1962) *Being and Time*, trans. J. Macquarrie and E. Robinson, Oxford: Basil Blackwell.
Husserl, E. (1998) *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy, First Book*, trans. F. Kersten, Dordrecht: Kluwer.
Hutcheson, F. (2002) *An Essay on the Nature and Conduct of the Passions and Affections, with Illustrations on the Moral Sense*, Indianapolis: Liberty Fund.
Mele, A.R. (2001) *Autonomous Agents, From Self-Control to Autonomy*, Oxford: Oxford University Press.
Meyers, D. (1989) *Self, Society, and Personal Choice*, New York: Columbia University Press.
—— (2002) *Gender in the Mirror: Cultural Imagery & Women's Agency*, New York: Oxford University Press.
Mill, J.S. (1962) "On Liberty," in M. Warnock (ed.), *Utilitarianism, On Liberty, Essay on Bentham*, Glasgow: Collins.
Nozick, R. (1974) *Anarchy, State and Utopia*, New York: Basic Books.
Oshana, M. (2003) "How Much Should We Value Autonomy?," in E. Frankel Paul, F.D. Miller, Jr., and J. Paul (eds.), *Autonomy*, New York: Cambridge University Press, pp. 99–126.
—— (2006) *Personal Autonomy in Society*, Hampshire: Ashgate.
Pettit, P. (1997) *Republicanism*, Oxford: Oxford University Press.
—— (2001) *A Theory of Freedom: From the Psychology to the Politics of Agency*, Oxford: Oxford University Press.
Raz, J. (1986) *The Morality of Freedom*, Oxford: Oxford University Press.
Rousseau, J.J. (1993) *The Social Contract and Discourses*, trans. G.D.H. Cole, London: Everyman.
Spector, H. (1992) *Autonomy and Rights*, Oxford: Oxford University Press.
—— (2010) "Four Conceptions of Freedom," *Political Theory* 38: 780–808.
Stendhal (2011) *Le Rouge et le Noir*, Paris: Le Livre de Poche.
Thalberg, I. (1989) "Hierarchical Analyses of Unfree Action," in J. Christman (ed.), *The Inner Citadel*, New York: Oxford University Press, pp. 123–36.
Watson, G. (1989) "Free Agency," in J. Christiman (ed.), *The Inner Citadel, Essays on Individual Autonomy*, New York: Oxford University Press, pp. 109–22.
Wolf, S. (1993) "The Real Self View," in J.M. Fischer and M. Ravizza (eds.), *Perspectives on Moral Responsibility*, Ithaca, NY: Cornell, pp. 151–69.
Young, R. (1986) *Personal Autonomy. Beyond Negative and Positive Liberty*, London: Croom Helm.

Further Reading

There are excellent collections that cover the main lines in the contemporary literature: J. Christman, *The Inner Citadel* (New York: Oxford University Press, 1989); E. Frankel Paul, F.D. Miller, Jr., and J. Paul (eds.), *Autonomy* (New York: Cambridge University Press, 2003); J. Christman and J. Anderson (eds.), *Autonomy and the Challenges to Liberalism* (New York: Cambridge University Press, 2005), and J.S. Taylor, *Personal Autonomy* (New York: Cambridge University Press, 2005). H. Spector, "Four Conceptions of Freedom," *Political Theory* 38 (2010) explores the connections among various conceptions of freedom, including conceptions of positive freedom that are related to personal autonomy.

51
POWER

Peter Morriss

1. The Intellectual Background of Modern Power Studies

Power has been surprisingly neglected in recent political and social philosophy, when compared with recent hot topics such as justice (see Section 4) and freedom (Chapter 49). Power is obviously central to the *practice* of politics; but that has not, by itself, made it appealing to many philosophers.

Instead, modern interest in power can be traced to debates in America during the cold war. It was claimed that America, and its allies, were democracies; the Soviet Union, and its allies, were tyrannies. One difference between these two systems, it was thought, was that in tyrannies power is concentrated in the hands of one person, or a few people, and the vast majority of people are powerless; in democracies, power is spread widely and more equally (Chapter 48). Quite why *this* difference was made central (rather than, say, how the systems benefitted and harmed their citizens) is not clear, but the result was to put power at the center of academic political polemic.

It was therefore somewhat of a shock when, in the 1950s, a number of empirical studies seemed to show that power in American society was very concentrated indeed. (The best known are Wright Mills 1956, and Domhoff 1967.) The response to this was fourfold. Three of these responses are not here relevant: questioning the studies on empirical grounds; reconsidering how we should think about democracy (e.g. Dahl 1956); and looking more closely at the *methodology* by which power should be studied (e.g. Dahl 1958). The fourth response was to look more closely at what power should be considered to *be*. If the study of the distribution of power should be at the center of political science, as many believed, then it would be a good idea to think carefully about what power *is*. Hence philosophical questions about the nature of power arose from disputes originally outside political philosophy, which perhaps explains the concept's marginal status within political philosophy.

By chance, this political requirement for an investigation of the concept of power coincided with the dominance in Anglo-American philosophy of "ordinary language philosophy," in which, as the name implies, so-called "ordinary language" was subjected to intense scrutiny (see e.g. Austin 1962). A concern with language was not new in philosophy: both Hobbes and Locke thought that it was crucial to use language properly and carefully, as have most philosophers before and since. The trend in the post-war period, however, was to make an analysis of meaning almost *the* central concern of philosophy. This approach filtered through to political scientists interested in power, but most ordinary language philosophers looked at other topics, and ignored power (and, indeed, the social sciences in general).

So the "power debate" in post-war political philosophy has consisted in attempting to produce a rigorous analysis of what power is, and most of this work has been carried out by non-philosophers. It was started by Dahl (1957) and Polsby (1963), who were criticized by Bachrach and Baratz (1962) and Lukes (1974). I myself have written a work in this vein (Morriss 1987). Other substantial contributions were Benn (1967) and Goldman (1972).

2. Power as a Dispositional Concept

A number of philosophical issues have emerged in this debate. One, about which there was dispute, but on which there is now some agreement, is that power should be thought of as a *dispositional* concept. (Some followers of Foucault have still failed to grasp this point.) That is to say, "power" does not refer to some event or occurrence in the world, but to a potential or possible event or occurrence. Thus, if your employer has the power to dismiss you without giving reasons, this is an important social fact (to you, at least)—but it says nothing about whether your employer *will* so dismiss you, or *has* dismissed you. If you are fortunate enough to serve out your time, undismissed, into a happy, and voluntary, retirement, your employer's decision never to dismiss you does not in any way contradict the statement that your employer *did* have this power. Just because a power is never exercised does not mean that the power was never there.

Dispositional concepts are important, but problematic. Since they do not refer to what *has* happened, but rather to what *could* happen, or *might* happen, their existence cannot be directly observed. It was therefore unfortunate that the post-war re-emergence of the study of power coincided with the dominance of an approach in social science called *behavioralism*. The motivation behind behavioralism was to create a discipline of political *science* after the model of natural science; this model was thought (following Popper) to require empirically testable propositions capable of falsification. So behavioralists demanded that practitioners of the empirical study of politics should take care to couch their empirical claims in ways that were capable, in principle, of falsification. For a proposition to be falsifiable, it must make a statement about the occurrence of some event—a statement that can (in principle) be shown to be either true or false. Therefore, it was thought, all concepts in social science should refer to things that are, in some way, directly observable.

Unfortunately political scientists (particularly in America) embraced this approach with a naive enthusiasm, and without delving deeply into the philosophy of science. By the time they were advocating behavioralism, almost all philosophers of science had rejected the view that science could purge itself of all not-directly-observable concepts. It was helpful, for instance, to describe zinc as soluble in sulphuric acid (and gold as not so soluble), even though a term like "soluble" does not directly refer to an observable event: a lump of zinc can be described as soluble not because it *has* dissolved in acid (because then there would be no lump of zinc to describe), but because (roughly) it *would* dissolve if placed in acid (which it hasn't been). But such talk about things we can't observe made behavioralists unhappy: they thought that we could not possibly *know* that this lump is soluble unless we had *observed* it dissolving. Similarly, behavioralists asserted, we can't *know* anyone is powerful if we haven't seen the power exercised. Thus Nelson Polsby—a writer who made an influential attempt to construct a behavioralist approach to power—asked (rhetorically)

> How can one tell … whether or not an actor is powerful unless some sequence of events, competently observed, attests to his power? If these events take place, then the power of the actor is not "potential" but actual. If these events do not occur, then what grounds have we to suppose that the actor is powerful? There appear to be no significant grounds for such a supposition.
>
> (Polsby 1963: 60)

But this is confused. It is, of course, true that one cannot tell whether an actor is powerful unless *some* set of observations "attests to" his power. But there is no reason at all why these observations should be of the actualization of that power. When I go to the zoo, I can see that a lion is powerful enough to eat me up by observing its jaws, teeth, and muscles, and combining these observations with my general knowledge of animals' masticatory performances. If I am still in doubt, I can observe what the lion does to a hunk of meat, and induce. Not even the most dogmatic behaviorist would declare that he couldn't know whether the lion could eat him up until it had actually done so.

This mistaken approach to philosophy of science (and common sense) got the study of power off on the wrong foot, and it was several decades before it became generally acknowledged that power is a dispositional concept.

Exactly *how* we should analyze dispositional concepts is still problematic in philosophy. The most obvious reading is in terms of counterfactual conditionals: that is, "*if* the zinc *were* to be placed in the acid, it would dissolve," or "if your employer *chose* to dismiss you, you would be dismissed." Philosophers have come up with some strong counter-examples to this analysis (see Fara 2009 for a survey), although not all are convinced by them. My own view is that these counter-examples do not pose a problem for usages of "power" in political analysis; but you should be aware that these are disputed philosophical waters.

But if the concept of power does not describe an event of some sort, why should we be interested in it? Why should it bother us what people *could* do, but don't? Answering these questions helps us to grasp when, and why, "power" matters—and when it does not matter. Consider again your (hypothetical) employer who could dismiss you at a whim—but has not done so. One *might* say that nothing has happened as a result of this power, so it makes no difference whether it is there or not: it makes no difference to you until you *are* dismissed. But that seems implausible. Supposing that you wish to remain in your job (and supposing that other equally desirable jobs are hard to find), then being subject to arbitrary dismissal would make your future uncertain; and that uncertainty might well matter. You would worry about entering into long-term commitments (such as buying a house or having children) that would only be possible if you remain in the job. If your job security was safer, you would have less uncertainty to worry about; and worry about uncertainty matters. Hence one way that being in someone's power matters is that you never know whether they might exercise that power tomorrow, at your expense. If your employer did not have such power, your life would be better. Hence an (unexercised) power does matter. This approach has been developed in much more detail by Lukes (2005), Pettit (1997) and Lovett (2010).

3. Power To and Power Over

The approach of ordinary language philosophy consists in analyzing how we use terms in ordinary language; so, applying this to power, we will need to examine how that term

is indeed used. Here there has been a difference of opinion that has lasted throughout the last 70 years, and it is a disagreement on the important question of what idea of power we should be thinking about when we think about power.

The traditional view was put clearly by Thomas Hobbes, in *Leviathan*, whose definition of "power" went as follows: "The power of a man … is his present means to obtain some future apparent good" (1651: ch. X; punctuation altered). However, when the study of power was re-started in the post-war debates, it was kicked off by Robert Dahl's definition, which went: "A has power over B to the extent that he can get B to do something that B would not otherwise do" (Dahl 1957: 80). What Dahl was here trying to define was not Hobbes's "the power of a man (person)," but the very different notion of someone's power *over* someone else. Surprisingly this change was accepted almost unquestioningly by Dahl's critics: they disagreed on *how* one should think of "A has power over B," but they did not disagree *that* we should be thinking of A's power over some B. The most sophisticated and well known such account is that by Steven Lukes (1974; see Lukes 2005: ch. 2 for a partial retraction).

This has become known (following Hannah Pitkin) as the debate between whether we should be thinking about "power to" or about "power over." (It might be better to characterize this as a "disagreement," rather than debate, as most protagonists have just asserted that one expression is their favorite, and then got on with investigating it.) "Power to" was always the traditional approach, as the passage from Hobbes illustrates. It also seems to be the more normal usage. For instance, in the latest (on-line) edition of the *Oxford English Dictionary*, the phrase "power to" occurs in the illustrative quotations more than 12 times as often as "power over" (1,672 against 138). If one checks entries on the web for the two terms, via google, the ratio is much the same, at about 10 to 1. Also, one can now search all the books that google has digitized, and produce a graph of the frequency of uses of the terms through time (using google's ngram facility); ever since 1624, "power to" is found to occur about five to ten times more frequently than "power over." Finally, exactly the same result occurs when one searches the British National Corpus, which contains over 100 million examples of ordinary English usage: when I searched in May 2011, "power over" occurred 360 times and "power to" 3,112. Of course, these are very crude measures and would need to be refined before they are any more than suggestive; for instance, the first example of "power over" produced by the British National Corpus was "After the election, Wilson handed power over to Heath" (i.e., Heath replaced Wilson as Prime Minister), which is not a use of "power over" in the usual sense. If anybody has a research assistant whom they really hate, they could ask him or her to trawl through these sources deleting the irrelevant entries; I do not have a research assistant, and so I have not done this. It might, however, be significant that in the google books data set, the usage of "power to" peaks at about 120 occurrences per million texts in the 1640s, with further peaks in the 1680s and second half of the eighteenth century (when it reached 67, from which it has been slowly declining ever since, to about 30); in other words, "power to" is used most often at times of heightened political activity, when more books about politics are being published. ("Power over" also peaks in the same decades in the seventeenth century, at the lower level of 21 per million; but since 1700 it has never exceeded 10 in any decade, and is now running at about 5.) So, at first blush, it seems plausible that "power to," and not "power over," has been the normal linguistic usage when people have been thinking about politics.

Why, then, have some social and political philosophers wanted to give priority to "power over"? One view is that social and political power (as distinct from other sorts

of power) inherently involves a relationship between people; hence we need both an A (the power-holder) and a B (the person with whom A is in a power-relationship). Thus Steven Lukes says (without, incidentally, producing any evidence for his empirical claim) that *social power* is different from A's power-to, which

> does not correspond to what, in common parlance and in the writings of philosophers, historians and social scientists, "power" is commonly taken to identify. In this more restrictive but widespread understanding, "power" is explicitly relational and asymmetrical: to have power is to have power *over* another or others.
>
> (Lukes 2005: 73)

And Keith Dowding, in rewording the distinction between "power to" and "power over" as that between "outcome power" and "social power," describes the latter term as "necessarily involv[ing] a social relation between at least two people" (Dowding 1991: 48). Both Lukes and Dowding, in different ways, want to emphasize power that involves conflict or domination.

But one has to be careful here, as Pitkin pointed out, in her excellent brief application of ordinary language philosophy to concepts of power. As she indicated, normally a relationship is something one has *with* another person, not *over* another. "And one does not 'exercise' a 'relationship' at all. What the social scientists mean by calling power relational, if I understand them, is that the phenomena of power go on among people" (Pitkin 1972: 276). This aspect of "going on among people" is adequately captured, for most purposes, by using the normal power concept of "power to," whilst ignoring those "powers to" (if there are any) that do not in some way—direct or indirect—involve other people.

Recently, Pamela Pansardi has breathed new life into this debate, by arguing (against both Dowding and myself) that the relational aspect of "power over" makes "power to" in some way a sub-category of "power over" (Pansardi 2012). I think that she makes an interesting and important case, but that the better conclusion of her argument is that there may be *more* linguistic uses of "power" than just "power to" and "power over" (see Morriss 2012b): that the phenomena of power go on among people in many diverse ways, and that "power to" and "power over" only capture some of them. For instance (as Pansardi points out) some recent feminist writing has stressed the importance of "power with" (see Allen 1999): here we have a relationship of power between at least two people, but not one that necessarily involves domination or conflict. Such a relationship might be that of a (good and empowering) teacher or psychotherapist. It seems that what we need here is more of the practice of ordinary language philosophy: in their "Saturday morning meetings," Austin and others used to pick a general concept, think of as many near-synonyms of it as they could, and then discuss how uses of these near-synonyms differed. It is perhaps a sign of the absence of philosophical input hitherto in the analysis of power, that there has not been consideration of a richer variety of power terms.

4. Why Does Power Matter?

It is part of the methodology of ordinary language philosophy that one should always ask *why* we are using a term—what job is it supposed to be doing? If we understand its

job—its purpose for us—then we can grasp much better how it should be used. These questions tended not to be asked when political scientists produced analyses of "power"; and the analyses suffered as a result. I want to suggest that there are (at least) four such uses. The first is the intensely practical use that we need to know who has power if we are going to engage in political action: as Brian Barry said (somewhat facetiously), the powerful people in any society must include those whom the CIA would want to bribe (Barry 1974: 189)—or, one might add, kill. A second use is common in moral philosophy and the law: it is a defense to a charge that one did something (or should have done something that one did not do) that one was quite powerless to do it. This is summarized in the slogan "ought implies can." I have said a little about these two uses in my book, but I will not discuss them further here.

5. Being in Another's Power

The third use is that of being in someone's power, which I illustrated earlier with the example of your powerful employer. This uses the concept of "power over," but it is important that it uses it in the passive: it is *being in the power of another* that matters, not *having* power over another. It is misleading, then, that both Dahl and Lukes have phrased their definitions in the active voice, as "A has power over B"—and both, I think, have been misled by taking this approach. So, thinking about "being in the power of another" is a useful way to proceed. However, this gives rise to complexities, which have not been properly resolved in the literature.

What is worrying for you in having your powerful employer is that he has the power to fire you at his will; and, as a consequence, for the last couple of hundred years trade unions have attempted to limit such a power of employers, so that they can only dismiss workers for some clear breach of the terms of their contract. Examples like this have led, somewhat naturally, to the belief that to be in someone else's power is bad, and that we necessarily do better by not being in the power of another. Yet that conclusion is false. We sometimes gain immeasurably by being in the power of another, providing that other is looking after our interests. This is again an insight made by feminist philosophy—and, like many such insights, it is obvious once stated, and shocking that it was previously overlooked. A child who is not in the power of a parent or teacher (or some equivalent) is likely to be in a very bad way indeed—and that remains true even though some few parents and teachers undoubtedly have used such a power to harm and not help. Adults, also, frequently benefit from being in some relationship with another in which the other has power over them: it might at some point be beneficial to you if, for instance, a doctor has the power to rip your heart out and replace it with that of a cadaver, which is an exercise of power over you if anything is. (We can try to blunt this example by talking of informed consent, and pretending that the doctor is somehow acting as our agent and it is we that have the power, but this is delusional; once we go under the anesthetic, we are as powerless as can be, and even when conscious we have to take an enormous amount on trust.)

So it seems that being in the power of another is sometimes beneficial. And that suggests that what is undesirable about your powerful employer is *not* that he is powerful, per se: it is that he has the power to harm (by firing) as well as help (by hiring in the first place). It might be tempting, then, to look at whether the power-holder does indeed act in a way that harms you rather than benefits you, and differentiate between the benevolent powerful employer and the ruthless one. This might often be a sensible

thing to do, but as soon as we concentrate on what the employers are *doing*, we are no longer focusing on their *power*—for, remember, "power" is a dispositional concept that refers to what *can* happen, not what *does* happen. There are, then, difficulties in providing an adequate analysis of why being in someone's power matters, in and of itself; and these difficulties have not yet been resolved.

6. Being Powerless or Powerful

I will now turn to the fourth use we have for "power," which is, I suggest, the use that was at work in the original dispute about power in America. This involves looking at who has the power to do what. That this is certainly an important matter can be illustrated by examining Amartya Sen's discussion of capabilities (see e.g. Sen 1993). Sen's theory has two key terms: functionings and capabilities. "Functionings" are closely connected to the familiar concept of welfare: states of a person that can be accepted as desirable, such as being adequately nourished and having good health. But Sen then says that knowledge of "functionings" is not enough. Sen often uses the example that we should not look merely at whether a person has enough food: food is, of course, essential for adequate functioning and welfare, but we draw a very different normative lesson about those who are starving because they choose to starve themselves (for instance, as part of a religious purification ritual) than we do about those unable to get hold of any food. That is to say, it is often less important what goods people *actually* have than what goods they are *able* to have. We should be concerned that people are in positions such that they can get whatever is necessary for a satisfactory life; not that they actually do get these things, should their wishes be to the contrary. We must, in short, include an assessment of what Sen calls people's "capabilities"—which is to say their powers (in the sense of power as "power to"). (It is unfortunate, and potentially misleading, that Sen talks of "capabilities" and not powers. For "capability" does not seem to mean what Sen wants it to mean: it means, as the *Oxford English Dictionary* has it "an undeveloped faculty or property" (my emphasis); to be capable is to be able to acquire a power, as I am (perhaps) capable of understanding Arabic, should I learn it, whilst I currently lack the ability to understand Arabic. That is not what Sen wants: he wants to understand by a capability something that one could indeed do now. It is one of the themes of this chapter that it is a great help to make sure that we get our vocabulary right at the start, and it is a pity that Sen failed to do that. By now, the semi-technical term "capability" has become entrenched in the discipline, in Sen's sense; but that is still regrettable.)

It should be clear that our power to eat an adequate diet is of immense importance to those of us who are lucky enough to have it; and perhaps of even more importance to those who lack it. Sometimes, no doubt, those who lack the power to feed themselves do so in circumstances in which nothing could be done about it: there is simply not enough food to go around. But Sen's own empirical work on famines (Drèze and Sen 1989) has shown that famines are rarely the consequence simply of lack of food: they are at least as often caused by a failure to distribute what food there is in an equitable way. When this occurs, it is crucial that the starving lack the power to feed themselves, and we are here certainly talking about power "going on among people" (in Pitkin's useful phrase), even if it is not clear what relationships are present and absent, and it is quite likely that neither "power over" nor Dowding's "social power" is present. It should certainly interest a political philosopher if there is a gross disparity in the powers that people have to feed themselves (as there is, at least on an international scale) when this disparity

could be different were the social arrangements different. Allocating powers in desirable ways is the rightful task of government, and it can (and should) be held to account if it fails; at a more general level, the political-economic system as a whole (e.g. capitalism) could be assessed by looking at how well it allocates such powers. (There is a pernicious doctrine dominant in many Western countries that "the market," unlike the state, is an impartial allocator of resources, and therefore beyond criticism; but "the market" is a human construct, and could be—and usually has been—constructed in a different way; it is therefore essential that we keep asking whether this method of running our economies is the most desirable one.)

Political power is a form of second-order power. I have said that social institutions allocate many of the powers that people have; political power is the power to control these institutions. How widespread this second-order power was was indeed the subject of the post-war "power debate," and it is perhaps even more relevant now. In the West most of us live in democracies, in which most adults have a vote, and the power of votes is usually more-or-less equal. But this still leaves undetermined how much control we have, through votes, over our major power-allocating institutions. I am writing this in Ireland, at a time when the government is in the middle of passing an austerity package which has been dictated to it by a combination of the International Monetary Fund and a consortium of banks; the (previous) government was coerced into surrendering its powers to regulate the economy by another consortium of banks. Each of us has a vote to determine who is in the government; but no vote to determine who runs the banks. Once again it is an open, and important, question how widespread the second-order power to control power-allocating institutions indeed is at present: how far the scope of democracy reaches. The power studies that were a thriving sub-discipline after the war are no longer being carried out; but it may be that they are now needed more than ever.

There is also an intimate connection, of some sort, between power and *freedom*: one cannot be free to do something if one is powerless to do it. There has recently been much important investigation of the concept of freedom by political philosophers (see, for instance, Carter 1999), but they have paid virtually no attention to the similarities, and differences, between a person's freedom and that person's power. Since these do refer in some ways to cognate ideas, it might be useful to explore them together. I have attempted to start doing this in Morriss (2012a) and elsewhere, but a lot more work needs to be done on understanding the *connections* between the concept of power and the other main terms in our political vocabulary.

7. Voting Power and Democratic Theory

Another area of academic research connecting political philosophy and power, which has been thriving recently, is the normative study of voting power. The path-breaking work of Dan Felsenthal and Moshé Machover (1998) for the first time established a rigorous framework for dealing with these issues (and also usefully summarized the previous literature). Subsequent works by Felsenthal and Machover, and by a number of younger scholars influenced by them, have produced further important results. These tend to be highly technical and mathematical, but one of the consequences of their work is to highlight gaps in the theory of democracy.

For instance, consider how we should construct a system of two-tiered voting. In two-tiered systems, the voters elect delegates or representatives, who in turn make deci-

sions. Parliaments usually work like that, of course, and, in electoral systems with single-member constituencies, democracy has been taken to require that the franchise should be extended as widely as possible, each voter should have one vote, and constituency boundaries should be drawn so that each constituency contains roughly the same number of people. But what should happen when the constituencies are pre-given, and are very different in population? Here, if we give the representatives one vote each, then voters in large constituencies will be under-represented; that would go against the principle that all voters should be equal. So we have to consider giving the representatives votes of *different* weight, as occurs in the Electoral College in the United States. But how should these different weights be determined? By now the mathematics of this has been largely resolved; but the political philosophy lags behind. (The best recent work on this has been done on the allocation of weights in the Council of Ministers in the European Union. For an introductory discussion, see Felsenthal and Machover 2000; the more recent developments in the EU are described in Felsenthal and Machover 2009.) It would be nice to have an algorithm from democratic theory, like the simple "one person, one vote"; but at present no such algorithm exists.

I will briefly mention three areas on which work is needed here. One issue is whether we should aim to ensure that each voter has *equal* power to determine a decision, irrespective of which country she is in. Or does democracy require that we *maximize* each person's voting power? Unfortunately, these two aims can produce different results. We can see this by realizing that often the only way of ensuring that each voter has *exactly* equal voting power is by giving each voter no power at all; this would satisfy the equality requirement, but defeat the whole object of the exercise. So the equality requirement is inadequate without including the further condition that more power is better than less; sometimes we have to sacrifice some equality. But it is still unclear how the two values of equalizing and maximizing voting power should be traded off against each other.

Another issue is when should the status quo be privileged by requiring a super-majority for change? At present, the EU requires such super-majorities, apparently because individual countries are more worried about changes being made at their expense than they are about their inability to initiate changes that would benefit them—the EU tends to be seen (by domestic politicians, at least) as more of a threat than an opportunity. When, if ever, is such bias in favor of the status quo acceptable?

The third issue is a bit more complicated: it is that it is not clear how much importance, if any, should be attached to the interests of a *state* as opposed to the interests of voters *in* that state. If we think of the important actors in the Union as the *states*, then presumably each state should have the same number of votes, irrespective of population (as occurs in the United Nations General Assembly, and also in the United States Senate). On this view, if some states are bigger than others, this is as irrelevant as that some people are bigger than others; in neither case should they get more votes because of it. The polar opposite position is that the grouping into states is simply irrelevant; the only actors that matter are people (the individual voters). This position is problematic, because then we should simply redraw constituency boundaries so that the constituencies are equal in size, as we do in national politics. If we *don't* want to do that, then we *are* saying that the existence of a state is somehow important; and presumably that importance has to be recognized when we allocate the weight of votes.

Related to this is how much normative relevance, if any, we should place on the existence of clear empirical voting patterns, representing shared interests. One view is that voting patterns have no normative relevance at all: each person should be given one

vote, which he or she can use however they wish. We should not construct the normative apparatus based on some expectation of how a voter will (or perhaps should) vote. If that is the case, then it has been shown that the weight of a representative should be given by the *square-root* of the number of electors she represents (see Felsenthal and Machover 2000: 22–6). But we get a different result if we assume that nations and states actually *mean* something to their constituents; that they have formed as they have to represent some common interest—or, perhaps, that their existence has *created* a common interest. If we assume that there is an identity of interest within each state, then the calculation that produces the square-root rule does not work; instead, the weight of the representative should be proportional to the actual number of voters they represent, not the square-root of that number. That makes a big difference in practice: within the EU, proportionally to population, Germany should have about 200 times Malta's vote, whilst by the square-root rule Germany should have only 14 times as many votes as Malta.

So we have seen that there are reasons to allocate votes to the representative in three starkly different ways: (a) one vote per representative, (b) a vote weighted in proportion to the population represented, and (c) a vote weighted in proportion to the square-root of the population represented. In the academic literature, option (c) has had most support. However, the assumptions that underlie it are identical to assuming that voters are allocated to states at random. But that is precisely *not* the case; there is supposed to be at least some element of identification with one's state, and states would have no *raison d'être* if that were not the case. We cannot first assume that states are important enough to provide the basis for grouping voters, and then assume that such a grouping has no importance at all. Yet it is not clear what alternative assumption we should make. A large gap in democratic theory has been exposed here.

We have returned to the role of power in democratic theory, and we have found a surprising amount of work left to do. Some of this will doubtless require new tools and new approaches, but there are also advances to be made by building on the foundations that have been put in place. The neglect of power by top-rate political philosophers has left many fruitful fields for further research.

Related Topics

Republicanism, The Capability Approach (and Social Justice), Freedom, Democracy

References

Allen, A. (1999) *The Power of Feminist Theory: Domination, Resistance, Solidarity*, Boulder, CO: Westview Press.

Austin, J.L. (1962) *How to Do Things With Words*, Oxford: Oxford University Press.

Bachrach, P. and M.S. Baratz (1962) "Two Faces of Power," *American Political Science Review* 56: 947–52.

Barry, B. (1974) "The Economic Approach to the Analysis of Power and Conflict," *Government and Opposition* 9: 189–223.

Benn, S. (1967) "Power," in P. Edwards, ed., *The Encyclopaedia of Philosophy*, New York: Macmillan, vol. 6, pp. 424–7.

Carter, I. (1999) *A Measure of Freedom*, Oxford: Oxford University Press.

Dahl, R.A. (1956) *A Preface to Democratic Theory*, Chicago, IL: University of Chicago Press.

Dahl, R.A. (1957) "The Concept of Power," *Behavioral Science* 2: 201–15.

Dahl, R.A. (1958) "A Critique of the Ruling Elite Model," *American Political Science Review* 52: 463–9.

Domhoff, G.W. (1967) *Who Rules America?*, Englewood Cliffs, NJ: Prentice Hall.

Dowding, K. (1991) *Rational Choice and Political Power*, Aldershot: Edward Elgar.
Drèze, J. and A.K. Sen (1989) *Hunger and Public Action*, Oxford: Oxford University Press.
Fara, M. (2009) "Dispositions," *The Stanford Encyclopedia of Philosophy*, at http://plato.stanford.edu/archives/sum2009/entries/dispositions.
Felsenthal, D.S. and M. Machover (1998) *The Measurement of Voting Power: Theory and Practice, Problems and Paradoxes*, Cheltenham: Edward Elgar.
Felsenthal, D.S. and M. Machover (2000) *Enlargement of the EU and Weighted Voting in its Council of Ministers*, London: London School of Economics, eprint 407; available on the LSE eprint web-site.
Felsenthal, D.S. and M. Machover (2009) "The QM Rules in the Nice and Lisbon Treaties: Future Projections," *Homo Oeconomicus* 26: 317–40.
Goldman, A.I. (1972) "Toward a Theory of Social Power," *Philosophical Studies* 23: 221–68.
Hobbes, T. (1651) *Leviathan* [many editions].
Lovett, F. (2010) *A General Theory of Domination and Justice*, Oxford: Oxford University Press.
Lukes, S. (1974) *Power: A Radical View*, London: Macmillan.
Lukes, S. (2005) *Power: A Radical View*, 2nd edn, Basingstoke: Palgrave Macmillan.
Morriss, P. (1987) *Power: A Philosophical Analysis*, Manchester: Manchester University Press (2nd edn, 2002).
Morriss, P. (2012a) "What is Freedom if it is Not Power?," *Theoria* 59 (132): 1–15.
Morriss, P. (2012b) "A Response to Pamela Pansardi," *Journal of Political Power* 5: 91–9.
Pansardi, P. (2012) "*Power to* and *Power over*: Two Distinct Concepts of Power?," *Journal of Political Power* 5: 73–89.
Pettit, P. (1997) *Republicanism: A Theory of Government and Freedom*, Oxford: Oxford University Press.
Pitkin, H.F. (1972) *Wittgenstein and Justice*, Berkeley: University of California Press.
Polsby, N.W. (1963) *Community Power and Political Theory*, New Haven, CT: Yale University Press.
Sen, A.K. (1993) "Capability and Well-being," in M. Nussbaum and A.K. Sen, eds., *The Quality of Life*, Oxford: Oxford University Press.
Wright Mills, C. (1956) *The Power Elite*, New York: Oxford University Press.

Further Reading

Dowding, K., ed., (2011) *Encyclopedia of Power*, Thousand Oaks, CA: Sage.
Haugaard, M. ed., (2002) *Power: A Reader*, Manchester: Manchester University Press.
Lukes, S., ed., (1986) *Power*, Oxford: Basil Blackwell.

52
AUTHORITY AND LEGITIMACY

Fabienne Peter

Political legitimacy is sometimes understood as a descriptive concept—one that captures people's faith in their political institutions and their political leaders (Weber 1964). But the focus here shall be on the normative dimension of legitimate political authority—on what justifies political institutions and the decisions made within them. This chapter will address the following main questions. First, under what conditions does an institution such as a national government have the right—as opposed to merely the power—to rule? And do these same conditions also create a duty to obey the commands of the political authority? Second, does legitimate political authority require democracy? If so, what kind of democracy? And finally, is there something specific about national governments or do problems of legitimacy also arise at the level of global governance institutions?

1. Legitimate Authority

Although some have recently proposed to interpret legitimate authority as a Hohfeldian *power* (Applbaum 2010; Edmundson 2010), according to the most common usage of the terms, a legitimate political authority has the *right* to issue laws and to enforce these laws if necessary. Without legitimacy, any attempt to rule is either an exercise of *de facto*, but unjustified, authority or a mere unjustified use of power. Even that concept of legitimate authority allows for very different conceptions of political legitimacy and authority. Differences stem from different interpretations of the conceptual relations between authority and legitimacy and between authority, legitimacy, and political obligations. I shall focus on legitimate authority in this section and on the relation between authority, legitimacy, and political obligations in the next.

On one influential definition, legitimacy is what justifies authority. Joseph Raz works with this understanding of legitimate authority (1988, 1995, 2006). Legitimate authority differs from *de facto* authority in that the former actually possesses the right to rule while the latter is merely effective at claiming to have this right.

Raz takes authority to be the basic concept. According to Raz, the directives of an authority give others a particular kinds of reasons for action. One feature of these reasons is that they are content-independent. They are reasons for action because of their source—because the authority is telling you so—not because of their content. In addition, the directives of a legitimate authority create reasons for actions that do not outweigh, but pre-empt or replace

your other reasons for action. Raz calls this feature the "pre-emptive thesis" (1988: 46). According to Raz's "dependence thesis" (1988: 46), the legitimacy of an authority depends on the quality of the reasons given by the directives of the authority. Raz's famous "normal justification thesis" draws on the pre-emptive and the dependence theses. It says that an authority is legitimate if

> the alleged subject is likely to better comply with the reasons which apply to him … if he accepts the directives of the alleged authority as authoritatively binding and tries to follow them, rather than by trying to follow the reasons which apply to him directly.
>
> (Raz 1988: 53)

On this view, legitimate authority serves those governed—Raz thus calls it the "service conception" of authority (Raz 1988: 56). Note that *de facto* but illegitimate authority does not serve those governed.

On the authority-based conception of political legitimacy, the use of coercive power is not the main defining feature of political institutions (Raz 1988: 24). As Leslie Green argues, following Raz: "[c]oercion threats provide secondary, reinforcing motivation when the political order fails in its primary normative technique of authoritative guidance" (Green 1988: 75). What characterizes a (*de facto*) authority is that it is effective at getting people to follow its commands.

Defenders of the coercion-based conception of legitimacy maintain that the fundamental normative problem of politics is, to use Arthur Ripstein's words (2004: 6), how people "may legitimately be forced to treat each other." On this view, coercive political power is thus not merely a means to secure some ideal of political authority. Instead, it is conceptually inseparable from political authority and the main target of justification. On this definition, legitimacy is what justifies the use of coercive political power (Rawls 1993, 2001; see also Ripstein 2004, 2009).

Kant is the principal historical reference for the coercion-based conception of legitimate authority. The civil state, according to Kant, establishes the rights necessary to secure equal freedom. The legitimacy problem that the civil state creates is that such a system of rights is both necessary and coercive. Kant defines coercion as a restriction of the freedom to pursue one's own ends. Any right of a person—independently of whether it is respected or has been violated—implies a restriction for others (cf. Kant, *Theory and Practice*, Part 2; Ripstein 2004: 8; Ripstein 2009; Flikschuh 2008: 389f.). Coercion, in this view, is constitutive of rights and hence of political authority.

The best known contemporary articulation of the coercion-based view is found in Rawls's *Political Liberalism*. For Rawls the problem of legitimacy arises from a tension between, on the one hand, the fact that in liberal, democratic societies political power "is regarded as the power of free and equal citizens" (Rawls 2001: 40) and, on the other, the reasonable pluralism that characterizes liberal, democratic societies (Rawls 1993: 63f.). Rawls asks: "in the light of what reasons and values … can citizens legitimately exercise … coercive power over one another" (Rawls 2001: 41)? His answer is that "political power is legitimate only when it is exercised in accordance with a constitution (written or unwritten) the essentials of which all citizens, as reasonable and rational, can endorse in the light of their common human reason" (Rawls 2001: 41). In this conception of political legitimacy the justification of coercion depends on the idea of public reason (Rawls 1999). Rawls's conception of legitimate authority thus differs from Raz's

with respect to both what is regarded as in need of justification and what counts as justification. According to Raz, the justification of authority depends on ordinary moral reasoning. Rawls's idea of public reason relies on the method of public—or "political" as opposed to "metaphysical"—justification that Rawls has developed in response to critics of his theory of justice as fairness (Rawls 1993). It is an essential feature of public justification that it is freestanding, i.e. independent of comprehensive doctrines. As Rawls puts it in his *Justice as Fairness: A Restatement* (2001: 27): "Public justification is not … simply valid argument from given premises (though of course it is that). … [W]hen the premises and conclusions are not acceptable on due reflection to all parties in disagreement, valid argument falls short of public justification." There is, however, much debate about how best to interpret Rawls's idea of public reason—what it contains, to whom and to which questions it applies, etc.—and its relation to political legitimacy (e.g. D'Agostino 1996; Gaus 1996, 2011; Reidy 2000; Quong 2004; Peter 2008; Weithman 2011).

Rawls's conception of political legitimacy is related to his conception of justice as fairness. According to Rawls, both justice and legitimacy draw on the same set of fundamental political values. But legitimacy imposes weaker requirements than justice. That is to say, political institutions and the decisions made within them may be unjust yet legitimate (Rawls 1995). One way to understand this is that legitimacy applies primarily to political institutions, whereas justice applies to the full set of political, social and economic institutions. On this view, legitimacy, although weaker than justice, is more fundamental. Legitimacy, as Paul Weithman (2011: chapter 10) characterizes this view, constrains the adoption of principles of justice—i.e. the conception of justice. On an alternative view, which Weithman favors, legitimacy merely formulates constraints on the application of principles of justice. On this view, to say that justice should not be imposed on people is to say that the implementation of principles of justice should not be imposed.

2. The Relationship between Legitimacy, Authority, and Obligation

Many would say that if a political authority is legitimate, then those under its rule have a duty to obey its commands. You might not agree with every aspect of how your government handles your tax money, but if the government agencies in question and the decisions they are making are legitimate, you still normally have a duty to pay your taxes. On this, quite widely held view, political legitimacy, political authority, and political obligations are three concepts that can interlock harmoniously. Both Rawls and Raz hold this view. On Rawls's account, legitimacy obtains if the citizens can justify the coercive power of their political institutions to each other. This also explains why citizens have a *pro tanto* moral duty to accept the constraints these institutions impose and obey the decisions made within them. On Raz's account, it follows from his service conception that if a political authority is legitimate, you are likely to better comply with reasons that apply to you if you obey the directives of that authority. The duty to obey the commands of a legitimate authority follows from the bindingness of the reasons that apply to you anyway.

Some philosophical anarchists, by contrast, argue that the three concepts cannot be made to fit together. Robert Paul Wolff's influential *a priori* critique of the possibility of a legitimate state, which focuses on a tension between autonomy and political obligation is of this kind (Wolff 1970; the distinction between *a priori* and *a posteriori*

forms of anarchism is from Simmons 2001: 104f.). Wolff starts with the observation that social contract arguments (Kantian arguments in particular) for the legitimacy of political authority presuppose the importance of individual autonomy. But, he argues, one cannot both be autonomous and be subject to the rule of a political authority. If individual autonomy is a presumption in arguments for the possibility of legitimate authority, then there cannot be a general—content-independent—duty to obey the commands of a political authority. Wolff concludes from this that legitimate political authority is impossible. What puts this conclusion into question, however, is that Wolff himself concedes that certain forms of democratic authority may be legitimate.

More moderate philosophical anarchists do not believe that the three concepts are in principle incompatible. But nor do they believe that legitimate or justified authority necessarily entails political obligations. One way to get to this is by interpreting legitimate authority, not as a right to rule, as philosophers commonly do, but as a Hohfeldian moral power to rule (Applbaum 2010; Edmundson 2010). Unlike a right, a moral power does not come with correlative duties—the relevant correlatives are liabilities.

A typical view of this kind is that what justifies political authority might not coincide with what creates political obligations. John Simmons probably has the clearest account of how one might envisage this possibility (Simmons 2001; see also Green 1988; Edmundson 1998). Simmons distinguishes between the justification of political authority and political legitimacy. According to Simmons, the authority of a political institution such as nation states depends on its moral justifiability—it has to be shown that having a state is better than not having one (Simmons 2001: 125). But showing that the authority of states is thus justified does not entail showing that you have a duty to obey any of the particular directives of the state in which you live. Whether or not you have such duties depends on that state's legitimacy—on its right to impose specific duties on you—and legitimacy depends on actual consent (Simmons 2001: 137). In the absence of consent, even a morally justified state will not have the right to force you to follow its directives. On Simmons' account, it is thus possible that a political authority is morally justified to rule yet lacks legitimacy. If this is the case, those under such a state's rule do not have an obligation to obey its commands.

Simmons takes his two-stage account to be an interpretation of Locke's approach to authority and legitimacy. Simmons argues that Rawls's conception of political legitimacy—and Kantian approaches in general—mistakenly blur the distinction between the justification of political authority and legitimacy. They wrongly collapse the problem of what justifies the state and the problem of why—if at all—one ought to obey the commands of the particular state one is living in.

It is not just Kant, however, but Thomas Hobbes and Jean-Jacques Rousseau too, who take a different approach to this issue. On their views, political authority is not transferred from the state of nature to the civil state; instead, political authority is created in the civil state. While Hobbes (see his *Leviathan*) did not sharply distinguish between political power, *de facto* political authority, and legitimate political authority, Rousseau and Kant elaborated on these distinctions. Both connect the problem of legitimacy to the coercive power of political institutions and anchor the justification process within the civil state. Rousseau, in *On the Social Contract*, argues that political authority is created through the implementation of a democratically justified set of laws, i.e. laws that reflect the general will. Attempts to rule that do not reflect the general will amount to an unjustified use of coercive power. Kant, by contrast, interprets the social contract hypothetically. As mentioned, his criterion for legitimacy is that each law should be

such that all individuals could have consented to it. Again, the justification of political authority and its legitimacy stem from the same normative source.

The difference between Simmons' two-step approach and the Rawlsian one-step approach is probably best understood as centering on the question of what counts as justification. Simmons' argument for a distinction between the justification of the state and political legitimacy hinges on a conception of justification in which the second-person standpoint is irrelevant. For Rawls, by contrast, justification in the context of legitimacy and justice is necessarily from the second-person standpoint (see also Gaus 1996, 2011).

3. Sources of Normativity

The idea that some form of individual consent is the ultimate source of legitimate authority took hold in Western political philosophy in parallel with the consolidation of the view that the sovereign state is the exclusive locus of legitimate political authority. Individual consent gradually replaced God's will as the main source of normativity (Skinner 1978; Tuck 1993). While consent theories still are important, they do not represent the only attempt to theorize the sources of legitimate authority. Other theories start from mutual obligations, beneficial consequences, and democracy. I shall discuss mutual obligations, consent, and instrumentalist theories in this section and democracy in the next.

Ronald Dworkin (1986) treats obligations as the fundamental normative concept for legitimate authority. What he calls "associative obligations" arise, not from legitimate political authority, but directly from membership in a political community. Many have, however, questioned the plausibility of his claim that obligations are conceptually prior to legitimate authority (e.g. Wellman 1997; Simmons 2001).

Margaret Gilbert has a fresh take on the idea of how mutual obligations may be the normative source for legitimate authority based on collective intentions (Gilbert 2006). She combines a Razian account of authority with the plural subject theory she developed earlier (e.g. Gilbert 1996). According to plural subject theory, social life is constituted by joint commitments between the members of social groups. Once a plural subject is formed, its intentions normatively override individual intentions. For as long as the plural subject exists, individuals ought not to follow their interests or inclinations independently of the aims of the plural subject. In her paradigmatic example of two individuals who, by joint commitment, are on a walk together, it is normatively inadmissible for one of them to just turn around and go home. Unless the joint commitment that created the plural subject is revoked—again a joint process—the person is bound to form intentions that support the aims of the plural subject. Gilbert argues that not only small-scale social groups, but whole societies can be understood as plural subjects. This is the background for her main thesis, which is that the best explanation for why individuals have a content-independent obligation to obey the commands of a political authority is that the authority represents the plural subject (Gilbert 2006: 248). The argument Gilbert presents is quasi-contractual in one respect: without joint commitments, there is no plural subject. But it is importantly different from consent theories because joint commitments do not, directly, give rise to obligations. An isolated individual's expression of a willingness to commit to a joint goal has no normative force.

Consent theories, by contrast, locate normativity in the acts or expressions of intentions of separate individuals. Consent theories vary with the type of consent that is

being invoked—actual or hypothetical, tacit or explicit, etc. More fundamentally, there are different ways in which the relation between consent and legitimate authority may be understood (Raz 1995: 356). According to Locke and contemporary followers such as Nozick (1974) or Simmons (2001), some form of actual consent of those governed is a necessary condition for the legitimacy of political authority. This is the most typical form consent theory takes. Rousseau defended a variant of this theory that linked the expression of consent to democratic decision-making. Hypothetical consent theories aim at some ideal of public reasoning. According to such theories, legitimacy requires that all individuals could have consented to the laws an authority issues. Kant invoked the social contract idea in this way. The social contract is a thought-experiment that functions as a test "of any public law's conformity with right" (Kant "On the Common Saying", AK 8: 294). Finally, there are theories that take the conditions of legitimate political authority to imply that those governed by that authority are under an obligation to consent. David Estlund (2008: 177ff.) defends a version of hypothetical consent theory that illustrates this version. What he calls "normative consent" is a theory that regards non-consent to authority, under certain conditions, as invalid, and that puts individuals under an obligation to consent. Estlund uses normative consent theory to explain why those who disagree with certain democratically approved laws nevertheless have a duty to obey these laws.

Although consent theory has been long dominant, there are many well-known objections to it, especially to Locke's version. Already David Hume, in his essay "Of the Original Contract," (Hume 1965 [1748]) and many after him, object to Locke that consent is not feasible, and that actual states have almost always arisen from acts of violence (see also Filmer 1991 [1680]). Hume's own solution was, like Jeremy Bentham (1987 [1843]) later, to propose to justify political authority with reference to its beneficial consequences. In the utilitarian view, legitimate political authority should be grounded on the principle of utility. This conception of legitimacy is necessarily a moralized one: the legitimacy of political authority depends on what morality requires. Bentham argues that the Hobbesian idea that political authority is created by a social contract gets things the wrong way around. According to Bentham, it is the state that creates the possibility of binding contracts. What this objection overlooks, however, is that actual consent is an agreement; it need not take the form of a contract (Gilbert 2006: 55).

Many are not convinced that utilitarian instrumentalist reasoning provides a satisfactory account of political legitimacy, however. Rawls (1971: 175f.) and Jeremy Waldron (1987: 143f.) object that an instrumentalist approach will ultimately only convince those who stand to benefit and will lack an argument to convince those who stand to lose.

Wellman counters this objection with a Samaritan account of political legitimacy. He claims that "what ultimately legitimizes a state's imposition upon *your* liberty is not merely the services it provides *you*, but the benefits it provides *others*" (Wellman 1996: 213; his emphasis). In his account, a state's legitimacy depends on it being justified to use coercion to enforce its laws. He suggests that it can be grounded in the Samaritan duty to help others in need. The thought is that because "political society is the only vehicle with which people can escape the perils of the state of nature" (Wellman 1996: 216), people have a Samaritan duty to provide to one another the benefits of a state. Associated restrictions of their liberty by the state, Wellman claims, are legitimate.

Raz also believes that the objection can be answered. His argument invokes his normal justification thesis. As explained above, Raz believes that if an authority is

legitimate, its directives are such that they help those governed to better comply with reasons that apply to them. According to Raz (1995: 359), "[g]overnments decide what is best for their subjects and present them with the results as binding conclusions that they are bound to follow." Samuel Hershovitz (2003) has criticized Raz's argument for being incompatible with democracy. He argues that in a democracy, governments are not above those they rule—the "binders" and those "bound" are identical. The members of the democratic constituency themselves decide what they consider to be best for them.

4. Democratic Legitimacy

Building on Rousseau and Kant, many contemporary political philosophers take democratic decision-making and/or some form of public reason or public reasoning to be the main source of legitimate authority. It is a particular feature of democracy that the right to rule is held by those who are ruled. Democratic authority is entailed by some account of the conditions under which citizens may legitimately exercise coercive power over one another and vice-versa. For this reason, political authority, in a democratic context, does not exist independently of its legitimacy.

Rousseau's articulation of this thought has been very influential for contemporary political philosophy. Rousseau addressed the difficulties with Locke's idea that legitimate authority could be based on tacit consent. According to Rousseau, democratic participation is necessary for legitimate political authority. Active participation by all might not generate a consensus, of course. So why would those who oppose a particular decision be bound by that decision? Rousseau's democratic theory offers an ingenious solution to this problem. Rousseau distinguishes among a citizen's private will, which reflects personal interests, a citizen's general will, which reflects an interpretation of the common good, and the general will, which truly reflects the common good. A democratic decision is always about the common good. In democratic decision-making, citizens thus compare their interpretations of the general will. If properly conducted, it reveals the correct view of the general will. This is the legitimate decision. On Rousseau's view, since the democratic decision, if conducted properly, correctly reveals the general will, those who voted against a particular proposal will recognize that they were wrong and will adjust their beliefs about what the general will is. In this way, individuals are bound only by their own will—their interpretation of the general will—but everyone is bound by a democratic decision.

Rousseau's approach appeals to many democratic theorists today because it combines both procedural and substantive considerations. On views of this sort, political institutions and decisions are legitimate if decisions are being made in the right way and if the decisions themselves are, in a certain sense, right. In contemporary Anglo-American political philosophy, Joshua Cohen has probably most influentially pursued this line of thought (Cohen 1997a, b). Cohen argues that unlike aggregative theories of democracy, deliberative democracy is uniquely able to solve the problem of political legitimacy that stems from the following dilemma. On the one hand, democrats will maintain that substantive judgments are legitimate only if they have been made through appropriate democratic procedures. On the other, however, sometimes the results of such procedures are in conflict with certain widely held values such that their legitimacy is undermined. Deliberative democratic procedures solve this problem because they ensure that reasons are given that are acceptable to everyone; not all substantive judgments will pass this

test. Cohen's argument builds on themes from Rawls's *Political Liberalism*, combined with a Habermasian understanding of deliberative democracy (Habermas 1996; see also Manin 1987; Knight and Johnson 1994). As Habermas argued: "Deliberative politics acquires its legitimating force from the discursive structure of an opinion- and will-formation that can fulfill its socially integrative function only because citizens expect its results to have a reasonable *quality*" (Habermas 1996: 304).

Estlund questions this line of argument altogether. He argues that such accounts of deliberative democracy cannot explain why those who disagree with a democratic decision are nevertheless bound by it (Estlund 2008: 85ff.). He believes that the problem with such accounts is that they fail to show how democratic decision-making could promote values such as justice or the common good. Where Estlund agrees with Cohen and other deliberative democrats, however, is that he also defends what I have called a rational proceduralist conception of democratic legitimacy (Peter 2008). All proceduralist conceptions differ from instrumentalist conceptions of political legitimacy in that they regard democratic procedures as non-instrumentally necessary for legitimacy. Typically, arguments in defense of proceduralism focus on the value of political equality. Thomas Christiano's book *The Constitution of Equality* (2008a) is a good example of this kind of defense. But only pure proceduralist conceptions regard democratic procedures as sufficient for legitimacy. Rational proceduralist conceptions, by contrast, regard a democratic decision as legitimate only if it has both been made through appropriate decision-making procedures and if the decision itself is of a certain quality.

Estlund proposes to define the quality of the outcomes of democratic decision-making in terms of their correctness—whether they are truly just or truly approximate the common good. But he also argues that the legitimacy of political authority cannot be established in terms of correctness only. He rejects such instrumentalist "correctness theories" of legitimacy (Estlund 2008: 99) because it will be impossible to solve the problem of who decides what the correct outcome is. Estlund's solution is to use correctness as a selection device among alternative decision-making procedures (for a related idea, see Pettit 2001; List 2006). He argues that among the procedures that qualify, democracy is the best—but not infallible—procedure. As Estlund puts it: a democratic decision is legitimate if it has been made through a procedure that "can be held, in terms acceptable to all qualified points of view, to be epistemically the best (or close to it) among those that are better than random" (2008: 98).

Pure proceduralist conceptions of democratic legitimacy put all the emphasis on procedural values. According to such conceptions, democratic decisions are legitimate as long as they are the result of an appropriately constrained democratic decision-making procedure.

Compared to Estlund, Gerald Gaus puts more emphasis on the fairness of the procedures and less on the outcome itself. In his view, "the voting mechanism constitutes a fair way to adjudicate deep disagreements about what is publicly justified," at least as long as "each citizen presents what he or she believes is the best public justification" (Gaus 1997: 234; see also Gaus 2011). I have built on this idea and argued that a pure proceduralist does not need to disregard epistemic values entirely (Peter 2008). The key step in the argument is the distinction between consequentialist and proceduralist social epistemologies (Goldman 1999). Social epistemologies explain the epistemic value of social practices—e.g. of those related to democratic decision-making. A consequentialist epistemology relates the epistemic value of democracy to the quality of the outcomes it generates. A proceduralist social epistemology, by contrast, highlights procedural epistemic values. The

thought is that the epistemic value of deliberative democratic decision-making can be understood as depending on its procedural features—on how inclusive it is, for example, or how thoroughly the knowledge claims on which particular proposals rest have been subjected to criticism. On this interpretation, democratic legitimacy is jeopardized not just by problematic influences on the democratic decision-making process of privileged access to social and economic institutions, but by the influence of privileged access to epistemic institutions as well. The resulting conception of democratic legitimacy regards democratic decisions as legitimate if they are the outcome of a deliberative decision-making process that satisfies some conditions of political and epistemic fairness (Peter 2008).

5. The Realm of Legitimate Political Authority

In contemporary political philosophy, discussions of political authority and legitimacy have until recently centered on the sovereign state and on relations between sovereign states. This has, of course, not always been the case. The question what constitutes legitimate authority has been around for much longer than the modern idea of the sovereign state. The debates in sixteenth- and seventeenth-century Europe, during which the contemporary understanding of legitimacy and authority was formed, focused on the competing claims to power and authority by secular and ecclesiastical institutions. It was the late seventeenth century that saw the end of the struggle against the Christian church and the consolidation of the view that the secular state is the exclusive site of legitimate political authority (Skinner 1978; Tuck 1993).

The process of globalization is currently changing the terms of the debate. Of course, ideas such as a law of nations, a league of nations, or even a global state are not new (see e.g. Kant's essay on "Perpetual Peace," AK 8: 341 and his "Doctrine of Right," AK 6: 203). What is new is that international and global institutions take on more and more global governance functions and, at least in certain policy areas, step into the role formerly played by national political agencies. Allen Buchanan and Robert Keohane, for example, argue that global governance institutions such as the WTO or the IMF "are like governments in that they issue rules and publicly attach significant consequences to compliance or failure to comply with them—and claim authority to do so" (Buchanan and Keohane 2006: 406). And Joshua Cohen and Charles Sabel argue that

> [t]o a substantial and growing extent … rulemaking directly affecting the freedom of action of individuals, firms, and nation states (and the making of rules to regulate this rulemaking) is taking place … in global settings created by the world's nations but no longer under their effective control.
>
> (Cohen and Sabel 2006: 765)

With reference to such developments, many challenge the view that there is something unique about the coercive power exercised by nation states or about the political authority of the state. This once again raises the question of how to conceptualize legitimate authority in the face of competing claims to power and authority.

While some authors defend statism as—at least instrumentally—necessary for global justice (e.g. Ypi 2008; Christiano 2008b), most authors argue for the need to rethink or even replace the state-centered approach to legitimacy. In the context of international

legitimacy, the state-centered approach treats states in the international context in an analogous way as it treats persons in the national context (Beitz 1979: 408). Person-centered approaches, by contrast, take features of individuals—their interests or their rights—as the normative anchor of legitimate authority, whether it is in the national, international, or global realm. Buchanan probably has the most comprehensive treatment to date of a person-centered approach to political legitimacy in the international realm (Buchanan 2003). Buchanan defends a moralized conception of political legitimacy, according to which entities are legitimate if they are morally justified to wield coercive political power. Specifically, Buchanan argues that entities are legitimate if they are minimally just, in the following sense: the entity did not come about through usurpation, it meets certain standards for how to interact with others, and, internally, it meets a basic human rights standard.

Human rights are one of the most important ingredients of the emerging world order (Raz 2010). Human rights increasingly set binding standards for both domestic political institutions and decisions, for international relations, and for the operation of global governance institutions. As such, human rights are an important element in a people-centered standard for global legitimacy. There is currently quite some controversy about what constitutes the best understanding of human rights. John Tasioulas (2009) has a good overview of this debate. On what is called the traditional conception, human rights are a subset of moral rights—they are natural rights that all individuals have simply in virtue of their humanity and justified by ordinary moral reasoning. On the competing political conception, human rights, as we understand them today, originated in the specific political-legal project that was launched post World War II and they are justified by public reason. This debate is of some significance for the question of what constitutes legitimate political authority in the global realm (Cohen 2008; Besson and Tasioulas 2010). Do human rights primarily stand for a type of moral authority that acts as a corrective force to the political authority of nation states? Or do human rights institutions constitute a new form of global political authority that is potentially in tension with the authority of nation states?

Related Topics

Contractualism and Political Liberalism, Human Rights and Cosmopolitanism, Global Justice and Politics, Equality, Freedom, Power, Democracy, Religion in Public Life

References

Applbaum, A. I. (2010) "Legitimacy without the Duty to Obey," *Philosophy and Public Affairs* 38(3): 215–239.

Beitz, C. (1979) *Political Theory and International Relations*, Princeton, NJ: Princeton University Press.

Bentham, J. (1987 [1843]) "Anarchical Fallacies," in J.Waldron (ed.) *Nonsense upon Stilts*, London: Taylor & Francis.

Besson, S. and J. Tasioulas (eds.) (2010) *The Philosophy of International Law*, Oxford: Oxford University Press.

Buchanan, A. (2003) *Justice, Legitimacy and Self-Determination*, Oxford: Oxford University Press.

—— and R. O. Keohane (2006) "The Legitimacy of Global Governance Institutions," *Ethics and International Affairs* 20(4): 405–437.

Christiano, T. (2008a) *The Constitution of Equality*, Oxford: Oxford University Press.

—— (2008b) "Immigration, Political Community, and Cosmopolitanism," *San Diego Law Review* 45: 933–962.

Cohen, J. (2008) "Rethinking Human Rights, Democracy, and Sovereignty in the Age of Globalization," *Political Theory* 36(4): 578–606.
—— (1997a) "Deliberation and Democratic Legitimacy," in J. Bohman and W. Rehg (eds.) *Deliberative Democracy*, Cambridge, MA: MIT Press, pp. 67–91.
—— (1997b) "Procedure and Substance in Deliberative Democracy," in J. Bohman and W. Rehg (eds.) *Deliberative Democracy*, Cambridge, MA: MIT Press, pp. 407–437. Reprinted from S. Benhabib (ed.) *Democracy and Difference*, Princeton, NJ: Princeton University Press, pp. 95–119.
—— and C. Sabel (2006) "Global Democracy?," *New York University Journal of International Law and Politics* 37(4): 763–797.
D'Agostino, F. (1996) *Free Public Reason*, Oxford: Oxford University Press.
Dworkin, R. (1986) *Law's Empire*, Cambridge, MA: Harvard University Press.
Edmundson, W. A. (1998) *Three Anarchical Fallacies*, Cambridge: Cambridge University Press.
—— (2010) "Political Authority, Moral Powers and the Intrinsic Value of Obedience," *Oxford Journal of Legal Studies* 30(1): 179–191.
Estlund, D. (2008) *Democratic Authority*, Princeton, NJ: Princeton University Press.
Filmer, R. (1991 [1680]) *Filmer: "Patriarcha" and Other Writings*, J. P. Sommerville (ed.), Cambridge: Cambridge University Press.
Flikschuh, K. (2008) "Reason, Right, and Revolution: Kant and Locke," *Philosophy and Public Affairs* 36(4): 375–404.
Gaus, G. F. (1996) *Justificatory Liberalism*, Oxford: Oxford University Press.
—— (1997) "Reason, Justification, and Consensus: Why Democracy Can't Have It All," in J. Bohman and W. Rehg (eds.) *Deliberative Democracy*, Cambridge, MA: MIT Press, pp. 205–242.
—— (2011) *The Order of Public Reason*, Cambridge: Cambridge University Press.
Gilbert, M. (1996) *Living Together*, Boston, MA: Rowman & Littlefield.
—— (2006) *A Theory of Political Obligation*, Oxford: Oxford University Press.
Goldman, A. (1999) *Knowledge in a Social World*, Oxford: Oxford University Press.
Green, L. (1988) *The Authority of the State*, Oxford: Oxford University Press.
Habermas, J. (1996) *Between Facts and Norms*, W. Rehg (transl.), Cambridge, MA: MIT Press.
Hershovitz, S. (2003) "Legitimacy, Democracy, and Razian Authority," *Legal Theory* 9: 201–220.
Hobbes, T. (1994 [1668]) *Leviathan*, E. Curley (ed.), Indianapolis: Hackett.
Hume, D. (1965 [1748]) "Of the Original Contract," in *Hume's Ethical Writings*, A. MacIntyre (ed.), London: University of Notre Dame Press.
Kant, I. (1999) *Practical Philosophy. Cambridge Edition of the Works of Immanuel Kant in Translation*, M. J. Gregor (ed.), Cambridge: Cambridge University Press.
Knight, J. and J. Johnson (1994) "Aggregation and Deliberation: On the Possibility of Democratic Legitimacy," *Political Theory* 22(2): 277–296.
List, C. (2006) "The Discursive Dilemma and Public Reason," *Ethics* 116(2): 362–402.
Locke, J. (1990 [1690]) *Second Treatise on Civil Government*, C. B. MacPherson (ed.), Indianapolis: Hackett.
Manin, B. (1987) "On Legitimacy and Political Deliberation," *Political Theory* 15: 338–368.
Mill, J. St. (1998) *On Liberty and Other Essays*, J. Gray (ed.), Oxford: Oxford University Press.
Nozick, R. (1974). *Anarchy, State, and Utopia*, Oxford: Blackwell.
Peter, F. (2008) *Democratic Legitimacy*, New York: Routledge.
Pettit, P. (2001) "Deliberative Democracy and the Discursive Dilemma," *Philosophical Issues* 11: 268–299.
Quong, J. (2004) "The Scope of Public Reason," *Political Studies* 52(2): 233–250.
Rawls, J. (1971) *A Theory of Justice*, Cambridge, MA: Harvard University Press.
Rawls, J. (1993) *Political Liberalism*, New York: Columbia University Press.
—— (1995) "Reply to Habermas," *The Journal of Philosophy* 92(3): 132–180.
—— (1999) *The Law of Peoples*, Cambridge, MA: Harvard University Press.
—— (2001) *Justice as Fairness: A Restatement*, Cambridge, MA: Harvard University Press.
Raz, J. (1988) *The Morality of Freedom*, Oxford: Oxford University Press.
—— (1995) *Ethics in the Public Domain: Essays in the Morality of Law and Politics*, Oxford: Clarendon Press.
—— (2006) "The Problem of Authority: Revisiting the Service Conception," *Minnesota Law Review* 90: 1003–1044.
—— (2010) "Human Rights in the Emerging World Order," *Transnational Legal Theory* 1: 31–47.
Reidy, D. (2000) "Rawls's Wide View of Public Reason: Not Wide Enough," *Res Publica* 6(1): 49–72.
Ripstein, A. (2004) "Authority and Coercion," *Philosophy and Public Affairs* 32(1): 2–35.

—— (2009) *Force and Freedom*, Cambridge, MA: Harvard University Press.
Simmons, A. J. (2001) *Justification and Legitimacy: Essays on Rights and Obligations*, Cambridge: Cambridge University Press.
Skinner, Q. (1978) *The Foundations of Modern Political Thought*, 2 vols, Cambridge: Cambridge University Press.
Tasioulas, J. (2009) "Are Human Rights Essentially Triggers for Intervention?" *Philosophy Compass* 4(6): 938–950.
Tuck, R. (1993) *Philosophy and Government, 1572–1651*, Cambridge: Cambridge University Press.
Waldron, J. (1987) "Theoretical Foundations of Liberalism," *The Philosophical Quarterly* 37(147): 127–150.
Weber, M. (1964) *The Theory of Social and Economic Organization*, Basingstoke: Macmillan.
Weithman, P. (2011) *Why Political Liberalism? On John Rawls's Political Turn*, Oxford: Oxford University Press.
Wellman, C. (1996) "Liberalism, Samaritanism, and Political Legitimacy," *Philosophy and Public Affairs* 25(3): 211–237.
—— (1997) "Associative Allegiances and Political Obligations," *Social Theory and Practice* 23: 181–204.
Wolff, R. P. (1970) *In Defense of Anarchism*, New York: Harper & Row.
Ypi, L. (2008) "Statist Cosmopolitanism," *Journal of Political Philosophy* 16(1): 48–71.

53
DEMOCRACY

Robert B. Talisse

Ann, Betty, and Cara wish to go to the movies together tonight. Their local cinema is showing two films; one is a bittersweet romantic comedy, and the other is an action-packed thriller. They must choose which one to see. How should they choose?

Suppose that Ann, Betty, and Cara are all indifferent. They want to see a film *together*, and it doesn't matter to any of them much which one. In that case, they could simply flip a coin, or employ some other strategy for picking one of the two films, and get on with it. But let's imagine that they are not indifferent. That is, let's imagine that for each member of the group, there is a movie among the two on offer that she most wants to see tonight. Of course, it could turn out that Ann, Betty, and Cara all agree on what movie to see; but let's suppose the group is not of one mind about what film to see. How should they decide? There's something natural about the thought that they should each state their preferred movie, and the group should go to whichever film the majority wants to see.

1. The Schoolyard View of Democracy

This is the simple view of democracy that we all learned as children. Call it the *schoolyard view* of democracy. The schoolyard view looks like a simple kind of *democracy* mainly because it embodies two norms that are most frequently associated with democracy in its full-blown variety, namely, (1) equality and (2) majority rule. In our example, the schoolyard view instantiates equality in that Ann, Betty, and Cara each get exactly one vote regarding which movie to see; it manifests majority rule in that the group goes to the movie that attracts the most votes. Putting these together, we say that the schoolyard view seeks to arrive at collective decisions in a way that treats each individual as an equal. That's democracy in a nutshell.

But the schoolyard view is in many ways simplistic. Indeed, when even slight complexities are introduced, it loses much of its intuitive force. For example, suppose Ann *really* wants to see the romantic comedy and really *dislikes* thrillers, whereas Betty and Cara have a slight preference in favor of the thriller, but are not opposed to the romantic comedy. Or consider the possibility that Ann's preference for the romantic comedy is the result of her having read several reviews by reliable critics, whereas Betty's and Cara's preference for the thriller is not informed by reviews or anything much at all. Further, imagine that Ann's preference for the romantic comedy is based on reliable reviews, and expresses her educated judgment concerning which film Betty and Cara are more likely to enjoy. Or imagine a case in which Ann strongly prefers the romantic comedy, and Betty strongly prefers the thriller, but Cara, who is generally indifferent,

votes for the thriller because she knows that Betty, unlike Ann, will get angry if she votes against the thriller. Finally, consider the case in which Ann realizes that both Betty and Cara are in some significant way *misinformed* about the thriller; they mistakenly believe that it stars their favorite actor and has an uplifting and patriotic message, while Ann knows that it stars an unknown actor and ends with a bleak affirmation of existential dread.

Things grow only more difficult when we consider even slightly larger groups with greater numbers of options. For example, what if the group discovers that the local cinema is showing *four* films this evening: a romantic comedy, a thriller, a mystery, and a documentary? What if Debbie and Emma join the group? Suppose that Debbie and Emma both most want to see the documentary, whereas Ann votes for the romantic comedy, Betty votes for the thriller, and Cara votes for the mystery. The group apparently should go to the documentary, even though the majority voted for some other option. What steps should be taken in order to avoid this kind of result? Perhaps some scheme should be devised by which each member of the group gets to vote for a particular *ordering* of films (e.g., mystery first; thriller second; and so on). But once such a system is introduced, we confront the difficult question of how the votes should be counted. Perhaps, alternatively, the group should nominate someone (or some sub-set) among them to be the decider with regard to films? But how should *that* decision be made? And upon what should the decider's decisions be based: her own preferences, her own judgments about which films are good, her judgments about which films the group would choose, or her own judgments about which films that group would like best? Additional difficulties of this kind could be added easily.

The point is that in any of these cases, it is not obvious that the simple democracy of the schoolyard is an appropriate way to make a collective decision. It is important to note that in the cases described above, the schoolyard view begins to look dubious not merely because it appears to be an inefficient way of making a good collective decision; rather, it seems inadequate from the point of view of equality. We are led to ask what, precisely, equality requires. Does the *one person, one vote* principle of the schoolyard view sufficiently model the equality of all the members of the group? Or must a process take into account additional factors, such as the *intensity* or the *order* of individuals' preferences, or the *informedness* or *rationality* of individuals' votes, in order to treat members of the group as equals?

2. Aggregative Views of Democracy

For most of the twentieth century, democratic theory was dominated by a family of views that can be generally characterized as *aggregative*. The aggregative view is, in essence, the schoolyard view of democracy writ large. Each political actor gets exactly one vote, each actor is expected to vote in accordance with his or her preference, each vote is given equal weight, and the majority rules. The principle guiding the view is that when interests collide and yet a collective decision must be made, we should follow a procedure that affords to each actor *equal input*, and aggregates the inputs into a collective decision. Insofar as a collective decision is produced by a fair process that aggregates the preferences of all, it can be said to represent the popular will or common good.

The aggregative view inherits all the difficulties of the schoolyard view. As we saw in the examples above, the problem quickly becomes "how to do the relevant math" (Shapiro 2003: 10), how to aggregate the inputs such that the result could be regarded

as a genuine expression of the majority will. By the middle of the twentieth century, this problem was shown to be insurmountable; famous results associated with the social choice theorist Kenneth Arrow (1951) demonstrate that there is no way to transform expressions of individual preference into a rational collective decision that could plausibly claim to represent the popular will. The details of the social choice challenge to aggregative views cannot be presented here. Suffice it to say that many hold that Arrow showed that, given commonsense views about fairness and rationality, there is no way to aggregate individual preferences into an acceptable collective decision.

3. Minimalist Views of Democracy

From these developments, many concluded that the very idea of a popular will or a common good is nonsensical or "empty" (Riker 1982: 239). Hence a range of *minimalist* views developed which retained the aggregative view's conception of the democratic process while jettisoning the idea that collective decision must represent the common good. One such view is the "Madisonian" view proposed by William Riker (1982: 9). Riker holds that democratic elections exist solely for the purposes of "restraining officials" (1982: 12) and enabling people to "get rid of rulers" (1982: 244) who have "offended" them (1982: 242). On Riker's view, then, democracy is "not popular rule, but rather an intermittent, sometimes random, even perverse, popular veto" (1982: 244).

The closely related *elitist* conception associated with Joseph Schumpeter (1962)—and, more recently, Richard Posner (2003)—holds that democracy is "that institutional arrangement for arriving at political decisions in which individuals acquire the power to decide by means of a competitive struggle for the people's vote" (Schumpeter 1962: 269). The elitist hence sees democracy as means for *producing* government by means of elections; citizens select "bosses" (1962: 269) who in turn must maintain a sufficient level of popularity or else face the "punishment" of losing the next election (Posner 2003: 182).

A third kind of minimalist view, *interest-group pluralism* holds with the elitist view that democracy is a market-like arrangement by which aspiring officials compete for votes (Dahl 1956: 3). The pluralist, however, introduces a view of power that is more nuanced than that of the elitist. The elitist sees political power as fundamentally residing in the offices of government; the pluralist, by contrast, sees power as distributed throughout society, especially in the various kinds of groups formed by individuals with shared interests. On the pluralist view, democracy is a constant negotiation among multiple interest groups, with competing groups exerting pressure on elected officials well beyond Election Day. The pluralist, that is, rejects the view that democracy is rule by elected elites, holding instead that democracy is rule by competing minorities (Dahl 1956: 133).

Despite their differences, minimalist views agree that the value of democracy is purely instrumental: democracy is the best means for maintaining social stability amidst masses of people with competing and conflicting interests. For the minimalist, then, the task of democratic theory is mainly descriptive rather than normative. The aim is to discern which institutions and rules best promote social stability. Consequently, these views do not speak to the familiar normative questions one may ask about democracy concerning the basis for the collective bindingness of democratic outcomes, the relation of democracy to justice, or the consistency of democracy with seemingly anti-majoritarian institutions such as judicial review. Indeed, some minimalist views deny that such normative questions even make sense.

The normative questions that the descriptive views had abandoned were revived in light of the political and social movements of the 1960s. Must one obey racist laws? Must one risk one's life in an unjust war if drafted? Is democracy best understood as the "dictatorship of the proletariat?" Why should one submit to rulers at all? That a leading democratic theorist at the time could assert that "political apathy may reflect the health of a democracy" (Lipset 1963: 32) is indicative of one of minimalism's principal shortcomings. Minimalism's strictly descriptive aspiration left it unable to speak to the political activity of the day, much of which was couched in the overtly normative language of justice, freedom, and equality. Furthermore, the claim that minimalism was non-normative became increasingly suspect. For the minimalists indeed adopt certain minimal normative commitments, such as that the rulers who "offend" the people do not deserve nonetheless to rule and that stability is a good. Lastly, minimalist views seem self-defeating. They ultimately rely on popular voting to select who shall become "boss" and who shall be "punished"; however, if the social choice results show that there is no "popular will," then there is no sense in which a vote could reveal popular dissatisfaction with a given official or policy. Put otherwise, if the social choice results are taken to show that the aspiration to discern a "popular will" and "common good" by means of the aggregation of individual votes is nonsense, then the idea that voting expresses a meaningful "veto" must also be discarded (Coleman and Ferejohn 1986; Cohen 1986: 30; Miller 1992).

4. Participatory and Deliberative Democracy

These dissatisfactions led in the 1970s and 1980s to the development of views that shifted the focus of democratic theory away from voting and towards other facets of democratic politics. To be more specific, there developed conceptions of democracy which regard the *one person, one vote* principle as insufficient to capture the kind of equality that democracy must manifest if its outcomes are to be collectively binding. These views can be classified into two general categories: *participatory democracy* and *deliberative democracy*. It is safe to say that deliberative democracy is the prevalent framework for contemporary democratic theory; even views that do not claim explicitly to be deliberativist emphasize the importance of public discussion and argument for democracy. For this reason, my discussion of participatory views will be brief, and I will focus considerably more attention on the deliberativist views.

According to *participatory democracy*, democratic equality is essentially tied to the ability of each individual to *participate* in public life *as an equal* citizen. This kind of equality differs from the *equal input* conception of the aggregative and minimalist views in that it sees political equality as a kind of *status*, namely, that of being a citizen. For the participationist, citizenship is a distinctive kind of office. A citizen is not simply a private individual acting in public on behalf of his or her preferences; rather, citizenship is understood as the activity of collective self-government.

Participatory democrats draw inspiration from Jean-Jacques Rousseau, who, at least on common interpretations, envisioned democracy as a process by which individual citizens put aside their private interests and judgments and adopted the perspective of the "general will" or common good. Hence democracy is not understood as a means of conflict management among individuals with conflicting interests, but rather as a social ideal which aims at the "transformation" of the "consciousness" of individual citizens (Pateman 1985: 156). When citizens participate in democracy, "*me* language"

is transformed into "*we* language" (Barber 1998a: 13); the public recaptures its "public voice" and "speaks in terms that reveal and elicit common ground, cooperative strategies, overlapping interests, and a sense of the public weal" (Barber 1998b: 116).

The participationists, then, see democracy as aimed at forging solidarity, consensus, community, and a sense of belonging among equal citizens. Hence the emphasis is moved away from the formal mechanisms of collective decision and towards the institutions and norms of civil society, where citizens interact and publicly engage matters of common concern.

Yet the participatory model seems to overlook, or at least downplay, the fact that democracy is always in large part a process of *collective decision*, the very need for which presupposes that those who must decide together are not of one mind with respect to what the group ought to do. Some participatory democrats go so far as to claim that deep disagreements among citizens are always merely *apparent*, or rather the *effects* of an insufficiently participatory democratic order and its corresponding lack of an "intersubjective identity" among citizens (Barber 1998a: 9). But were such an identity achieved, there would be little need for politics, democratic or otherwise. Participatory democracy, it seems, gets things precisely backwards.

Beginning in the 1980s, democratic theory took what is aptly characterized as a "deliberative turn" (Dryzek 2000: 1; Goodin 2008: 2). According to *deliberative democracy*, democratic equality requires democratic outcomes to in some way give equal consideration to the reasons citizens offer for their favored political positions. At present there is a wide variety of views which claim the label *deliberative democracy* (cf. Bohman 1998 and Freeman 2000), and nearly all contemporary conceptions of democracy invoke some aspect of deliberation.

All versions of deliberative democracy hold that deliberation contributes to democratic legitimacy, that deliberation is a central element of the collective bindingness of democratic decisions (Cohen 2009: 21; Benhabib 1996: 69; Gutmann and Thompson 2004: 10; Dryzek 2000: 1). Unsurprisingly, this core commitment is interpreted differently among deliberative theorists, depending upon their more general orientation within democratic theory. One way to get a handle on the various versions of deliberative democracy, then, is to categorize them according to the three broad programs within contemporary democratic theory: liberalism, civic republicanism, and radical democracy. It should be stressed that these categories are loose and non-exclusive. They should be taken as different *approaches*; the point is to convey the flavor of the views in currency.

5. Liberal Deliberativism

Liberal deliberativism follows principally from the later work of John Rawls (2005 [1993]), and among its main exponents are Joshua Cohen (2009) and Amy Gutmann and Dennis Thompson (1996, 2004). As it is a variety of liberal democratic theory, liberal deliberativism shares with non-deliberative liberal views the idea that a liberal society is a "fair system of cooperation among free and equal citizens" (Rawls 2005: 22). The question is how political authority and coercion can be consistent with the equality of all citizens.

According to liberal deliberativists, equality "requires more than that the interests of all be given equal consideration in binding collective decisions" (Cohen 2009: 243). Again, the *one person, one vote* principle is insufficient for equality. The liberal deliberativists hold that in order to satisfy the requirement of treating all citizens as equals, democratic

decisions must be in some way sensitive to the reasons citizens can advance in support of their favored policies; indeed, some theorists hold that properly designed processes of public deliberation are themselves expressions of political equality (Christiano 2008: 200). Accordingly, decision making must be preceded by episodes of open and inclusive public deliberation in which citizens attempt to justify to each other, by means of reasons accessible to all, their views concerning how political power should be exercised. In this way, those whose views lose out can still affirm the legitimacy of the outcome, because it was produced by an egalitarian process that is receptive to their reasons.

Of course, liberal deliberativists often propose this model of deliberative legitimation as an *ideal* of democratic politics. Cohen holds that democratic decisions are legitimate insofar as they would have been the outcome of an ideal deliberation among citizens; hence actual democratic politics should "mirror" ideal deliberation "as far as possible" (Cohen 2009: 23). This means that democratic governments must seek to "facilitate" and sustain "favorable conditions" for public deliberation among citizens (Cohen 2009: 224). Gutmann and Thompson claim that it is a "requirement" for legitimacy that "citizens or their representatives actually seek to give one another mutually acceptable reasons to justify the laws they adopt" (2004: 100). Some champion ambitious reconstructions of democracy, envisioning new institutions which aim to realize the deliberative ideal. Bruce Ackerman and James Fishkin (2003) propose a new national holiday for carefully structured nationwide deliberation events and "deliberative polls" designed to reveal how a well-informed populace would vote on a given issue (Fishkin 1997); others call for the formation of a fourth, "deliberative" branch of government (Leib 2004).

6. Civic Republicanism

Although civic republicanism has its roots ultimately in pre-modern sources (see Skinner 1998), contemporary civic republican deliberativists frequently draw inspiration from Rousseau, John Dewey, and Hannah Arendt. Philip Pettit (1995, 2003) has proposed the most sophisticated form of civic republican deliberativism, so we will focus on his views.

Pettit holds that citizens are free when they are not *dominated*, that is, subject to arbitrary interference from others (2003: 152). From this, the problem arises of how citizens can ward off the domination of the state. The state is, after all, a necessarily powerful entity that is charged with interfering with its citizens. How can its interference be kept non-arbitrary?

According to Pettit, the key to effectively constraining the state is popular contestation. This requires political conditions in which citizens "have access to the reasons" which drive state action, and can reasonably "expect to be heard" when they advance objections to those reasons (2003: 152). Republican liberty thus requires deliberative democracy. Pettit calls this the "contestability argument" for deliberative democracy (2003: 153). As Pettit notes, this argument dovetails nicely with considerations advanced by fellow republican, Cass Sunstein, who in a long list of related publications, has proposed the view that the US Constitution "was designed to create a deliberative democracy" (1993: 19), a "republic of reasons" (1993: 20). According to Sunstein, democratic government acts legitimately when it can offer reasons "that can be intelligible to different people operating from different premises" (1993: 24), and thus be held publicly accountable.

7. Radical Democracy

There is a group of deliberativist views which sees deliberative democracy as an *alternative* to liberal and civic republican democratic theory, rather than a supplement to them. I gather these views under the category of *radical democracy*, though the label is admittedly somewhat artificial.

Drawing often from varied sources in critical theory, feminist philosophy, and environmentalism, radical democratic deliberativists often begin from a critique of both liberal and civic republican views. The critique begins with the charge that liberal and republican versions of deliberative democracy are focused exclusively on the state and its actions, and consequently can operate "only on the surface of the political economy" (Dryzek 2000: 21). On radical views, the state and its apparatus are not the main sites of democratic politics; democracy resides primarily in *civil society* or the *public sphere*, arenas of citizen interactions outside of the state (Benhabib 1996: 75; Habermas 1996: 26). The radical democrats' objection is that liberal and republican views are unfit to address concerns regarding the ideological, economic, and structural inequalities which flourish in the public sphere of contemporary democratic societies (Benhabib 1996: 75); continuing, they charge that encouraging public deliberation under such conditions tends simply to further entrench the status quo and provide further advantage to agents who are already disproportionately powerful (Young 2001: 670). This general line of criticism encourages a closely related charge that, by focusing on the state, liberal and republican views abandon the aim of democratizing civil society, including especially the workplace and economic relations more generally (Dryzek 2000: 27; but see Pettit 2003: 153).

One way to capture the contrast between radical democratic deliberativism on the one hand and liberal and civic republican deliberativisms on the other is to say that the latter views appeal to deliberation for the purposes of legitimating state coercion and reaching consensus on state policy, whereas the former appeals to deliberation as a way of rendering society more *authentically* democratic (Dryzek 2000: 29). The authenticity of a democracy is measured by the degree to which the collective institutions of society are controlled by "autonomous and competent actors" (Dryzek 2000: 29), who are actively engaged in various forms of social "critique" of (Benhabib 1996: 87), and "struggle" against (Young 2000: 50), the status quo. Yet the aim of such activity is not the establishment of a new, perhaps more inclusive and just, status quo, but rather the continuation of the struggle by means of "democratic iterations," acts of challenge, resistance, criticism, and defiance designed to disrupt and transform the status quo (Benhabib 2006: 48). In this way, radical democratic deliberativists are often closely allied with post-Marxist agonists like Chantal Mouffe, who see democratic politics as perpetual agitation and dissent against the very aspiration of reaching a reasoned consensus about the terms of political association (Mouffe 1996: 254f.; cf. Mouffe 2000: ch. 4). Consequently, radical democratic deliberativism often joins with various modes of political activism (Young 2001), green political theory (Dryzek 2005), and global democracy (Benhabib 2006; Bohman 2007).

8. Deliberative Democracy Criticized

Let us turn next to a serious concern about deliberative democracy in general. In any of its versions, deliberative democracy requires citizens to engage in collective deliberation

with those with whom they disagree. This raises concerns regarding the social dynamics of face-to-face disagreement. Lynn Sanders has argued that the same patterns of gender and racial privilege which prevail in the broader society tend to replicate themselves in deliberative contexts (Sanders 1997: 370).

Sanders notes, for example, in jury deliberations, "men talk more" and are more likely to be selected to lead juries; moreover, in interracial deliberative groups, the contributions of white males tend to be more influential (1997: 363). This is explained in large part by the fact that deliberation is "a certain kind of talk" which naturally puts a premium on speakers' rhetorical prowess, including the ability to speak in ways that seem calm, eloquent, and authoritative, characteristics culturally associated with being white and male (Sanders 1997: 370). The challenge, then, is that deliberative democracy prizes modes of social engagement that tend to systematically disenfranchise, exclude, or disadvantage persons of color, women, immigrants, the uneducated, the poor, and the working-class. Deliberative democracy seems to be simply the rule of political elites, because deliberation as understood by most deliberative democrats is the activity of political elites (Young 2001: 677).

To meet this challenge, some deliberativists have proposed broadened conceptions of deliberation that include modes of communication that do not focus on reason giving. Iris Young has moved the furthest in this direction; she contends that properly democratic communication aims primarily at "understanding" across differences of gender, class, race, and ethnicity rather than the justification of coercion (1996: 127). Accordingly, Young proposes a model of deliberation in which reason giving is joined by greeting, rhetoric, and storytelling (Young 1996: 129f.; 2000: ch. 2); she also holds that democratic deliberation must be understood to include acts that are "rowdy" and "disorderly" such as "street demonstrations and sit-ins, musical works and cartoons" (Young 2001: 688). She contends that a model of deliberation that includes these non-rational styles of communication is necessary in order to counteract exclusionary deliberative norms.

There are several difficulties with Young's proposal. First is that Young provides no reason for thinking that her additional components of deliberation—greeting, rhetoric, and storytelling—are free from the distorting influences of social privilege. Why should we think that Young's broadened conception of deliberation is any less prone to distortion than the more usual conceptions? Indeed, it seems likely that the same hierarchies will surface once reason giving is supplemented with non-rational modes of communication (Dryzek 2000: 67).

Young concedes in later work that storytelling, rhetoric, and greeting might serve significant functions in helping to establish conditions favorable to mutually respectful and inclusive argument, but they cannot serve as substitutes for argument (2000: 79). Yet if she holds that storytelling, rhetoric, and greeting are *supplements* to deliberation rather than deliberation itself, Young proposes nothing that deliberative democrats deny (Dryzek 2000: 68). More importantly, she has not addressed Sanders' concerns. Even when aided by the kinds of non-rational communication Young identifies, deliberation is still focused on reason giving, and, if Sanders is correct, deliberation is subject to systematic distortions. The lesson, then, is that deliberative democracy is still in need of a theory of political action, especially in real-world contexts where the deliberativist ideal of egalitarian, respectful, and inclusive public reasoning will never be realized.

David Estlund suggests a way forward. Estlund begins by noting that deliberation includes norms of calmness, orderliness, and, in a word, civility because the point of

deliberation is to block "power's interference with reason" (2008: 193). Accordingly, the idea of civil collective reasoning can be upheld as an *ideal* without being taken as a prescription for how citizens must act in real-world contexts where power does indeed interfere with reason. To explain: Estlund holds that as an *ideal* the deliberativist image of civil collective reasoning can be used to identify "breakdowns" in real-world contexts (2008: 200). (After all, it is precisely by reference to such an ideal that Sanders is able to identify the ways in which real-world deliberative contexts manifest objectionable tendencies.) However, in real-world contexts, where power is unequally distributed and allowed influence in deliberative contexts, the deliberativists' ideal of civility proposes inappropriate constraints on political action. Put otherwise, deliberativists propose the norms of civility because they are concerned to block power's interference with reason; so in contexts where power does interfere with reason, those norms are canceled.

This does not mean, however, that in real-world contexts anything goes (2008: 192). Estlund argues that in real-world contexts, a more permissive set of norms prevails; more specifically, real-world deliberation will often permit exercises of "countervailing power" designed to "remedy epistemic distortion wrought by the initial insertion of power" (2008: 193–194). These norms may allow "sharp, disruptive, and even suppressive" tactics (2008: 185), including "conscientious suppression of an overrepresented message … or the introduction, through sharp transgressive methods, of an underrepresented message" (2008: 194). It seems that these more permissive norms would also allow the various forms of political activism and public protest identified by Young (2001). Accordingly, we need not abandon or reject the deliberative ideal in order to accommodate the concerns identified by Sanders.

Not all deliberativists share Estlund's strongly epistemic orientation. However, nearly all deliberativists make *some* appeal to the epistemic benefits of deliberation. Deliberation is said to make decisions more "justifiable" (Gutmann and Thompson 2004: 21; Christiano 2008: 193), "rational" (Benhabib 1996: 71; Dryzek 2000: 174), "wise" (Young 2000: 30), and "intelligent" (Anderson 2006: 13); and some argue for deliberative democracy on the grounds that it is the political manifestation of the epistemic norms that should govern our cognitive lives as such (Misak 2004: 15; Talisse 2009). But once it is admitted that democratic deliberation has this epistemic function, we must introduce additional epistemic elements into the conception of deliberation (Cohen 2009: ch. 11).

Yet there has been relatively little attention paid to the epistemological issues that the deliberativist program raises. Perhaps this is due to the fact that when we turn to the empirical literature regarding the public's epistemic abilities, things look bleak. As one might expect, deliberative groups consisting of average citizens are epistemically a mixed-bag. Hence it seems that, in light of the deliberative turn, it is necessary to bring deliberative democratic theory into closer contact with the emerging fields of social epistemology and argumentation theory. A handful of theorists (Gaus 1996, 2011; Sunstein 2006; Estlund 2008; Goodin 2008; Talisse 2009) have begun this work, but there remains ample room for further development.

Related Topics

Liberalism, Equality, Freedom, Authority and Legitimacy, Social Choice Theory, Discourse Theory

References

Ackerman, B. and J. Fishkin (2003) *Deliberation Day*, New Haven, CT: Yale University Press.
Anderson, E. (2006) "The Epistemology of Democracy," *Episteme* 3: 9–23.
Arrow, K. (1951) "A Difficulty in Social Welfare," *Journal of Political Economy* 68: 328–346.
Barber, B. (1998a) *A Passion for Democracy*, Princeton, NJ: Princeton University Press.
—— (1998b) *A Place for Us*, New York: Hill and Wang.
Benhabib, S. (1996) "Toward a Deliberative Model of Democratic Legitimacy," in S. Benhabib, ed., *Democracy and Difference*, Princeton, NJ: Princeton University Press.
—— (2006) *Another Cosmopolitanism*, New York: Oxford University Press.
Bohman, J. (1998) "The Coming of Age of Deliberative Democracy," *Journal of Political Philosophy* 4: 418–443.
—— (2007) *Democracy across Borders*, Cambridge, MA: MIT Press.
Christiano, T. (2008) *The Constitution of Equality*, New York: Oxford University Press.
Cohen, J. (1986) "An Epistemic Conception of Democracy," *Ethics* 97(1): 26–38.
—— (2009) *Philosophy, Politics, Democracy*, Cambridge, MA: Harvard University Press.
Coleman, J. and J. Ferejohn (1986) "Democracy and Social Choice," *Ethics* 97: 6–25.
Dahl, R. (1956) *A Preface to Democratic Theory*, Chicago, IL: University of Chicago Press.
Dryzek, J. (2000) *Deliberative Democracy and Beyond*, New York: Oxford University Press.
—— (2005) *The Politics of the Earth*, New York: Oxford University Press.
Estlund, D. (2008) *Democratic Authority*, Princeton, NJ: Princeton University Press.
Fishkin, J. (1997) *The Voice of the People*, New Haven, CT: Yale University Press.
Freeman, S. (2000) "Deliberative Democracy: A Sympathetic Comment," *Philosophy and Public Affairs* 29: 371–418.
Gaus, G. (1996) *Justificatory Liberalism*, New York: Oxford University Press.
—— (2011) *The Order of Public Reason*, Cambridge: Cambridge University Press.
Goodin, R. (2008) *Innovating Democracy*, New York: Oxford University Press.
Gutmann, A. and D. Thompson (1996) *Democracy and Deliberation*, Cambridge, MA: Harvard University Press.
—— (2004) *Why Deliberative Democracy?* Princeton, NJ: Princeton University Press.
Habermas, J. (1996) "Three Normative Models of Democracy," in S. Benhabib, ed., *Democracy and Difference*, Princeton, NJ: Princeton University Press.
Leib, E. J. (2004) *Deliberative Democracy in America*, College Park, PA: Penn State University Press.
Lipset, S. M. (1963) *Political Man*, New York: Doubleday.
Miller, D. (1992) "Deliberative Democracy and Social Choice," *Political Studies* 40: 54–67.
Misak, C. (2004) "Making Disagreement Matter," *Journal of Speculative Philosophy* 18(1): 9–22.
Mouffe, C. (1996) "Democracy, Power, and 'The Political'," in S. Benhabib, ed., *Democracy and Difference*, Princeton, NJ: Princeton University Press.
—— (2000) *The Democratic Paradox*, New York: Verso.
Pateman, C. (1985) *The Problem of Political Obligation*, Berkeley: University of California Press.
Pettit, P. (1995) *Republicanism*, New York: Oxford University Press.
—— (2003) "Deliberative Democracy, the Discursive Dilemma, and Republican Theory," in James Fishkin and Peter Laslett, eds., *Debating Deliberative Democracy*, Oxford: Blackwell.
Posner, R. (2003) *Law, Pragmatism, and Democracy*, Cambridge, MA: Harvard University Press.
Rawls, J. (2005) *Political Liberalism*, expanded edition, New York: Columbia University Press.
Riker, W. (1982) *Liberalism Against Populism*, New York: W. H. Freeman and Co.
Sanders, L. (1997) "Against Deliberation," *Political Theory* 25: 347–377.
Schumpeter, J. (1962) *Capitalism, Socialism, and Democracy*, New York: Harper Books.
Shapiro, I. (2003) *The State of Democratic Theory*, Princeton, NJ: Princeton University Press.
Skinner, Q. (1998) *Liberty before Liberalism*, Cambridge: Cambridge University Press.
Sunstein, C. (1993) *The Partial Constitution*, Cambridge, MA: Harvard University Press.
—— (2006) *Infotopia*, New York: Oxford University Press.
Talisse, R. B. (2009) *Democracy and Moral Conflict*, Cambridge: Cambridge University Press.
Young, I. M. (1996) "Communication and the Other: Beyond Deliberative Democracy," in S. Benhabib, ed., *Democracy and Difference*, Princeton, NJ: Princeton University Press.
—— (2000) *Inclusion and Democracy*, New York: Oxford University Press.
—— (2001) "Activist Challenges to Deliberative Democracy," *Political Theory* 29: 670–690.

54
RIGHTS

Jonathan Quong

The language of rights is pervasive in moral, political and legal debates, but there is no clear consensus on what rights are and how they work. Do all rights have the same underlying function? Should a theory of rights be conceptual or normative? Can rights conflict with each other? I cannot tackle these questions in sufficient depth here, and I have to ignore many other interesting questions altogether. I will try, however, to provide a roadmap to a few of the central debates about rights—hopefully offering a clear picture of both what's at stake in these debates and why philosophers have developed starkly different answers to the questions above.

1. A Framework for Rights

Wesley Hohfeld famously showed that rights can be distinguished into four different types or incidents (Hohfeld 1919). Hohfeld focused on legal rights, but his framework applies to moral and political rights as well. Below I explain Hohfeld's typology, though some of the further terms I use are not Hohfeld's. (For further expositions of Hohfeld see Kramer (1998), Thomson (1990: ch. 1), and Wenar (2005).)

First, and most centrally, rights can refer to *claims*. Betty might have a claim that Albert not enter her room, which is another way of saying that Albert is under a duty to Betty not to enter her room. Claims thus always correlate with duties—whenever one person has a claim, at least one other person owes a duty to the claim-holder. (There may be other kinds of duties that have nothing to do with individual claims, but they are not our focus.) Claims are the most familiar form of right. For example, Albert's moral claims not to be murdered, raped, and assaulted mean that every other person is under a duty not to murder, rape, or assault Albert. If someone breaches one of these duties then that person has failed to uphold a claim of Albert's and, in doing so, that person has *wronged* Albert. The opposite of having a claim is having a *no-claim*.

Now consider the following assertion: "I have every right to try to persuade Carl to change his mind." The imagined speaker need not be presenting himself as possessing a claim and thereby insisting that someone else is under a duty. Instead, the speaker can plausibly be understood to be declaring that he himself is under *no duty* not to try to use persuasion to change Carl's mind. To be under no duty not to perform some action, X, is to have a *liberty or privilege* to do X. (I will use the term "liberty.") This is sometimes what people mean when they assert a right: they are at liberty to perform the action in question since no one else has a claim against them performing the action.

Although claims and liberties often travel together, it's important not to confuse them. For example, in boxing matches each boxer has a liberty to try to punch the

other: each is under no duty not to do this. But each boxer also has a liberty to try to stop the other from punching him. Liberties thus don't necessarily confer claims—sometimes we are at liberty to try to do something, but other people are also at liberty to try to stop us from doing that thing. Of course liberties are often accompanied by claims; for example, when we say that Albert has a right to free speech, this usually means that Albert has both a liberty to say whatever he likes (no duty not to do so) but also that Albert has certain claims, for instance, claims that entail government officials are under a duty not to prevent or censor Albert's speech. We thus often apply the term "right" in ordinary discourse to what is in fact a cluster or conjunction of Hohfeldian claims and liberties (Thomson 1990: 54–6).

Claims and liberties cover "first order" assertions of rights—they account for the cases where our rights talk refers to what duties individuals are, or are not, under with regard to other individuals' claims. But there is also a "second order" of rights discourse: assertions about how we can alter our first order claims and liberties. Consider this statement: "Betty has the right to determine who can enter her apartment." If this statement is true, then Betty possesses what Hohfeld calls a *power*—the authority to alter the claims or liberties of others, in this case, the liberty of others to enter Betty's apartment. Other examples of powers include: the power to declare two people legally married; the power to transfer ownership of a car to someone else; and the power of a police officer to detain a suspect. To lack a particular power is to have a *disability*.

When someone lacks the power to alter some other person's claims or liberties, then the latter has what Hohfeld calls an *immunity*. For example: "The government has no right to prohibit same-sex couples from adopting children." If this statement is true, then same-sex couples possess an immunity with regard to the government; the government lacks the power to affect their right to adopt children. It's widely believed that individuals possess immunities with regard to a number of important claims; for instance, it's widely believed that no one else could revoke or modify my claims against being murdered, raped, or tortured. To lack an immunity is to have a *liability*.

In summary, there are four Hohfeldian incidents, each of which has a correlative and an opposite (see Figure 1).

This Hohfeldian framework is valuable because it provides a precise set of conceptual tools for thinking about rights. If Albert says, "I have a right to kill my attacker in self-defense," we want to know exactly what this alleged right involves. Does it simply mean Albert has a liberty (no duty not to) to kill his attacker in self-defense? Does it mean, more strongly, that Albert has a claim that others not interfere with him when he attempts to kill his attacker in self-defense (others owe him a duty of non-interference)?

Hohfeldian Type	**Correlative**	**Opposite**
Claim held by A	B is under a duty	A has a no-claim
Liberty held by A	B has a no-claim	A is under a duty
Power held by A	B has a liability	A has a disability
Immunity held by A	B has a disability	A has a liability

Figure 1

Does Albert have the power to forfeit or transfer his right to act in self-defense? Does he possess an immunity from having this right revoked or altered by others? These are all important questions that can be asked about most alleged rights, and Hohfeld's framework enables us to be more precise about what's at stake whenever a right is invoked.

2. The Function of Rights

Many philosophers believe that there is a single account of the connection between claims and duties—a single explanation of exactly what makes it the case that a duty-bearer owes his duty to the claim-holder, and not to some other person. Two different theories of this relationship have dominated the philosophical literature.

Some argue that a normative proposition only merits the label of a "right" when there is a right-holder who is uniquely empowered to make choices about what other people's duties are with regard to some action or object (Hart 1955; Hart 1982: 162–93; Steiner 1994: ch. 3; Steiner 1998). Suppose Bert has the right to that computer—he owns it. We generally think this means Bert has the authority to determine how people can behave with regard to the computer. Others are likely under duties not to touch or use the computer, and Bert is the person who can either demand these duties be observed, or choose to waive these duties by giving others the permission to touch or use the computer. Bert is also surely the person who can decide to sell the computer to someone else or the one who can choose to destroy the computer if he wants to. The *will* or *choice theory* of rights reflects these judgments since it identifies the right-holder as the person who holds all the unextinguished second-order *powers* with regard to the duties correlated with a given claim. Those powers include the power to: (a) waive compliance, (b) demand compliance, (c) waive compensation for breaches, (d) demand compensation for breaches, (e) waive enforcement, and (f) demand enforcement (Steiner 1998: 240–5).

The will theory, as a general account of rights, has a number of virtues. First, it coheres with many of our beliefs about property rights (e.g. Bert and his computer), and so has a certain intuitive appeal. And it is consistent with the view that a set of legal or moral rights must be *compossible*, that is, that all the duties entailed by any set of rights are consistent and capable of fulfillment without conflict (Steiner 1994: 86–101). If rights necessarily identify the agent whose choices uniquely determine the enforcement or waiver of claims, then the domain of rights must be divided in such a way that for every alleged duty only one agent ultimately holds the power to choose whether the duty will be observed. Alternative theories of rights, for example the interest theory considered below, seem to allow for the possibility of rights conflicts since such theories do not require that each right be held by a single agent with the power to demand or waive enforcement.

Finally, the will theory offers a theoretical framework for a compelling normative theory of justice. On one Kantian view, justice should allocate to each of us the largest possible sphere of freedom compatible with a similar freedom for all. If spheres of freedom are measured by the domains over which an individual has the authority to decide what actions are required or permitted, then rights as understood by the will theory are the fundamental units of analysis in a Kantian conception of justice (Steiner 1994).

Despite these attractive features, the will theory also faces several serious objections. Most obviously, if right-holders must be choosers, then creatures incapable of rational choice cannot have rights. So, children below a certain age, people with serious mental

disabilities, and non-human animals are all apparently precluded from being right-holders on the view offered by the will theory. This generates three distinct objections: (a) it fails to cohere with our ordinary legal and moral discourse, which does attribute rights to children, the mentally disabled, and (more controversially) animals; (b) it is normatively implausible to claim that creatures incapable of rational choice do not have rights; and (c) if creatures incapable of choice don't have rights, there will be insufficient normative constraints regarding the treatment of those creatures. Critics of the will theory often advance one or more of these objections (Kramer 1998: 69–70; MacCormick 2007: 121–5; Wenar 2005: 239–40).

The conceptualization of the right-holder as the agent with the power to demand or waive enforcement of a duty yields further counter-intuitive results in the criminal law. For example, in the criminal law of most countries a law-breaker is liable to punishment regardless of the wishes of the victim. The victim of an assault, for example, lacks the power to choose whether his assailant will be prosecuted—this power lies in the hands of state officials. Thus, the will theory declares that individuals lack many rights within the criminal law that we assume individuals do in fact possess: rights not to be assaulted, murdered, or raped for example (though the civil law does allocate these rights to individuals in the manner required by the will theory). The will theory instead declares it is the relevant law enforcement officials who hold these rights, since these are the officials with the powers to choose whether alleged offenders will be prosecuted. Furthermore, the will theory does not allow for the possibility of inalienable rights (MacCormick 1977: 198–9). For every right, according to the will theory, there must ultimately be someone who has the power to waive the right. So the will theory denies, for example, that you have an inalienable right against being enslaved—you, or someone else, must have the power to waive this claim, that is, the power to *extinguish* the duty that others previously were under not to enslave you.

The will theory's main rival is the *interest theory* of rights (Kramer 1998; Lyons 1994: 23–46; MacCormick 1977; Raz 1986: ch. 7). According to the interest theory, a necessary condition for one person being a right-holder is that the alleged right serve some interest of the person in question. Rights are thus not identified by determining who holds the power to waive or demand enforcement of a claim, but rather by determining who benefits from the existence of the right.

The interest theory's central idea that a right ought to advance some interest of the right-holder coheres with many of our ordinary beliefs about rights. For example, the interest theory can easily explain why children, the mentally disabled, or animals have rights, since being a rational chooser is not a necessary condition for being a right-holder according to the interest theory. So long as these beings have interests that can be protected by claims, they can be right-holders. It also has the virtue of being able to easily explain how individuals have rights under the criminal law even when they lack the power to waive or demand enforcement, and also how some rights might be inalienable—it might not be in the right-holder's interest to have the power of waiver.

But the interest theory also faces serious objections. For instance, there is the awkward issue of third-party beneficiaries (Hart 1982: 187–8; Steiner 1998: 284–6). Suppose that Albert agrees to pay $100 to Betty in exchange for piano lessons. Betty intends to spend this money on a gift for her nephew, but she cannot afford to buy her nephew a gift unless she receives the money from Albert. It seems odd to suppose that because the nephew is a third-party beneficiary to the agreement between Albert and Betty he is therefore a right-holder with regard to Albert's payment of the $100—that Albert owes

this duty not only to Betty, but also to the nephew. Yet this is the conclusion that some versions of the interest theory reach.

There also seem to be cases where an individual can be a right-holder without this right serving to protect any of the right-holder's interests. For example, if I inherit an elderly donkey from my Uncle Ned that I'm unable to sell (no demand for elderly donkeys), I have legal rights to the donkey, but this could be a serious inconvenience and not a benefit of any kind.

Finally, the interest theory must posit that whenever public officials possess a right, this is explained by the fact that the right serves some interest of the public official. But this view seems either false or else trivially true in the sense that it's usually advantageous to possess rights.

Of course, will and interest theorists believe the objections to their respective theories can be met (Kramer and Steiner 2007). But in light of the difficulties faced by both theories, some philosophers have chosen instead to search for a third theory of rights, one that either fuses the interest and will theories (Sreenivasan 2005, 2010) or eschews both views by abandoning the search for a single-function theory of rights (Wenar 2005, 2008).

3. Theories of Rights: Conceptual or Normative?

The debate between the will and interest theories is often presented primarily as a *conceptual* dispute about rights (Kramer 1998: 91–101; Steiner 1998: 293–8). On this view, each theory is an attempt to uncover the most coherent conception that underlies the everyday practice and discourse of rights, making do with as few normative assumptions as possible. If we characterize the debate in this way, however, both theories seem seriously flawed, for as we've already seen both are, in different ways, often inconsistent with the everyday practice and language of rights.

Maybe the reason that both the will and interest theories appear inconsistent with ordinary rights talk is because both theories seek a *singular* explanation of the function of rights. If ordinary rights discourse is too complex to be reduced to a single function, perhaps the solution is to abandon the effort to find a single function that rights perform, and instead be more faithful to ordinary language by embracing a multi-function theory of rights (Wenar 2005, 2008).

Although I think there is a lot to be said in favor of a multi-function theory of rights, I do not believe theories of rights should be developed and evaluated primarily on the basis of how well they accord with ordinary language. If theories of rights only aim to faithfully reflect ordinary language, these theories might not provide enough normative guidance—the kind of guidance an account of rights ought to provide. Suppose, as is often the case, we want to answer a normative question about rights—we want to know whether individuals *ought* to have a right of some type X. If our competing theories of rights only aim to uncover the best conception underlying everyday practice and language, each theory would only purport to tell us whether having a right to X is consistent with that theory's account of how the term "right" is used in ordinary discourse. But, unless we hold a very peculiar normative theory—one where substantive normative claims about what we have reasons to do can be determined by how people use words—this won't help us address the normative question. For example, if I say "as a matter of moral fact, animals have certain rights, and so we have reasons not to treat animals in certain ways" the truth of this normative statement is not threatened by a

purely conceptual version of the will theory, since all that theory can do is insist that my claim is inconsistent with the best construal of ordinary rights talk. I can plausibly reply, "So much the worse for how people have been talking about rights; it remains true that animals have important interests in not suffering pain, and these interests ground claims that are correlated with duties not to harm them in various ways" (see also Simmonds 1998: 212–13). Even if this statement is not consistent with the best construal of ordinary rights discourse, it might still be true. Surely a theory of rights should offer some significant normative guidance in moral and political philosophy—not merely telling us how we use words, but also whether we're using those words to make valid or true normative assertions.

In light of these worries, maybe we could make better progress by focusing on the following question: what normative role should rights play in theories of justice or morality? (See also Dworkin 2011: ch. 8 on interpretive concepts.) The will theory might then tell us that rights should mark each person's fair share of freedom—the domain within which each person has the normative authority to determine what will occur. The interest theory, on the other hand, might tell us that rights should protect or promote whatever interests are sufficient to warrant placing some person or persons under a duty (Raz 1986: 166).

Although some of the most prominent will and interest theorists reject this construal of the debate, I think the normative dimension cannot be ignored when developing and evaluating theories of rights. Many of the most important questions about rights are normative: Do animals have rights? Should women have an unconditional right to an abortion late in the second trimester? Is basic medical care a human right? The best theory of rights ought to cohere with whatever we believe the correct answers are to these and other substantive questions about rights (Simmonds 1998). Here, as in other areas of moral and political theory, I think we do best by adopting the method of reflective equilibrium: going back and forth between our more abstract theories or principles, and our considered convictions about individual cases with the aim of achieving a reasonably coherent fit between them (Rawls 1999: 42–5). I am suggesting it is not only normative moral and political theories that require the method of reflective equilibrium, but also our theories of normative concepts. This process of reflective equilibrium might lead us to see that some version of either the will theory or the interest theory (or some other alternative) provides the best conception of rights—the account that seems to best cohere with our considered judgments about which moral and political rights there really are.

A possible objection to this proposal is that our deep normative disagreements might make it very difficult, even impossible, to understand how competing *conceptions* of rights are nevertheless recognizable versions of the same normative *concept*. (For this distinction see Rawls 1999: 5.) Would we have any reason to believe that the things called "rights" in my normative theory are in some way relevantly similar to things labeled "rights" in your very different normative theory? If our normative theories are fundamentally different, won't you and I simply be talking past one another when I assert that all individuals possess some right, and you deny this assertion? In order for us to be having a genuine disagreement, we have to share some notion of what rights are that is independent of our differing normative frameworks. What could that shared normative notion be?

I think the answer is roughly as follows (Thomson 1990: 212–23). First, rights are moral constraints on the actions of agents; they constrain the behavior of individuals

who can understand and act for moral reasons. Second, rights are grounded in the fact that individual right-holders (whoever they turn out to be) have their own aims and interests that are distinct from the aims and interests of others, and distinct from what would be best from some collective point of view. If we didn't each have distinct lives to lead—if each individual's good was always identical to the good of the collective—then the idea of individual rights would seem unnecessary. But we do have distinct aims and interests: killing Albert and redistributing his organs to Betty, Carl, and Debbie (who will all die without Albert's organs) might be the best thing for Betty, Carl, and Debbie, but it doesn't seem to be the best thing for Albert. Rights reflect the fact that other individuals can be the source of claims on us, apart from whatever value individuals' lives may have from an impersonal or collective perspective. To be clear, I am not proposing this as an alternative to the will or interest theories—I'm suggesting that these are the two key features of a *shared* moral concept that underlies competing conceptions of rights. Let's call this view of rights *the constraint view*.

The constraint view is compatible with a wide range of theories in moral and political philosophy. In particular, adopting the constraint view does not require adopting a nonconsequentialist or deontological view of morality or justice, though it is consistent with such theories. The constraint view tells us that individuals have claims on the behavior of others in virtue of the fact that they are individuals with distinct aims or interests, and so it posits a sphere of claims whose normative force is not reducible to aggregate goodness and thus places some moral barriers on the pursuit of aggregate goodness. This view is potentially compatible with many consequentialist theories, provided those theories make room for the idea that aggregate goodness is not all that matters.

Of course, the constraint view is only a very thin concept of rights—it only aims to identify a normative core that all plausible conceptions of rights have in common. In conjunction with Hohfeld's conceptual framework, the constraint view can give us some limited sense of what any theory of rights ought to look like, but it leaves most important questions about rights unanswered. In particular, it does not tell us anything about what specific rights individuals possess, or what the grounds for those rights are. Those arguments, I have suggested, must be provided by a more general normative theory.

There are two further features that I believe any sound conception of rights ought to possess, but both features are controversial and may be specific to *moral* rights, and so I do not include them as necessary features of the constraint view. First, rights constrain our behavior in the sense that they enter into our deliberations about how we ought to behave—they provide reasons to do, or forbear from doing, various actions. I therefore believe that rights apply only to actions that arise out of a person's voluntary agency: we cannot have claims against the involuntary movements of others, or against certain states of affairs occurring that no person's voluntary agency could have prevented (for the alternative view see Thomson 1991). If, for example, Albert stands at the bottom of a well, he can have a claim against Betty that she not try to jump down the well and land on him in order to kill him. But Albert cannot have a claim against the same state of affairs occurring as a result of some non-agential force, for example, a claim against Betty being unforeseeably thrown by a giant gust of wind down the well towards Albert (see McMahan 1994: 276–7; Otsuka 1994: 79–84). This latter state of affairs is no doubt very bad for Albert, but because it does not arise as a result of anyone's agency it cannot be the sort of event against which Albert could have a valid claim, just as Albert could not have a valid claim against a tiger's attack or against the wind blowing his hat

off. Rights require that others take us into account when they decide how to act, and so claims can only correlate with duties that can be voluntarily performed. This has important implications for many normative questions, but discussion of these implications is beyond our scope here.

Second, I believe we ought to accept that rights can conflict with one another, and the remainder of this chapter is devoted to a brief defense of this idea.

4. Rights Conflicts and the Role of Rights

Albert's Problem: Albert is an artist and he simultaneously (via email) promises to give Betty the painting of her choice tomorrow, and promises to give Carl the painting of his choice tomorrow. Tomorrow arrives, and Betty and Carl choose the same painting.

I believe this is an example where rights can conflict. Albert has a duty to give the painting to Betty and he has a duty to give the painting to Carl, but he lacks the ability to fulfill both duties. Since each person has a claim against Albert (and, let's assume, the associated powers to waive or demand enforcement) it would be wrong for Albert to default on either duty without first trying to get the consent of one or both parties to some alternative arrangement. But suppose Albert can't get Betty or Carl to agree to any alternate arrangements. Now what? I think that whatever Albert does now, he is going to default on a duty he owes to at least one person, and he therefore *wrongs* at least one of the people involved.

Some philosophers, however, deny that rights can conflict. Why take this position? The most important argument takes the following form (Steiner 1994: 86–101). If Albert owes a duty to give Betty the painting, then this is what Albert *ought* to do. It therefore cannot be true that Albert has a duty to give Betty the painting and a duty to give Carl the painting since this would entail a contradiction—it would entail that Albert both ought to, and ought not to, give each person the painting. Since it cannot be true that Albert both ought to and ought not to do the same action, rights cannot conflict. Let's call this argument against the possibility of rights conflicts *the argument from contradiction*.

I think we can reject the argument from contradiction by rejecting the premise that rights generate oughts, or at least by rejecting the premise that rights generate all-things-considered oughts (Thomson 1990: ch. 3; see also Kramer 2009; Sreenivasan 2010). I think we do better by adopting a view where rights provide us with *pro tanto* reasons for action. If Betty has a claim that correlates with Albert being under a duty to give her the painting, this means that Albert has a reason to give her the painting. But it would be very implausible to infer from the fact that Albert owes this duty to Betty to the conclusion that what Albert ought to do is give her the painting. In order to know what Albert ought to do, we need much more information about what options Albert faces, what the consequences of those options are, and what other claims might be affected by Albert's choices. Suppose that Albert's giving the painting to Betty will result in the death of twenty innocent people—can we really say that this fact is *irrelevant* to our judgment about what Albert ought to do, and that only Betty's claim is relevant? I think it's clear we cannot.

How can the proponent of the argument from contradiction respond? The proponent might reply that, of course, rights do not entail facts about what we ought to do *all-things-considered*, but they do entail facts about what we ought to do within whatever normative domain the rights in question are found. So, for example, a claim that

correlates with a legal duty for Albert tells Albert what he ought to do as a matter of law, and a claim that correlates with a duty of justice for Albert tells Albert what he ought to do as a matter of justice, and so on. The opponent of rights conflicts can then insist that the argument from contradiction merely resurfaces at a different point. If legal rights can conflict, this will result in contradictions regarding what the law permits or requires. If duties of justice can conflict, this entails contradictions regarding what justice permits or requires. And it's just as problematic to have a theory of rights that permits logical contradictions about the demands of law or justice.

But this further revision to the argument from contradiction is vulnerable to the same objection as the original version of the argument. It's not credible to suppose that legal rights represent conclusions about what we ought to do as a matter of law, or that claims of justice represent conclusions about what we ought to do as a matter of justice. For example, as a matter of justice, individuals have claims not to be non-consensually killed by others. But this fact, on its own, does not tell us whether a bomber in a war is permitted, as a matter of justice, to drop a bomb on an enemy target when this will result in the non-consensual killing of several innocent civilians nearby. In order to know whether the bomber's actions are just or unjust, we need to find out whether his actions are consistent with the correct principles of just war, for example, whether his actions have a just aim, and whether the deaths of the civilians will be proportionate relative to the good achieved by the bombing. The same point applies with regard to legal rights.

The proponent of the argument from contradiction might insist that rights only appear to conflict when we fail to fully specify the conditions of particular rights. On this view, individuals do not have general rights not to be killed, they only have very complex rights not to be killed under various circumstances, and those circumstances will not include being killed as a proportionate and unintended consequence of a just attack in a just war. I think this response fails both because it makes rights explanatorily inert (more on this below), but also because it's normatively implausible. I think when innocent people are killed in just wars, it remains true that those innocent people had rights not to be killed, and this fact partly explains why just conduct in war aims to minimize the number of innocent people who are killed in the course of military action. This view also allows us to easily explain why the innocent victims of justified violence in war might nevertheless be entitled to compensation—they are owed compensation because they had rights against the harm imposed.

In sum, it's no more plausible to suppose that rights entail conclusions about law or justice than it is to suppose rights entail conclusions about all-things-considered practical reason. Instead rights are most plausibly seen as generating *pro tanto* reasons about what we ought to do. How weighty these reasons are will then depend on the particular type of right at issue; for example, my claim not to be killed by you is a lot weightier than my claim that you not pinch me on the arm.

Both these proposals—that rights can conflict, and that rights generate *pro tanto* reasons whose weight varies in accordance with the type of right at issue—might appear to lend support to the interest theory of rights, but they don't necessarily have this implication. It's true that these proposals appear incompatible with certain versions of the will theory (Steiner 1994, 1998), but some proponents of the will theory appear to accept both these points (Hart 1955: 185–6). My purpose in briefly arguing for these conclusions has not been to take a side in the debate between the interest and will theories, but rather to defend a particular conception of the role that rights ought to play in our moral and political reasoning. If rights were conclusions about what to do,

we would have difficulty explaining how it can sometimes be just for a bomber to kill innocent civilians in war, or how it can be morally justified to default on a debt if this is the only way to save someone's life. Instead, rights make more sense if we see them as one important set of inputs into our moral and political reasoning; as generating strong, but not necessarily decisive, reasons to act in various ways.

Not only does this view of rights allow us to make more sense of various examples where rights appear to conflict either with each other or with other considerations, it also allows rights to play an important explanatory role in moral and political theory. If saying "I have a right that you do X" was just another way of saying "morality requires you to do X," then rights would not play any part of the story in explaining *why* morality requires you to do X. All the normative work would be done by other moral concepts. But on the view of rights I've defended, rights can and do play an important explanatory role. If morality requires you to do X, one of the reasons for this fact might be that someone else has a claim that you do X. Most of us, in our moral and political reasoning, take rights to have this kind of explanatory power. We believe that individuals have claims not to be killed, raped, tortured, assaulted, or coerced, and these rights generate powerful reasons to treat people in particular ways. If, like me, you think people have these rights and this fact partly explains *why* the requirements of morality and justice take the shape they do, then you should accept that rights are not the final word in our deliberations, but rather are an important source of reasons in moral and political philosophy.

Acknowledgments

I'm very grateful to Fred D'Agostino, Jerry Gaus, Hillel Steiner, Zofia Stemplowska, and Rebecca Stone for comments on previous drafts.

Related Topics

Liberalism, Natural Law and Rights Theory, Left Libertarianism, Libertarianism, Human Rights and Cosmopolitanism

References

Dworkin, R. (2011) *Justice for Hedgehogs*, Cambridge, MA: Belknap Press of Harvard University Press.

Hart, H.L.A. (1955) "Are There Any Natural Rights?" *The Philosophical Review* LXIV: 175–91.

—— (1982) *Essays on Bentham: Jurisprudence and Political Theory*, Oxford: Clarendon Press.

Hohfeld, W. (1919) *Fundamental Legal Conceptions*, W. Cook (ed.), New Haven, CT: Yale University Press.

Kramer, M.H. (1998) "Rights Without Trimmings," in Matthew H. Kramer, N.E. Simmonds, and Hillel Steiner, *A Debate Over Rights*, Oxford: Oxford University Press.

—— (2009) "Consistency is Hardly ever Enough: Reflections on Hillel Steiner's Methodology," in Stephen de Wijze, Matthew H. Kramer, and Ian Carter (eds.), *Hillel Steiner and the Anatomy of Justice: Themes and Challenges*, New York: Routledge.

—— and Steiner, H. (2007) "Theories of Rights: Is There a Third Way?" *Oxford Journal of Legal Studies* 27: 281–310.

Lyons, D. (1994) *Rights, Welfare, and Mill's Moral Theory*, Oxford: Oxford University Press.

MacCormick, N. (1977) "Rights in Legislation," in P.M.S. Hacker and Joseph Raz (eds.), *Law, Morality, and Society: Essays in Honour of H.L.A. Hart*, Oxford: Clarendon Press.

—— (2007) *Institutions of Law: An Essay in Legal Theory*, Oxford: Oxford University Press.

McMahan, J. (1994) "Self-Defense and the Problem of the Innocent Attacker," *Ethics* 104: 252–90.

Otsuka, M. (1994) "Killing the Innocent in Self-Defense," *Philosophy & Public Affairs* 23: 74–94.

Rawls, J. (1999) *A Theory of Justice: Revised Edition*, Oxford: Oxford University Press.
Raz, J. (1986) *The Morality of Freedom*, Oxford: Clarendon Press.
Simmonds, N.E. (1998) "Rights at the Cutting Edge," in Matthew H. Kramer, N.E. Simmonds, and Hillel Steiner, *A Debate Over Rights*, Oxford: Oxford University Press.
Sreenivasan, G. (2005) "A Hybrid Theory of Claim-Rights," *Oxford Journal of Legal Studies* 25: 257–74.
—— (2010) "Duties and Their Direction," *Ethics* 120: 465–94.
Steiner, H. (1994) *An Essay on Rights*, Oxford: Blackwell.
—— (1998) "Working Rights," in Matthew H. Kramer, N.E. Simmonds, and Hillel Steiner, *A Debate Over Rights*, Oxford: Oxford University Press.
Thomson, J.J.T. (1990) *The Realm of Rights*, Cambridge, MA: Harvard University Press.
—— (1991) "Self-Defense," *Philosophy & Public Affairs* 20: 283–310.
Wenar, L. (2005) "The Nature of Rights," *Philosophy & Public Affairs* 33: 223–53.
—— (2008) "The Analysis of Rights," in Matthew Kramer, Claire Grant, Ben Colburn, and Anthony Hatzistavrou (eds.), *The Legacy of H.L.A. Hart*, Oxford: Oxford University Press.

55
TOLERATION

Peter Jones

As an idea that inhabits contemporary western political thinking, toleration has its origins principally in the European Reformation of the sixteenth century. Ideas and practices of toleration have a much longer pedigree going back to the ancient and medieval worlds (Nederman 2000) and beyond the European world (Lauren 1999), but, in the west of the modern era, the Reformation and the religious divisions it created provided the main stimulus for thought about toleration. Religious difference remains an important focus for the theory and practice of toleration in virtually all parts of the contemporary world, but the idea of toleration is now applied much more widely, for example, to differences of culture, ethnicity, political opinion, life-style, sexuality and sexual conduct, forms of dress, and so on. "Tolerant," "tolerance" and "toleration" are generally used as terms of commendation, while "intolerant" and "intolerance" are typically pejorative terms, but that should not mislead us. There is much that we should not tolerate and therefore much of which we are rightly intolerant.

1. The Idea of Toleration

We tolerate when we endure or permit something to which we object. If we find something unexceptionable, we have no occasion to tolerate it. At least, that is what "toleration" means in its orthodox usage. Nowadays, "toleration" is sometimes used in a looser sense, in which to be tolerant means to be accepting and perhaps even approving of the differences of others (e.g. Apel 1997; Walzer 1997: 11–12). I shall comment later on this more relaxed understanding of toleration and its possible justification but, for the time being, I shall keep faith with the orthodox understanding of toleration. If we depart from that orthodoxy, we lose what is distinctive and interesting about toleration: the way in which it combines negative appraisal with a commitment to not preventing or suppressing what it appraises negatively.

The negative appraisal of the object of toleration—of what it is that we tolerate—can be of various sorts. Toleration is most commonly toleration of the conduct of others and we might disapprove of that conduct because we believe it to be immoral, irreligious or imprudent. But we might also tolerate conduct that we find irritating, unpleasant or distasteful, in other words, conduct that we merely dislike. We can extend toleration to non-human beings and to objects; we might, for example, tolerate a barking dog or a squeaking door.

Toleration itself is most commonly explicated as not preventing: we tolerate when we refrain from preventing what we disapprove of or dislike. In political contexts, "not preventing" most commonly takes the form of refraining from legal prohibition, but it

can encompass the many other ways in which we might remove or diminish a person's freedom. Intolerance includes seeking to prevent as well as actually preventing; laws, for example, can be instruments of intolerance even though they do not always succeed in preventing what they prohibit. Toleration can also consist in not compelling: religious toleration, for example, requires us not to compel people to comply with a faith they do not hold, as well as not to prevent their complying with the faith they do hold.

Merely expressing disapproval or dislike is usually thought to be consistent with toleration. Given that disapproval or dislike is an essential feature of toleration, it would be odd to deem people intolerant merely because they disapprove or dislike or because they make that known. Seeking to change people's thinking through rational persuasion is also usually thought consistent with toleration. But here there are lines that may be crossed. When disapproval or dislike becomes the oppressive and stigmatizing social opinion that so worried John Stuart Mill (1978), it can clearly be an instrument of intolerance. So too can "persuasion" when it ceases to be rational and becomes brow-beating, emotional blackmail, manipulation or some other form of mind-changing technique that is at odds with the autonomy of those at whom it is directed.

Toleration is commonly associated with not preventing because, if *A* does nothing to prevent *B*'s conduct, it would be odd to hold that he fails to tolerate *B*'s conduct. That is why *A*'s indicating disapproval of *B*'s conduct but doing nothing to impede it, would not normally count as intolerance. Yet there is one form of intolerance in which prevention and its variants are not at stake. Consider the act of placing a pig's head in a synagogue or mosque. We would normally consider that an instance of intolerance even though it neither prevents nor compels. That example indicates that intolerance can sometimes take the form of non-preventative persecution—an act that aims to inflict suffering rather than to prevent or coerce.

Another feature of toleration is that we can tolerate only what we are able not to. Suppose you behave in a manner that I dislike but that I am unable to prevent. In that case, my toleration is neither here nor there; it is not an option I enjoy. I can tolerate only when I am genuinely able to prevent but choose not to. So, if intolerance is not an option genuinely available to me, neither is tolerance. This is commonly known as the "power" condition of toleration.

It is, however, a condition that needs qualification. We describe people as tolerant in relation to their dispositions, attitudes and beliefs as well as their conduct. Someone can have a tolerant disposition, attitude or belief even in respect of circumstances that she is powerless to affect. We might deem her tolerant insofar as she would not prevent the objectionable conduct of another even if she could.

2. Justifying Toleration

How might toleration be justified? We should begin by noting that arguments against intolerance are not always arguments for toleration. Consider the position of someone who opposes intolerance of abortion and homosexual conduct. His opposition to that intolerance is very likely to be grounded in his belief that there is nothing wrong with abortion or homosexuality, so that intolerance of either is unjustified. If that is the argument he presents to those who oppose abortion and homosexuality, he is not presenting them with reason to be tolerant; rather, he is providing them with reason to abandon their objection to abortion and homosexuality so that those practices will cease to be possible objects of their toleration. Insisting that others should tolerate a practice

because they are wrong to object to it is not really arguing for toleration at all. An argument for toleration would leave the opponents' disapproval of a practice in place but give them reason not to prevent the practice in spite of their disapproval.

That is not to imply that it is better for people to be disapproving so that they can be tolerant. If someone's disapproval is misplaced, we normally think it better that they should correct their error than that they should persist with it but be tolerant. Certainly those who are objects of disapproval would typically prefer not to be disapproved of, than to be disapproved of but tolerated. The point is simply that an absence of intolerance does not always signal the presence of toleration.

The notion that it can be good or right to allow the bad or the wrong is sometimes described as the "paradox of toleration" (e.g. Mendus 1989: 18–21). How can it be right not to prevent a preventable wrong? Why should it be good to permit bad conduct? Questions like these can indeed make toleration seem paradoxical but, in fact, there is nothing very puzzling about it. Toleration is a coherent and intelligible position just in case our reason for not preventing outweighs or overrides our reason for preventing. What sort of reason for non-prevention could that be?

People's reasons for being tolerant can be prosaic and self-interested. I might, for example, tolerate the bad behavior of my neighbor's children simply because I want to retain my neighbor's friendship, or I might tolerate my boss's ill-mannered conduct simply because my promotion depends upon his patronage. Similarly, a government might tolerate another country's abuse of human rights merely because it wants to retain that country as an ally. But the reasons for toleration that have preoccupied political philosophers have been of a more general and principled kind.

There is no finite list of such justifications, but four types of reason have been particularly prominent in recent thinking about toleration as an ideal that should shape the life and the arrangements of a society.

One of these focuses on the good consequences of toleration, especially peace and the many other goods the peace makes possible (Gray 2000: 1–33, 105–139; Horton 2011). Historically, intolerance has frequently led to conflict and repression, both within societies and between them, and it is still a major cause of death, bloodshed and suffering. Sometimes a society's embracing toleration is the only way in which it can secure peace and the fruits of peace.

In the development of religious toleration, a grudging acceptance that "live and let live" was a necessary condition for securing peace amongst conflicting groups of believers often provided the initial motivation for toleration. Only later did people come to conceive toleration amongst those of different denominations and faiths as a right and proper condition. Because it can take the form of a *modus vivendi* to which the warring parties agree only as a temporary expedient, toleration grounded only on advantage can prove unstable. If the parties to a conflict tolerate one another only if and because neither can prevail over the other to its own advantage, each has reason to abandon toleration if circumstances should change such that it can prevail to its own advantage. But there is no reason why the consequences of toleration need be valued only in such a narrowly self-interested and amoral fashion. If we have strong moral reason to avoid or to minimize human death and suffering, and if toleration is the most effective way of securing that goal, we have strong moral reason, which should be matched by a similar motivation, to be tolerant even if we could be intolerant to our own advantage. Nor need we conceive this consequentialist valuing of toleration only in grand terms of war and peace. If, for example, in frustrating someone's desire to live as she wishes or as she

believes she should, we cause her anguish and unhappiness, those consequences give us reason, at least pro tanto, to desist from that frustration.

A second sort of justification for toleration focuses upon skepticism or the modesty with which we should hold our beliefs (e.g. Barry 1995: 168–188; Rorty 1989: 73–95). Issues of toleration and intolerance are most frequently associated with conflicts of belief. Person *A* believes that the conduct of person *B* is wrong and wrong in a way that ought not to be tolerated. But if *A* has reason to be less than fully confident about his objections to *B*'s conduct, he also has reason to hesitate about not tolerating it. Skepticism can therefore give us reason to refrain from intolerance.

But can skepticism really ground toleration? Toleration, as we have seen, entails allowing what we disapprove of or dislike and, in the most challenging cases, what we believe to be wrong. But if *A* doubts his own beliefs, so that he simply does not know whether *B*'s conduct is wrong, and if he refrains from interfering for that reason, we cannot really say that *A* tolerates *B*. However, that dismisses the relevance of skepticism to toleration a little too easily. *A* can find a position plausible or credible, and more plausible or credible than the position embraced by *B*, but still have reason to accept that he could be wrong and that *B* could be right. Thus, even though *A* has reason to regard *B*'s belief as inferior to his own, modesty about the status of his belief can give him reason to refrain from imposing it upon *B*. That argument is of particular force in relation to non-evidential beliefs, such as religious faith. It is true that the certainty with which people hold their beliefs is sometimes inversely related to the evidence supporting them. But we are concerned here with reasons, and people might have *reason* to be modest about their beliefs even if they themselves fail to recognize that.

A third argument for toleration focuses on autonomy as an essential constituent of the good life (e.g. Mill 1978; Oberdiek 2001: 111–142). The essential idea here is that a form of life can be a good life for a person only if that person genuinely embraces it. We might think that there is a better form of life the person could and should live, and we might be right. But that "better form" will not be realized as a good life for the person if he is coerced or pressurized into living it and has no commitment to it himself. A good life has to be lived "from the inside"; it cannot be imposed by outsiders (Dworkin 2000: 237–284). Again, we could question whether this is really an argument for toleration since it enjoins us to recognize the limits of our capabilities: we might be able to conceive a better life than the life a person is living but we cannot make that person's life better through imposition. Nevertheless, the argument remains an important one for toleration in that it gives us reason to refrain from coercively preventing or altering something of which we disapprove.

A variant of this argument has provided one of the most common justifications for religious toleration, at least within Christianity. Proponents of religious toleration, such as John Locke (1968), argued that religious faith could be of value—and of value particularly to God—only if it was sincerely held. External acts unaccompanied by sincere belief were spiritually worthless. Hence, the state should not coerce people in religious matters since, while it could alter their external conduct, it could not change what really mattered: their internal convictions. However, this argument could be used to opposite effect. Thomas Hobbes argued that, since what mattered to God was sincere *belief*, people had no reason to dissent or resist if their ruler required them merely to *act* contrary to their religious belief (Hobbes 1957: ch. 42).

Finally, justifications for toleration can be deontological in character (e.g. Forst 2003; Larmore 1987: 40–68; 1996: 121–151; Nagel 1991: 154–168). Rather than focusing on

the plural nature of good lives and what, all things considered, can make a life go well or badly, these justifications focus on the status of the person and the respect others owe him. Adult individuals are capable of thinking, reflecting and choosing for themselves. When we ignore their wishes or beliefs about how they should live and impose upon them others' conceptions of how they should live, we treat them as less than persons. We deny them the equal status to which they are entitled and the freedoms that respecting their status entails. This approach is often described as Kantian and is particularly important in contemporary liberal defenses of toleration. It provides a peculiarly apt form of argument for toleration, since the ground upon which it justifies toleration—people's *status* as persons—is categorically different from what it is that requires toleration—the *merit* (or demerit) of the forms of life that people live.

3. The Limits of Toleration

However strong a justification of toleration, there will be a limit to what it justifies. Some sorts of human conduct are intolerable and rightly not tolerated. How then do we decide what they are? The answer might be thought to lie in the adage that we should tolerate everything except intolerance. However, a moment's reflection reveals how unsatisfactory that adage is. Consider acts that we regard, routinely and uncontroversially, as intolerable, such as murder, assault, rape, fraud, and defamation. It would be odd to characterize the wrongfulness of those acts as "intolerance." The same is frequently true of more controversial cases. If, for example, we believe abortion or gay marriage or the eating of meat or ritual circumcision (male or female) to be intolerable, we would still be hard put to represent those practices as "intolerant" and our own opposition to them as "intolerance of intolerance." The simple truth is that the domains of the intolerable and the intolerant are not co-extensive: not everything that is intolerable is intolerant and not every instance of intolerance is intolerable. There is no simple formula for setting the limits of toleration; all of the many and different reasons that might justify our limiting someone's freedom will also be reasons that justify intolerance, in spite of the pejorative nuances that often accompany the word "intolerance."

4. The Psychology and the Philosophy of Toleration

We have grown used to the idea that we now live in a tolerant age and one that is more tolerant than its predecessors. Yet it is easy to exaggerate how tolerant we are. Our normal mode of moral or political argument is to identify practices as good or bad, right or wrong, wise or unwise. If we appraise them positively, we want them to be permitted or promoted; if we appraise them negatively, we want them stopped. Consider, for example, contemporary arguments about abortion, the use of embryos or animals in scientific research, gay marriage, and the hunting of animals for sport. In these arguments, each side is primarily concerned to assert the rightness of its own view and the error of its opponents, on the assumption that the right view should determine what should be done. Only more rarely does the argument turn to what we may have reason to tolerate in spite of its being wrong.

It could also be that we care less about many matters, such as religious differences, than did our forebears, so that we exhibit greater indifference, rather than greater tolerance, than they did. Here, however, there is a danger of confusion arising out of two different ways of thinking about toleration: psychological and philosophical. We can

think of toleration psychologically, as an experience that people undergo. If we do, people's toleration will be calibrated according to the extent to which they are willing to endure what they object to. The more they object to something, the more tolerant they will be if they endure it. This way of thinking can generate the "paradox of the censorious tolerator": the more censorious a person, the more tolerant he can be, since his greater censure gives him more to tolerate (Horton 1996: 33–36; Newey 1999: 107–111). There is nothing paradoxical about the censorious tolerator if we conceive toleration as involving a sort of psychological struggle: it will simply be true that those who object more strongly have more to overcome and will exhibit greater tolerance if they overcome it. What makes the case of the censorious tolerator appear paradoxical is the supposition that toleration is virtuous, with the consequent implication that the more censorious we are, the more virtuous we can be. But we need find nothing positively virtuous in people's tolerating what they have no good reason to censure.

The psychological understanding of toleration has merit if we are assessing toleration as a feature of people's characters, but as a way of understanding a commitment to toleration it can be misleading. Identifying different reasons, weighing them against one another, and arriving at an overall conclusion, is not the same as undergoing psychological turmoil and internal torment. We can undertake that philosophical assessment with equanimity. If the balance of considerations comes out in favor of toleration, we can also accept that conclusion with equanimity. If we return to the case for religious toleration, nowadays many people in liberal democratic societies think that case is so well established that they accept it without hesitation. Their ready acceptance of religious toleration need not therefore signal that they are more indifferent to religious matters than their predecessors but simply that they are more persuaded of the case for religious toleration, and it would be perverse to hold that, as their commitment to toleration has increased, so the reality of their toleration has diminished.

5. Toleration, Equality and Rights

A quite different way in which doubt might be cast upon the place of toleration in the contemporary world concerns its compatibility with equality and rights. Toleration is often said to be patronizing, condescending and inegalitarian (e.g. Addis 1997; Brown 2006; Phillips 1999). As we have seen, we tolerate only what we are able to prevent but choose not to. Thus toleration can seem to presuppose an unequal relationship in which the tolerator extends toleration to a subordinate as a matter of grace and favor. The inequality and condescension of that relationship does not comport well with a world in which people possess equal standing and equal rights. Nor is this indictment of toleration of only recent vintage. Immanuel Kant held that an enlightened prince, who appreciated he was duty-bound not to prescribe anything to his population on matters of religion, would spurn the "presumptuous title of *tolerant*" (Kant 1991: 58). Similarly, Thomas Paine insisted that the constitution established by the French Revolution of 1789 had, in establishing universal rights of conscience, abolished rather than instantiated the "despotism" of toleration (Paine 1989: 94).

The understanding of toleration implicit in these criticisms has led some to characterize toleration as supererogatory (Heyd 2008; Newey 1999). Supererogatory acts are good to do but not wrong not to do (or good not to do but not wrong to do); they exceed the demands of duty. Thus, on this analysis of toleration, if you have the right to do as you please and I have a correlative duty to allow you to do as you please, my perform-

ance of that duty will not be an exercise in toleration. I will be tolerant only when I allow you to do something that I have no duty not to prevent your doing; in other words, I tolerate only when I allow you to do what you have no right that I should not prevent your doing. This understanding of toleration may be consistent with its retaining a significant place in personal relationships, but it drives toleration to the margins of contemporary political life. Insofar as modern citizens relate to one another as the possessors of reciprocal rights and duties, they will have no need of one another's toleration; nor will toleration be a significant feature of the political systems under which they live. Liberal democracies, even ideally conceived, will not instantiate toleration; they will supersede it.

Is then toleration obsolete in the contemporary world, at least for political purposes? The monarchs who ruled post-Reformation Europe did treat toleration as a matter of grace and favor. A Protestant monarch who tolerated his Catholic subjects, and a Catholic monarch who tolerated his Protestant subjects, each did something that he had no obligation to do and to which his subjects had no right. Toleration was a form of indulgence. It is easy to suppose therefore that, once democracy overtook monarchy and rights replaced indulgence, toleration became redundant. There would indeed be something very odd in a modern democratic government's aping an early modern monarch in extending toleration to its population, since a democratically elected government is supposed to be the servant not the master of its population. However, we need not think of toleration on that "vertical" model. Rather, in the context of democracy, we can think of toleration as something that citizens extend to one another. They do so by refraining from using democratic political power to impose upon others beliefs and ways of life that those others do not accept. Once we conceive toleration on that "horizontal" model, it becomes clear that it does not presuppose an asymmetric relationship. That is not to say that, in democratic circumstances, toleration cannot be asymmetric and condescending. A democratic majority may extend toleration to a minority, which that powerless minority is unable to reciprocate and which, arguably, reinstates the "vertical" model. But the ideas of democratic *equality* and *liberal* democracy are more consistent with a conception of political toleration as horizontal, multilateral and mutual (Jones 2007).

Even if toleration can be egalitarian, we are left with the objection that rights render toleration redundant: once people have a right to do x, their doing x ceases to be a possible object of others' toleration. That claim is also mistaken. It rests upon a misguided elision of the empirical and the moral in setting the possibility conditions of toleration.

We have seen that, strictly speaking, we can tolerate x only when we are able not to tolerate x. In that sense, power or ability is a possibility condition of toleration. It is easy to elide that condition with another putative condition: we can tolerate x only when we are morally free not to tolerate x. If I am morally unfree to prevent x, not tolerating x is not a moral option I possess. Thus, we might conclude, if you have a moral claim-right to do x and I have a corresponding moral duty not to prevent your doing x, my tolerating your doing x will disappear as an option: I cannot be tolerant in merely doing my duty.

What has gone wrong here is that we have conflated an empirical condition of toleration (power or ability) with a moral criterion of what we are morally free or unfree to do. We are not morally free to do what there is good moral reason not to do; in particular, we are morally unfree to infringe the moral rights of others. But that sort of

consideration is quite different from the empirical question of what, as a matter of fact, we are able to prevent. When there is good moral reason not to prevent, there is good moral reason to tolerate. When people have a moral right not to be prevented, others have a particularly compelling moral reason, an "exclusionary reason" (Raz 1975: 35–48), to tolerate. But that is what these reasons are—moral reasons for toleration. It would be bizarre to hold that strong reasons must transform toleration into something else, so that toleration can remain toleration only so long as it is weakly justified.

Thus toleration can sit comfortably amidst equality, rights and justice. Indeed, a society can instantiate its commitment to toleration by establishing rights for its members to live as they see fit in spite of the disapproving judgments of others (Jones 2012). That conception of toleration is well illustrated by the work of John Rawls, whose *Political Liberalism* (1993) is arguably the most important book written on toleration during the last fifty years and for whom toleration is a requirement of justice. His *The Law of Peoples* (1999) extends his conception of toleration to the international world: just as, intranationally, justice requires citizens to tolerate one another's preferred ways of life, so, internationally, it requires peoples to tolerate one another's preferred forms of society.

6. Toleration, Identity and Difference

There is yet another respect in which toleration can seem out of kilter with the contemporary world. Historically, toleration has been most associated with religious differences and religious toleration has always been the paradigm case of toleration. There is more than historical reason for that. To be a Christian is to reject Islam and to be a Muslim is to reject Christianity. Similarly, to be a Protestant is to reject Catholicism and to be a Catholic is to dissent from Protestantism in all of its many forms. Thus, different religious beliefs are not merely different; they are conflicting. The disapproval or rejection that is a necessary part of the circumstances of toleration is also a necessary feature of religious diversity. That is why toleration has been so closely bound up with religion and why mutual toleration is the best we can hope for in circumstances of religious difference.

We are sometimes enjoined to *celebrate* rather than merely tolerate religious differences but, for anyone who has a serious religious belief, that injunction must be nonsensical. If I am a Muslim, how can I celebrate the fact that the majority of humanity fails to recognize that Mohammed was God's Prophet? If I am a Christian, how can I celebrate beliefs that reject the truth that Jesus was the Son of God and that faith in Him provides the path to salvation? Certainly people can often find reason for toleration within their faith as well as outside it, but a reason for toleration is not a reason for approval and celebration.

Religious differences are not the only intrinsically conflictual differences. Rawls characterizes citizens in modern democratic societies as possessing different "comprehensive doctrines," which lead them to embrace different "conceptions of the good," that is, different conceptions of the ends we should pursue in life. These comprehensive doctrines include philosophical and moral as well as religious doctrines, and they too conflict so that mutual toleration is needed even when people's doctrinal differences are not religious (Rawls 1993: xvi–xxvii, 10–15).

But how far is the diversity that characterizes modern plural societies really *doctrinal* in character? There is nothing doctrinal about differences in race, gender or sexual orientation, even though those differences might become the concerns of doctrine (as, for example, gender and sexual orientation are for many religions). Doctrinal differences

can be associated with ethnic and cultural differences but sociologically those differences are usually more significant as markers of identity than as manifestations of belief. Even religious differences nowadays often seem to matter more as markers of identity than as rival claims to truth, particularly when they are associated with racial, national or ethnic differences.

If we confront, or understand ourselves to confront, a plurality of identities rather than doctrines, how does that bear on toleration? Potentially, very significantly. Toleration can seem inappropriate for differences of identity, not because those differences are intolerable, but because the negative appraisal that toleration presupposes can be quite inappropriate when we move from doctrines to identities. First, if we see the world in terms of identities, we are more likely to see people's differences as *mere* differences, to which no exception should be taken. Second, a negative reaction to an identity is more likely to be a reaction of dislike or hatred than a disapproving judgment and we might think that dislike and hatred should be overcome rather than indulged. Third, the identities that figure prominently in contemporary societies are often "ascriptive" identities—differences that people have through no choice of their own, such as differences of race, ethnicity, culture, gender or disability. Viewing those differences negatively is especially reprehensible. So the toleration that is desirable and praiseworthy in the case of doctrinal differences can be unedifying and unwelcome when it is directed at differences of identity.

Nowadays, it has become common to call for the recognition, rather than the toleration, of identities (e.g. Taylor 1994). To accord an identity recognition is to give it a form of positive endorsement, either in respect of its status or its merit, although in reality calls for recognition are often calls for the abandonment of misrecognition or non-recognition—calls for people to cease viewing or treating particular identities negatively. Recognition can be combined with toleration (Galeotti 2002); if, for example, we recognize a person as having a status equal to our own, that can give us reason to tolerate that person's living in ways of which we disapprove. But it has become common for people to demand endorsement for their *particular* identities as, for example, "Muslim" or "gay," in which case it becomes impossible both to recognize and to disapprove but tolerate (Jones 2006).

In spite of these compelling considerations, toleration can remain significant even in a world of identities. Whatever the normative logic of "identity," the propensity of differences in identity to stimulate antagonism, conflict and even genocide, have been amply illustrated by recent history. As long as different identities prompt hatred and enmity, we shall have reason to urge the merits of toleration. In addition, identities are often caught up with the sorts of difference of belief and value for which toleration is entirely appropriate. That is most obviously true of religious differences; indeed, it might be a mistake to suppose that differences in identity were any less significant for the religious conflicts of the post-Reformation era than they are for the conflicts of our own age (Creppell 2003). Belief and value may also figure importantly in cultural differences. Some might see these differences as making no call upon toleration: there is merely "our way" and "their way." But for others cultural differences might be sites of deep disagreement about what constitutes right conduct, about the status we should accord different sorts of human being, and so on, which present them with the most testing circumstances of toleration.

A final point takes us back to the meaning of "toleration." As we have seen, in orthodox usage toleration entails disapproval or dislike. But consider a person whose society

is multicultural and who resolves not to take exception to the differences she encounters. While those around her respond with suspicion and hostility to the unfamiliar and the alien, she makes a determined effort not to. If she has instincts of aversion or dislike, she keeps them in check. When she finds something odd or unsavory about another's way of life, she assumes that is to be explained by her own lack of understanding rather than by the error or inadequacy of others. How should we describe such a person? "Tolerant" would seem the obvious adjective, and yet here we have someone who does not endure what she disapproves of or dislikes; rather, she simply refrains from disapproval or dislike. Nowadays the term "tolerant" is often used in this more generous sense. For example, when politicians and others aspire to a "tolerant society," their ideal often seems to be a society whose members cease to view one another negatively, rather than one in which they persist with their negativity but put up with one another even so. If we give toleration a meaning that brings it closer to "acceptance," it will clearly be both relevant and desirable even for differences of identity. But there are hazards in this usage. It can imply that merely to disapprove is, in *an*y circumstance, to be intolerant; if we continue to believe that we can have good reason to disapprove along with good reason not to act on our disapproval, that implication will be a step too far.

Related Topics

Locke, Kant, Mill, Liberalism, Pluralism, Multiculturalism, Freedom, Religion in Public Life

References

Addis, Adeno (1997) "On Human Diversity and the Limits of Toleration," in Ian Shapiro and Will Kymlicka (eds), *Ethnicity and Group Rights*, New York: New York University Press.

Apel, Karl-Otto (1997) "Plurality of the Good? The Problem of Affirmative Tolerance in a Multicultural Society from an Ethical Point of View," *Ratio Juris* 10(2): 199–212.

Barry, Brian (1995) *Justice as Impartiality*, Oxford: Clarendon Press.

Brown, Wendy (2006) *Regulating Aversion: Tolerance in the Age of Identity and Empire*, Princeton, NJ: Princeton University Press.

Creppell, Ingrid (2003) *Toleration and Identity: Foundations in Early Modern Thought*, New York: Routledge.

Dworkin, Ronald (2000) *Sovereign Virtue: the Theory and Practice of Equality*, Cambridge, MA: Harvard University Press.

Forst, Rainer (2003) "Toleration, Justice and Reason," in Catriona McKinnon and Dario Castiglione (eds), *The Culture of Toleration in Diverse Societies*, Manchester: Manchester University Press.

Galeotti, Anna Elisabetta (2002) *Toleration as Recognition*, Cambridge: Cambridge University Press.

Gray, John (2000) *Two Faces of Liberalism*, Cambridge: Polity Press.

Heyd, David (2008) "Is Toleration a Political Virtue?" in Jeremy Waldron and Melissa S. Williams (eds), *Toleration and its Limits*, New York: New York University Press.

Hobbes, Thomas (1957) *Leviathan*, ed. Michael Oakeshott, Oxford: Blackwell.

Horton, John (1996) "Toleration as a Virtue," in David Heyd (ed.), *Toleration: an Elusive Virtue*, Princeton, NJ: Princeton University Press.

—— (2011) "Why the Traditional Conception of Toleration Still Matters," *Critical Review of International, Social and Political Philosophy* 14(3): 289–305.

Jones, Peter (2006) "Toleration, Recognition and Identity," *Journal of Political Philosophy* 14(2): 123–143.

—— (2007) "Making Sense of Political Toleration," *British Journal of Political Science* 37(3): 383–402.

—— (2012) "Legalising Toleration: A Reply to Balint," *Res Publica* 18(2): forthcoming.

Kant, Immanuel (1991) *Kant: Political Writings*, ed. Hans Reiss, Cambridge: Cambridge University Press.

Larmore, Charles E. (1987) *Patterns of Moral Complexity*, Cambridge: Cambridge University Press.

—— (1996) *The Morals of Modernity*, Cambridge: Cambridge University Press.

Lauren, John Christian (1999) *Religious Toleration: "The Variety of Rites" from Cyrus to Defoe*, Basingstoke: Palgrave Macmillan.
Locke, John (1968) *A Letter on Toleration*, ed. Raymond Klibansky, Oxford: Clarendon.
Mendus, Susan (1989) *Toleration and the Limits of Liberalism*, Basingstoke: Macmillan.
Mill, John Stuart (1978) *On Liberty*, ed. Gertrude Himmelfarb, Harmondsworth: Penguin.
Nagel, Thomas (1991) *Equality and Partiality*, New York: Oxford University Press.
Nederman, Cary J. (2000) *Worlds of Difference: European Discourses of Toleration, C.1100–C.1550*, University Park, PA: Pennsylvania State University Press.
Newey, Glen (1999) *Virtue, Reason and Toleration: the Place of Toleration in Ethical and Political Philosophy*, Edinburgh: Edinburgh University Press.
Oberdiek, Hans (2001) *Tolerance: Between Forbearance and Acceptance*, Lanham, MD: Rowman and Littlefield.
Paine, Thomas (1989) *Thomas Paine: Political Writings*, ed. Bruce Kuklick, Cambridge: Cambridge University Press.
Phillips, Anne (1999) "The Politicisation of Difference: Does this Make for a More Intolerant Society?" in John Horton and Susan Mendus (eds.), *Toleration, Identity and Difference*, Basingstoke: Macmillan.
Rawls, John (1993) *Political Liberalism*, New York: Columbia University Press.
—— (1999) *The Law of Peoples*, Cambridge, MA: Harvard University Press.
Raz, Joseph (1975) *Practical Reason and Norms*, London: Hutchinson.
Rorty, Richard (1989) *Contingency, Irony and Solidarity*, Cambridge: Cambridge University Press.
Taylor, Charles (1994) "The Politics of Recognition," in Amy Gutman (ed.), *Multiculturalism: Examining the Politics of Recognition*, Princeton, NJ: Princeton University Press.
Walzer, Michael (1997) *Toleration*, New Haven, CT: Yale University Press.

Part VII

APPROACHES

56
SOCIAL EVOLUTION

Gerald Gaus and John Thrasher

It is a matter of dispute how far back evolutionary explanations of social order should be traced. Evolutionary ideas certainly appear in the work of the ancient Greek philosophers, but it seems reasonable to identify the origins of modern evolutionary thinking in the eighteenth-century natural histories of civil society such as Rousseau's *Discourse on the Origin and Foundations of Inequality Among Men* (1750: Pt III), Adam Ferguson's *An Essay on the History of Civil Society* (1767), and Adam Smith's *Wealth of Nations* (1776: Bk III). In these eighteenth-century works, the explanation of current social institutions as an unplanned and generally adaptive development out of earlier and simpler arrangements gained traction. Germany too had a tradition of *Naturphilosophie* employing general evolutionary ideas, as well as Hegelian-influenced thinking on the development of societies. In 1863, four years after Darwin's *On the Origin of Species*, August Schleicher's *Die Darwinsche Theorie und die Sprachwissenschaf*, drew on these traditions as well as Darwin's book to present an evolutionary account of the development of families of languages (Taub 1993), an endeavor that was carried on by a number of scholars in the later part of the nineteenth century.

1. Spencer

Although Herbert Spencer is generally known as a "Social Darwinist" his early evolutionary account predates *Origin of Species*. In an 1857 article in the *Westminster Review* Spencer sketches a comprehensive account of evolution of the inorganic, organic, and human and social realms. "In respect to that progress which individual organisms display in the course of their evolution," Spencer argues, the investigations of "Wolff, Goethe, and von Baer, have established the truth that the series of changes gone through during the development of a seed into a tree, or an ovum into an animal, constitute an advance from homogeneity of structure to heterogeneity of structure" (1857: ¶2). He continues:

> On passing from Humanity under its individual form, to Humanity as socially embodied, we find the general law still more variously exemplified. The change from the homogeneous to the heterogeneous is displayed in the progress of civilization as a whole, as well as in the progress of every nation; and is still going on with increasing rapidity. As we see in existing barbarous tribes, society in its first and lowest form is a homogeneous aggregation of individuals having like powers and like functions: the only marked difference of function being that which accompanies difference of sex. Every man is warrior, hunter, fisherman,

> tool-maker, builder; every woman performs the same drudgeries. Very early, however, in the course of social evolution, there arises an incipient differentiation between the governing and the governed … no sooner does the originally-homogeneous social mass differentiate into the governed and the governing parts, than this last exhibits an incipient differentiation into religious and secular—Church and State; while at the same time there begins to be differentiated from both, that less definite species of government which rules our daily intercourse—a species of government which, as we may see in heralds' colleges, in books of the peerage, in masters of ceremonies, is not without a certain embodiment of its own. Each of these is itself subject to successive differentiations. … Simultaneously there has been going on a second differentiation of a more familiar kind; that, namely, by which the mass of the community has been segregated into distinct classes and orders of workers. While the governing part has undergone the complex development above detailed, the governed part has undergone an equally complex development, which has resulted in that minute division of labour characterizing advanced nations.
>
> (Spencer 1857: ¶¶ 8–9)

This account is not selectionist, in the sense of postulating a filtering mechanism (such as "survival of the fittest") that eliminates, say, societies that are not heterogeneous in favor of those that are. Instead, Spencer conceives of evolution as a law of increasing complexity of all things based on the "all-pervading principle" that "[e]very active force produces more than one change—every cause produces more than one effect" (1857: ¶22). Evolution, then, is a process of increasing complexity over time.

In his 1860 essay on "The Social Organism" Spencer makes much of the idea that societies, like individual organisms, "spontaneously evolved" to form complex systems of interaction characterized by the division of labor (1860: ¶3). For Spencer, the "social organism" is itself subject to evolutionary developments as a separate entity from its constituent individuals, though individual and social organisms have distinct characteristics. Some have taken Spencer's belief in a developing social organism as evidence that Spencer was willing to sacrifice individuals for the sake of the greater social whole, but this is a mistake. In these initial statements of his social evolutionary views, competition between individuals and families for survival is not the driving force of social evolution: it is a manifestation of the inner workings of all organisms toward more complex internal organizations and mutual dependence between the constituent elements. Ultimately, there is nothing very Darwinian in this account. Instead, it is an expression of a universal metaphysical principle characterizing all order, a general theory of development. Moreover, the idea that societies could be understood as organisms with their own evolutionary tendencies was widespread in late nineteenth- and early twentieth-century thought, from Spencer to left-leaning reformers such as L.T. Hobhouse and even to anarchists like Peter Kropotkin (Gaus 1983: ch. 2). Evolution was in the air and developmental thinking can be found in a variety of different fields from Hegel at the beginning of the nineteenth century through Marx, Schopenhauer, and Nietzsche, in philosophy to the British Whig historians such as Trevelyan and Acton, and even into the understanding of the mind itself in the writing of Freud. It was an age obsessed with understanding the progress of history and making sense of the historical laws of development, even independent of the influence of Darwin. Indeed, Darwin can be seen, as one of the most important

and successful of these evolutionary thinkers of the period, but in no sense was his approach unique.

2. Darwin

Darwin, unlike Spencer, was notoriously cautious in applying evolutionary principles to humans, and only does so in a systematic way in *The Descent of Man* in 1871 and, to a lesser extent, in *The Expression of the Emotions in Man and Animals* in 1872. In both works, he provides an evolutionary account of the development and importance of sympathy and moral sense. In a crucial passage he seems to appeal both to direct reciprocity (Darwin 1871: ch. 5) and group selection (Darwin 1871: ch. 3) in explaining the rise of sympathy:

> [a] we are led by the hope of receiving good in return to perform acts of sympathetic kindness to others; and there can be no doubt that the feeling of sympathy is much strengthened by habit. In however complex a manner this feeling may have originated, as it is one of high importance to all those animals which aid and defend each other, it will have been increased, through natural selection; [b] for those communities, which included the greatest number of the most sympathetic members, would flourish best and rear the greatest number of offspring.
> (Darwin 1871: 82)

Darwin's passage raises a theoretical point that should be emphasized. Note that claim [a] is consistent with an account of natural selection that operates only at the level of individual organisms: those who help others when they can expect help in return outcompete non-reciprocating egoists in the struggle for existence. Organism *A*'s act ϕ of helping *B* is to the evolutionary advantage of *A* if the expected benefits from future help by *B* exceed the costs of ϕ to *A* (Henrich and Henrich 2007: 42; Hamilton 1964). So on this analysis sympathy, which can lead one to help another, does not produce (from an evolutionary point of view) genuine altruism, since ϕ is to *A*'s long-term advantage. On the other hand at [b] Darwin appears to be positing a group selection mechanism; communities composed of sympathetic individuals will outperform communities of egoists without sympathy. Here we confront great controversy. Although group selectionist accounts were widely accepted in the nineteenth century and the early part of the twentieth century, in the 1970s the idea was discredited and, indeed, identified as a simple and manifest error. The problem is obvious: even if a group of sympathetic individuals outperform a group of non-sympathetic egoists, a non-sympathetic egoist within a group of sympathetic individuals will outperform the rest of their group. The free-riding non-reciprocator receives all the benefits of living in a sympathetic reciprocating group, but never has to bear any of the costs of reciprocating, in the sense of sacrificing some of its own fitness to help other group members. If such defectors thrive in the group, they will eventually dominate it, outcompeting the reciprocators. In the last few decades, however, evolutionary theorists have reevaluated the possibility of group selection; under certain conditions it is certainly a possibility and could, indeed, be an important factor in evolution (Sober and Wilson 1999; Okasha 2006; Richerson and Boyd 2005; Bowles and Gintis 2011; Wilson 2012). It is important, however, to distinguish between group selection in cultural evolution and group selection in biological evolution: the former is considerably less controversial.

3. The Late Nineteenth-Century Debate: Competitive or Cooperative Evolution?

The great debate in social and political philosophy in the late nineteenth century was not whether Darwin's views were applicable to the study of society, but whether the study of social evolution appropriately stresses competition among individuals within a society, or the cooperative nature of successful societies—ideas, as we have just seen, that can be found side-by-side in Darwin. Although Darwin was always generally cautious in his pronouncements, especially when they involved the application of his theory to human society, his views tended toward the competitive side of the debate. It is important to stress in this regard that Darwin's key idea of the "struggle for existence" was itself taken over from Thomas Malthus's 1798 essay on population, which stressed that social life is always a struggle between increasing population and the constraints on food supply. To give food to a non-producer over a producer was not only to increase the burdens on the food producers in their struggle for existence, but also to handicap society in its efforts to produce as a whole (Malthus 1798: ch. III, ¶6; Ritchie 1891: 2). In a letter Darwin remarks "What a foolish idea seems to prevail in Germany on the connection between Socialism and Evolution through Natural Selection" (F. Darwin 1887: 237). Yet, in the end, Darwin also provides support for a more cooperative reading :

> Important as the struggle for existence has been and even still is, yet as far as the highest part of man's nature is concerned there are other agencies more important. For the moral qualities are advanced, either directly or indirectly, much more through the effects of habit, the reasoning powers, instruction, religion, &c., than through natural selection; though to this latter agency the social instincts, which afforded the basis for the development of the moral sense, may be safely attributed.
>
> (Darwin 1871: 404)

Besides Spencer, who certainly became associated with the competitive interpretation, perhaps the most famous exponent of this view was William Graham Sumner. According to Sumner, "man is born under the necessity of sustaining the existence he has received by an onerous struggle against nature, both to win what is essential to his life, and to ward off what is prejudicial to it" (1860: ¶2). The crux of this argument, though, is essentially Malthusian: our struggle is against nature and only derivatively against each other, as we "are struggling each to win from nature the material goods necessary to support life, and are carrying on this struggle side by side" (Sumner 1881: ¶11). To Sumner, socialism was simply a sentimental philosophy that denied unpalatable reality. For Sumner as for Spencer, the study of sociology was to help us cope with hard reality with which nature confronts us. With Sumner we see an especially clear example of a fairly swift move from an account of social evolution to social policy, which could not negate the inherent competition among individuals.

The late nineteenth-century response by those friendly to socialism was generally not to dismiss the relevance of evolutionary analysis to social and political philosophy, but to insist on the social-cooperative interpretation. In his important *Darwinism and Politics* (1891), D. G. Ritchie insisted that while an individual struggle for material existence might characterize lower levels of evolution, in advanced civilizations social evolution is moved by the selection of ideas as embodied in institutions that make for a coopera-

tive order. Socialism, according to Ritchie, was the sort of arrangement that, in contrast to the *laissez faire* of Spencer and Sumner, is favored in advanced civilizations. L. T. Hobhouse, one of the founders of the "new liberalism," stressed the case for expanded poor relief, unemployment insurance, and old-age pensions. He also took an explicitly cooperative view of social evolution. Like Ritchie he insisted that the "higher" phases of social evolution are guided by a mental evolution—the evolution of a social mind. He argued, "the turning point in the evolution of thought … is reached when the conception of the development of humanity enters into explicit consciousness as the directing principle of human endeavor" (Hobhouse 1911: 155; see also Hobhouse 1901). Hobhouse's conception of the development of societies and thought towards higher stages of cooperation and more reflective consciousness greatly influenced his student, Morris Ginsberg, who advocated evolutionary sociology in the first half of the twentieth century. Ginsberg believed that the outcome of social evolution is that "men have become increasingly conscious of a fundamental unity of purpose and a good common to all mankind" (1931: 223). Although Ritchie sought to make much of his account consistent with a Darwinian orientation, this is not true of many other similar thinkers. For the most part, these cooperative accounts are explicitly non-Darwinian; social evolution is viewed as a principle of development or "growth" in ways that are not radically different from Spencer's pre-Darwinian formulation.

4. F. A. Hayek

Although Ginsberg and a few others championed social evolutionary thought through the 1930s and later, willingness to apply evolutionary analysis to societies and social policy radically dropped off after World War I, and even more dramatically after the rise fascism. The ideas of struggle for existence or survival of the fittest were often depicted as imperialistic or aggressively nationalistic ideas; combined with the association of the social application of Darwinian principles with eugenics, social evolutionary thought was generally thought to be discredited in the early and mid-twentieth century (Hodgson and Thorbjørn 2010: 16). During the 1950s, when the idea of social evolution was deeply unpopular, Hayek started to revive the approach in his University of Chicago interdisciplinary seminar. Participants in the seminar included the geneticist Sewall Wright, himself a group selectionist (Caldwell 2004: 299). This was the beginning of a number of works by Hayek in the 1960s and 1970s, wherein he developed a sophisticated analysis of the relation between complex phenomena, spontaneous orders, and cultural evolution (Gaus 2006). These accounts employed selectionist ideas but were not directly Darwinian.

Hayek repeatedly refers to "the twin ideas of evolution and spontaneous order" (Hayek 1967c: 77; 1978: 250; 1973: 23, 158; 1988: 146). The starting point is Hayek's work on complex phenomena; in contrast to the relatively simple phenomena studied by classical mechanics, biological organisms and social orders are complex systems that give rise to emergent properties, which can only be predicted within very broad ranges (Hayek 1967b). Throughout his long career—and certainly since the 1950s—Hayek's overriding concern was the analysis of the emergent property he called "the order of actions":

> It is the resulting overall order of actions but not the regularity of the actions of the separate individuals as such which is important for the preservation of the group; and a certain kind of overall order may in the same manner contribute

> to the survival of the members of the group whatever the particular rules of individual conduct that bring it about.
>
> (1967c: 68)

Hayek's fundamental claim is that the survival of a society depends on the emergent property of orderly cooperation of different individuals, which has a complex relation to the rules of conduct individuals follow. Some theories of spontaneous order (e.g. economics) explain how the complex order of social actions can be self-organizing and self-maintaining—they are essentially what Hayek calls "models" of the complex phenomenon of the social order of actions, providing general accounts of how the elements relate (Hayek 1967a: 14ff.). But to understand the workings of a complex social order of actions—spontaneously organized and self-regulating complex phenomena—is, still, essentially a static explanation. Hayek is so attracted to evolutionary accounts of the order of actions because they hold out the promise of providing general explanations of the alteration and development of complex orders without needing the assumption that anyone fully understands the working of the order. As Hayek sees it, evolutionary accounts provide the real alternative to design or "constructivist theories" (Hayek 1973: ch. 1). Evolutionary accounts articulate precisely the type of explanations based on the principles of change, rather than the prediction of system states, that are appropriate to complex systems (Hayek 1967b: 31ff.). In biology, Darwinian theory allows us to understand the principles that regulate the development of species, shows us that some developments are outside the possible range of possible values (e.g. that horses will suddenly give birth to winged offspring), but it is unable to generate specific predictions about the future of individuals or species (Hayek 1967b). Hayek believes that the same types of explanations are appropriate in other complex systems such as the economy or society as a whole.

In Hayek's social evolutionary analysis, then, the *explanandum* (that which is to be explained) is the rise and development of an emergent property, namely, the social order of actions. As Hayek says, "the selection process will operate on the order as a whole" (1967c: 71). This social order of actions is what Hayek calls the "Great Society": an overall spontaneous order of adaptations that allows for coordinated action (Hayek 1973: 2ff.). The *explanans* (that which does the explaining) is an evolutionary account whereby the rules and institutions that give rise to this order (i.e., this emergent property) are selected via a competition within and among social orders. The emergent property, we have seen, arises out of a system of rules; therefore, their constituent rules and institutions determine the competition among these social orders as they operate in specific environments. Constituent rules and institutions differentiate social orders; variation in the rules and institutions can provide a competitive advantage in the competition between social orders, leading to selection of a social order of actions with certain sets of rules.

Hayek suggests a number of selection mechanisms, one of which is group survival:

> [t]he rules of conduct have … evolved because the groups who practiced them were more successful and displaced others. They were rules which, given the environment in which men lived, secured that a greater number of the groups or individuals who practiced them would survive.
>
> (Hayek 1973: 18; see also Hayek 1988: 25)

This passage presents difficulties. By simply talking of "groups," and then adding at the close that "greater number of the groups or individuals" survive, Hayek appears to be

directly falling prey to problems of group selection and collective action. It looks as if Hayek is claiming that an individual's chance of survival is maximized if she belongs to a group that maximizes its own chance of survival (Hayek 1967c: 72). But this raises familiar problems of collective action and the rationality of free-riding that we saw in Section 2. If individuals are confronting prisoner's dilemmas, it might be rational for each person not to do the thing that is good for the group, even though this leads to a situation that is disadvantageous for all. One of the lessons of game theory is that what is good for the group might not be rational for any particular person in the group.

Hayek also, however, writes that a person's "thinking and acting are governed by rules which have by a *process of selection been evolved in the society in which he lives*" (Hayek 1973: 11, emphasis added). Understood thus, it looks now as if his project is to explain how each rule (not the order of actions itself) evolved within the society. That this project might rely not only on the rule's ability to produce a competitive order of actions, but the rule's attractiveness to individuals who live under it, is suggested by Hayek's remark that "[t]he competition on which the process of selection rests must be understood in the widest sense. It involves competition between organized and unorganized groups *no less than competition between individuals*" (Hayek 1960: 37, emphasis added). This stress on individual competition and the evolution of *rules* suggests that, instead of a competition between social orders, Hayek has in mind a competition between individuals within a social order that leads to the *selection and evolution of rules*. So we seem to have two evolutionary competitions, pressuring rules from two different directions: (a) A competition exists between social orders that are, as it were, carriers of rules (as individual organisms are carriers of genes). The rules give a social order a certain competitive advantage, but the rules are only selected insofar as they are part of the evolved social order of actions. (b) There is also a competition between individuals and groups within a social order, and *this* competition selects certain rules as conducive to individual/group success.

5. Game Theoretic Modeling of Evolution

Although Hayek and Darwin saw quite clearly the implications of their evolutionary approach to understanding human institutions, they lacked the analytical tools to fully explore their central insights. Hayek's goal, as we have seen, was to develop a dynamic analysis of social evolution that could replace what he considered to be outmoded forms of social theory. Advances in the understanding of biological evolution in the mid-twentieth century were combined with advances in mathematical techniques, most notably in game theory and the modeling of complex, dynamic systems. The first development of the ideas that would become evolutionary game theory were made by, quite possibly, the most important biologist in the twentieth century, R. A. Fisher. His work on sex ratios in *The Genetical Theory of Natural Selection*, a key book in the development of the modern evolutionary synthesis of genetics with natural selection, developed the basic idea of evolutionary equilibrium as a kind of competitive stability (Fisher 1930). Competitive stability could be modeled in game theoretic terms. This insight led to the development of evolutionary game theory.

The irony, as John Maynard Smith points out on the opening page of his monumental *Evolution and the Theory of Games* is that "game theory is more readily applied to biology than to the field of economic behavior for which it was originally designed" (1982: vii). This is because the basic solution concept of traditional game theory, the

Nash Equilibrium, has some drawbacks in terms of modeling individual rational strategic behavior. As Maynard Smith notes, there are good reasons to doubt that humans will behave rationally in the way that traditional game theory suggests (1982: vii). Furthermore, the unrealistic requirements of common knowledge required in traditional game theory make it unsuitable as the basis of an empirical, prediction-based science of human behavior (Sugden 1991).

When applied to a biological context, however, game theory is a powerful tool. Instead of analyzing the rational strategy for an individual agent, we look at stable strategies across populations. This has several implications. *First*, the payoffs to the game which, in traditional game theory, were expressed in the somewhat murky notion of individual utility are transformed into the much more straightforward idea of replication rates. Successful strategies propagate themselves at a higher rate than unsuccessful ones. This solves the perennial problem of interpersonal comparisons of utility by locating payoffs along a single dimension of reproductive success. *Second*, the idea of a mixed-strategy, a key component to the power of the Nash existence theorem (which shows that all finite games have at least one Nash Equilibrium), was never well motivated in the case of individual single-play games. As many have pointed out, it is hard to see what rationally accessible reason a person could have for actually playing mixed-strategy equilibria (Pettit 1996: 291–3). At the population level, however, this problem is solved by understanding mixed strategies as populations divided into different strategies. To have one-third of a population be composed of strategy A agents and two-thirds of the population be composed of strategy B agents makes sense in a way that an individual agent playing strategy A with a one-third probability and B with a two-thirds probability in a single-play game does not. *Third*, the fundamental solution concept of evolutionary game theory, the "evolutionarily stable strategy" (ESS) seems more robust than the various solution concepts in traditional game theory.

The ESS is, again, based on population level dynamics. As Maynard Smith puts it, "evolutionary game theory is a way of thinking about evolution at the phenotypic level when fitness of particular phenotypes depend on their frequency in the population" (1982: 1). In evolutionary game theory, phenotypes are represented as particular strategies. Each strategy has a given payoff in terms of fitness when paired against any other strategy in a population, including itself. A strategy is an ESS if, as Herbert Gintis puts it, "a whole population using that strategy cannot be invaded by a small group with a mutant [phenotype]" (Gintis 2009b: 229). Similarly, a social "phenotype," a set of social norms, rules, or practices, will be evolutionarily stable if a group of mutant innovators cannot move the entire population away from the equilibrium strategy.

To better see this, consider the classic Hawk–Dove game given in Figure 1:

	Hawk	**Dove**
Hawk	–5 –5	0 10
Dove	10 0	5 5

Figure 1 A Specific Hawk–Dove Game

Suppose that Hawks and Doves are phenotypes in the population. A Hawk always battles for territory until either it is injured or its opponent retreats. A Dove engages in display battle: if it meets a Hawk it quickly retreats without injury; if it meets another Dove, there is a 0.5 probability that it will retreat—in no case does it sustain injury. The expected payoff of a Hawk playing another Hawk is –5; the expected payoff of a Hawk meeting a Dove is 10; the expected payoff of a Dove meeting a Hawk is 0; and the expected payoff of a Dove meeting a Dove is 5, with payoffs representing differential reproduction of the phenotype in the next generation. Though the players do not vary their strategies (and so they cannot make "moves" in the sense of traditional game theory), we can understand the population as "moving" in the sense that if being a Hawk has a higher expected payoff than being a Dove, the population will move towards more Hawks and fewer Doves. In evolutionary terms we can think of this as a *replicator dynamic* in which those strategies that tend to have higher average payoffs increase in the population and so displace lower payoff strategies (Skyrms 1996). Neither an all-Hawk nor an all-Dove strategy is evolutionarily stable. For a population of all Hawks the average expected payoff of the population is –5; since a mutant Dove would have a payoff of 0 it would outperform the population average and increase. For a population of all Doves, the average population payoff is 5; a mutant Hawk would receive a payoff of 10, thus again outperforming the average of the Dove population. In Figure 1 an evolutionarily stable equilibrium would be a mixed population evenly split between Doves and Hawks. At that mix the average Hawk and average Dove payoffs are the same (2.5), and so neither population can grow at the expense of the other. This is equivalent to a Nash Equilibrium of mixed strategies.

The basic tools of evolutionary game theory can be extremely powerful, modeling the dynamics of social evolution in ways that Hayek could never have imagined. Indeed, some have gone so far as to argue that evolutionary game theory is essential to a unified understanding of the social and behavioral sciences (Gintis 2009a). The basic hope is that, given this powerful analytical tool, social theorists can model the complex world of social interaction and the development of different social evolutionary equilibria. Much work has already been done along these lines and evolutionary game theory has helped clarify important social practices such as the development of social norms (Bicchieri 2006), property in both animal and human populations (Gintis 2007), and altruism (Trivers 1971) among others.

6. The Co-evolution of Genes and Culture

Another set of powerful formal models of cultural evolution based on population dynamics has been developed by Robert Boyd and Peter Richerson (1985, 2005). Boyd and Richerson's approach seeks "a systematic framework" accounting "for all the processes by which the distribution of beliefs, attitudes and values in a population are modified" (1985: 11–12). The aim is to model cultural evolution in terms of demographic dynamics, accounting for social, ecological and psychological processes. On their analysis the *forces of cultural evolution* can be divided into (a) the inertial, (b) those that induce variation, and (c) those resulting from natural selection (Richerson and Boyd 2005: 68). The *inertial forces* are those that induce towards stability, such as the accurate copying of ideas and models from one person to the next, which accurately replicate cultural forms. In contrast, *cultural factors that induce variation* lead to cultural changes, and can be subdivided into random and decision-making variations. Random

variations include, say, simple errors in copying cultural forms leading to mutations, as well as a cultural version of genetic drift (if a society has some skill that is practiced by a small group and, by chance, this group dies off in a generation, the society will lose that cultural form). Boyd and Richerson place great stress on decision-making factors that induce variation. What they call "guided variation" involves social learning and invention which are successfully reproduced. Some individual, confronting a problem, might discover a superior method that is then copied by others. On their view, however, a great amount of cultural variation stems from "biased transmission," where certain sorts of cultural forms are more likely to be reproduced than others. For example, cultural ideas and practices followed by high-prestige individuals are more likely to be copied, as are more common forms, or simply those that are simpler and easier to remember.

Fundamental to Boyd and Richerson's view is the hypothesis of the co-evolution of culture and genes (see also Henrich and Henrich 2007: ch. 2). The variation of cultural forms is affected by their interaction with natural selection: cultural forms that have advantages in natural selection (say, language) can spread. According to the crucial co-evolution thesis, the relation between natural selection and cultural evolution is reciprocal: natural selection primes us for some culture variations, which might be favored by natural selection, yet those cultural forms change the environment in which natural selection operates.

Boyd and Richerson's account is straightforwardly selectionist, with selection operating at both individual and group levels, on both cultural and biological dimensions. Thus some cultural form (say, a new way of farming), might be selected because individuals who employ it are seen to do better, and so are copied by others, but also because groups that employ the method outperform groups that do not, and so eventually are copied by, or eliminate, other groups (Boyd and Richerson 2005: ch. 11). Furthermore, as we have just seen, a form may spread because it is seen as useful (and so selected by at the cultural level), but it may also spread because it gives those who use it a biological evolutionary advantage (and so selected by natural selection).

An especially important part of the work of Boyd and Richerson, as well as Gintis, Ernst Fehr and others has been the analysis of the evolution of cooperation in large groups (Boyd and Richerson 2005: pt. 3; Gintis, Bowles, Boyd and Fehr 2005; Henrich and Henrich 2007; Bowles and Gintis 2011). In many ways the beginning of rigorous analysis of the evolution of cooperation was Robert Axelrod's (1984) analysis of tit-for-tat in iterated prisoner's dilemmas. Axelrod showed that if we model cooperation in terms of direct reciprocity (each person reciprocates with other individuals who have cooperated with her, but refuses cooperation with those who have not cooperated with her), then strategies such as tit-for-tat can account for the evolution of cooperation among egoists. From the perspective of game theory, Binmore (2005) has effectively shown that the family of possible strategies in equilibrium is *much* larger than tit-for-tat and its near variants. The crucial findings of Boyd, Richerson and others, however, is that direct reciprocity, which places heavy information demands on those who employ it, is not a plausible model for the evolution of cooperation in large groups (Boyd and Richerson 2005: chs. 8–10; Henrich and Henrich 2007: chs. 6–7; Gaus 2011: ch. 2; Bowles and Gintis 2011: chs. 5–6). It seems that the most plausible hypothesis is that the evolution of large-scale human cooperation is best explained by the evolution of "altruistic punishers" or what might be called "rule-following punishers"—individuals who follow norms when most others do so, and who are also willing

to expend resources to punish non-cooperators. Groups of such rule-following punishers outcompete groups composed of more egoist individuals and also resist invasion by defectors (Gaus 2011: ch. 3). This explanation, if correct, provides an answer to the problem that Darwin posed in 1871 about the possibility of morals and cooperation co-evolving.

Despite the power of the co-evolutionary account of cultural evolution proposed by Boyd and Richerson, there is some controversy over what exactly an account of cultural evolution is about. That is, what is culture and how is it transmitted? In many ways, this controversy is the cultural evolutionary version of the traditional "levels of selection" debate in the philosophy of biology. This debate arises because of the abstract nature of the idea of natural selection. This, combined with the hierarchical nature of biological organisms and the complex nature of adaptation, makes it unclear what should be considered the replicator that is selected in certain cases (Okasha 2006: ch. 1). The fuzzy nature of what constitutes culture, norms, principles, and social rules makes this problem even more acute. In his groundbreaking 1976 defense of gene level selection, *The Selfish Gene*, Dawkins first introduced the idea of the "meme" which acts as the analogue to the gene in social evolution (1976: ch. 11). Dawkins insists that cultural evolution is driven by gene-like memes at the genotypic level (see also Blackmore 1999; Distin 2005).

Dan Sperber proposes a model of cultural evolution based on what he calls "epidemiology of representations" (Sperber 1996: 1). Sperber's goal is to build a naturalistic theory of culture and its evolution through an epidemiological model that explains "population-scale macro phenomena … as the cumulative effect of micro-processes" (Sperber 1996: 2). Sperber disagrees with Dawkins and, to a certain extent, with Boyd and Richerson, that cultural evolution can be understood along Darwinian lines. For Sperber, culture is made up of micro-level mental representations in the minds of all of the individuals within a social unit (Sperber 1996: 70–1). Furthermore, "the human mind is susceptible to cultural representations in the same way as the human organism is susceptible to disease" (Sperber 1996: 57). Some cultural representations are more "contagious" and have more staying power than others. The debate between the Sperberian account of cultural evolution and Boyd and Richerson has many dimensions, but the main thread of the controversy is over the unit and process of cultural transmission. Sperber argues that we need to understand the genotypic process of cultural evolution; in this case the mental representations of culture, to understand how culture changes and is maintained. Gene-culture co-evolution is "too slow a process" according to Sperber to explain cultural changes over the last couple of thousand years (Sperber 1996: 114). Instead of co-evolution, we should focus on the psychological nature of the transmission and stability of representations from one person to another. Following a theory of mental representations heavily informed by the work of Chomsky and Fodor, Sperber builds a basically "Lamarckian" account of social evolution from the inside out, that is, from the nature of genotypic mental representations to culture as a whole. Boyd and Richerson, on the other hand, have developed a "Darwinian" model of social evolution that is more interested in explaining phenotypic selection while remaining silent on genotypic causes (Richerson and Boyd 2005: 80–1). Far from privileging one approach over the other, it seems fruitful to pursue both lines of research simultaneously. Either one will prove more explanatorily productive than the other or they will converge; in either case the pursuit will have been useful.

7. Conclusion

The study of the nature of social evolution manifests our desire to understand how we arrived where we are and, on some views, perhaps where we are headed. The contemporary interest in social evolutionary models is not only part of the twenty-first century's resurgent general Darwinism, but a deeper recognition of the complexity of social orders and the search for adequate explanatory approaches. A better understanding of the nature of social evolution is important for political and moral philosophers because it can help draw the limits of the feasible; it shows us the contingency of our particular social and moral institutions, while providing a framework for showing why all societies arrive at types of institutions that perform certain crucial functions.

Much is still unresolved. The debate over competitive vs. cooperative evolution that was started in the nineteenth century is still raging—and this is connected to the larger debate over the levels of selection. As we have seen, there is still controversy over the relation between, and relative importance of, biological and social evolution in analyzing culture. In many ways, then, the field of social evolution is just beginning (once again), with many different promising research programs.

Related Topics

Late Nineteeth- and Early Twentieth-Century British Thought, Rational Choice Theory

References

Axelrod, Robert (1984) *The Evolution of Cooperation*, New York: Basic Books.

Bicchieri, Cristina (2006) *The Grammar of Society: The Nature and Dynamics of Social Norms*, Cambridge: Cambridge University Press.

Binmore, Ken (2005) *Natural Justice*, Oxford: Oxford University Press.

Blackmore, Susan (1999) *The Meme Machine*, Oxford: Oxford University Press.

Bowles, Samuel and Herbert Gintis (2011) *The Cooperative Species: Human Reciprocity and its Evolution*, Princeton, NJ: Princeton University Press.

Boyd, Robert and Peter J. Richerson (1985) *Culture and the Evolutionary Process*, Chicago, IL: University of Chicago Press.

—— (2005) *The Origin and Evolution of Cultures*, Oxford: Oxford University Press.

Caldwell, Bruce (2004) *Hayek's Challenge*, Chicago, IL: University of Chicago Press.

Darwin, Charles (1871) [1981] *The Descent of Man, and Selection Related to Sex*, introd. John Tyler Bonner and Robert M. May, Princeton, NJ: Princeton University Press.

—— (1872) [2002] *The Expression of the Emotions in Man and Animals*, introd. Paul Ekman, Oxford: Oxford University Press.

Darwin, Francis (ed.) (1887) *The Life and Letters of Charles Darwin, Including an Autobiographical Chapter*, London: John Murray, volume 3.

Dawkins, Richard (1976) *The Selfish Gene*, Oxford: Oxford University Press.

Distin, Kate (2005) *The Selfish Meme*, Cambridge: Cambridge University Press.

Ferguson, Adam (1767) [1995] *An Essay on the History of Civil Society*, ed. Fania Oz-Salzberger, Cambridge: Cambridge University Press.

Fisher, Roald J. (1930) *The Genetical Theory of Natural Selection*, Oxford: Clarendon Press.

Gaus, Gerald F. (1983) *The Modern Liberal Theory of Man*, New York: St. Martin's.

—— (2006) "The Evolution of Society and Mind: Hayek's System of Ideas," in *The Cambridge Companion to Hayek*, ed. Ed Feser, Cambridge: Cambridge University Press.

—— (2011) *The Order of Public Reason: A Theory of Freedom and Morality in a Diverse and Bounded World*, Cambridge: Cambridge University Press.

Ginsberg, Morris (1931) "The Concept of Evolution in Sociology," *Proceedings of the Aristotelian Society*, ns 31 (1930–1): 201–224.

Gintis, Herbert (2007) "The Evolution of Private Property," *Journal of Economic Behavior and Organization* 64(1): 1–16.
—— (2009a) *The Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences*, Princeton, NJ: Princeton University Press.
—— (2009b) *Game Theory Evolving: A Problem-Centered Introduction to Modeling Strategic Interaction*, Princeton, NJ: Princeton University Press.
——, Samuel Bowles, Robert Boyd and Ernst Fehr (2005) *Moral Sentiments and Material Interests: The Foundations of Cooperation in Human Life*, Cambridge, MA: MIT Press.
Hamilton, W. D. (1964) "The Genetical Evolution of Social Behaviour I," *Journal of Theoretical Biology* 7: 1–16.
Hayek, F. A. (1960) *The Constitution of Liberty*, London: Routledge.
—— (1967a) "Degrees of Explanation," in F. A. Hayek, *Studies in Philosophy, Politics, and Economics*, Chicago, IL: University of Chicago Press, pp. 3–21.
—— (1967b) "The Theory of Complex Phenomena," in F. A. Hayek, *Studies in Philosophy, and Economics*, Chicago, IL: University of Chicago Press, pp. 22–42.
—— (1967c) "Notes on the Evolution of Systems of Rules of Conduct," in F. A. Hayek, *Studies in Philosophy, Politics, and Economics*, Chicago, IL: University of Chicago Press, pp. 66–81.
—— (1973) *Law, Legislation, and Liberty*, vol. 1: *Rules and Order*, Chicago, IL: University of Chicago Press.
—— (1978) "Dr. Bernard Mandeville," in F. A. Hayek, *New Studies in Politics, Economics and the History of Ideas*, London: Routledge, pp. 249–266.
—— (1988) *The Fatal Conceit*, ed. W. W. Bartley III, Chicago, IL: University of Chicago Press.
Henrich, Natalie and Joseph Henrich (2007) *Why Humans Cooperate*, Oxford: Oxford University Press.
Hobhouse, L. T. (1901) *Mind in Evolution*, London: Macmillan.
—— (1911) *Social Evolution and Political Theory*, New York, Columbia University Press.
Hodgson, Geoffrey and Thorbjørn Knundon (2010) *Darwin's Conjecture*, Chicago, IL: University of Chicago Press.
Malthus, Thomas Robert (1798) *An Essay on the Principle of Population, as it Affects the Future Improvement of Society, with Remarks on the Speculations of Mr. Godwin, M. Condorcet, and Other Writers*, London: J. Johnson.
Okasha, Samir (2006) *Evolution and Levels of Selection*, Oxford: Oxford University Press.
Pettit, Philip (1996) *The Common Mind: An Essay on Psychology, Society, and Politics*, Oxford: Oxford University Press.
Richerson, Peter J. and Robert Boyd (2005) *Not by Genes Alone: How Culture Transformed Human Evolution*, Chicago, IL: University of Chicago Press.
Ritchie, D. G. (1891) *Darwinism and Politics*, London: Swan Sonnenschein.
Rousseau, Jean Jacques (1750) [1997] *Discourse on the Origin and Foundations of Inequality Among Men*, ed. & trans. Victor Gourevitch, Cambridge: Cambridge University Press.
Skyrms, Brian (1996) *Evolution of the Social Contract*, Cambridge: Cambridge University Press.
Smith, Adam (1776) [1976] *An Inquiry Into the Nature and Causes of the Wealth of Nations*, ed. R. H. Campbell, A. S. Skinner and W. B. Todd, Oxford: Clarendon Press.
Smith, John Maynard (1982) *Evolution and the Theory of Games*, Cambridge: Cambridge University Press.
Sober, Elliot and David Sloan Wilson (1999) *Unto Others: The Evolution and Psychology of Unselfish Behavior*, Cambridge, MA: Harvard University Press.
Spencer, Herbert (1857) [1891] "Progress: Its Law and Cause," in *Essays: Scientific, Political, and Speculative. Library Edition, containing Seven Essays not before republished, and various other Additions*, London: Williams and Norgate.
Sperber, Dan (1996) *Explaining Culture: A Naturalistic Approach*, Oxford: Blackwell Publishers.
Sugden, Robert (1991) "Rational Choice: A Survey of Contributions from Economics and Philosophy," *The Economic Journal* 101(407): 751–785.
Sumner, William Graham (1860) [1914] "The Challenge of Facts," in William Graham Sumner, *The Challenge of Facts and other Essays*, ed. Albert Galloway Keller, New Haven, CT: Yale University Press.
—— (1881) [1919] "Sociology," in *War and Other Essays*, ed. Albert Galloway Keller, New Haven, CT: Yale University Press.
Taub, Liba (1993) "Evolutionary Ideas and 'Empirical' Methods: The Analogy between Language and Species in Works by Lyell and Schleicher," *The British Journal for the History of Science* 26: 171–193.
Trivers, Robert (1971) "The Evolution of Reciprocal Altruism," *The Quarterly Review of Biology* 46(1): 35–57.
Wilson, Edward (2012) *The Social Conquest of the Earth*, New York: Liveright.

57

THE PRAGMATIST PROJECT IN POLITICAL PHILOSOPHY

Cheryl Misak

1. Introduction

Pragmatism originated in Cambridge Massachusetts in the late 1860s in the Metaphysical Club—a short-lived reading group whose members included Oliver Wendell Holmes, William James, and Charles Sanders Peirce. Exciting ideas about science and philosophy were already very much in the air in Europe and they most certainly had drifted across the Atlantic. August Comte's positivism was all the rage in France from the mid-1880s. Darwin's *On the Origin of Species* was published in 1859 and it and its implications were being intensely discussed by philosophers. By the time the fledgling pragmatists started to talk to each other, the thinking person was already worrying about the problematic relationship between science and God. Science seemed to entail the abandonment of the worldview that had God and religious absolutes at its center.

The early pragmatists lined up fully behind the new scientific worldview, but one of the distinguishing features of their approach was the attempt to find a way to preserve a place for value within it. They strongly tended toward what we today call "holism"—they took their view of truth and inquiry to encompass science, mathematics, logic, religion, ethics, law, and politics. Unlike their empiricist predecessors (think of David Hume, for instance), they fenced off no realm of inquiry from the principles they set out.

Another distinguishing feature of their approach was to put practice at the center of philosophy. As Peirce stated the pragmatic maxim, we must look to the upshot of our concepts in order to understand them (*CP* 5: 4). Pragmatism is committed to "taking a look," as Ian Hacking has more recently put it—to keeping philosophy connected to first order inquiry, to real examples, to action, and to real-life expertise (2007: 36).

If we bring together these two distinguishing features, we can see that pragmatists are naturalists in that they want their explanations and ontology down-to-earth (natural as opposed to supernatural). They require philosophical theories and norms to arise from our practices. But they reject any naturalism that gives ontological priority to matter or physicality—they want to consider whether value, generality, chance, etc. might be part of the natural world.

One way this orientation towards practice manifests itself is that pragmatists hold that the theory of truth must be driven by what happens in inquiry. They argue that

truth is not a relationship between our beliefs and the believer-independent world but, rather, is the best we human inquirers could do.

Given the commitment to take ethical, legal, and political discourse to fall under our cognitive scope, the pragmatist view of these matters, at its very heart epistemic. It treats ethics, politics, and law as kinds of deliberation or inquiry. Like other kinds of inquiry, these should gather as much evidence as they can. But that evidence will not look exactly like evidence in science. In political inquiry, for instance, the input or voice of the people seems as important as the input or voice of experts. Thus, pragmatism seems to have a built-in argument for democratic politics being the best kind of inquiry.

The overarching issue for pragmatism is Kant's problem. How can we make sense of our standards of rationality, truth, and value as being genuinely normative or binding while recognizing that rationality, truth, and value are profoundly human phenomena? Attempts to answer this question shape the great debate within pragmatism. That debate is one between those who suggest (or whose view entails) that there is no truth or objectivity or standards to be had anywhere, and those who take pragmatism to promise a naturalist position that preserves our cognitive aspiration to getting things right.

On the one side of the debate we have Richard Rorty and his classical predecessors (William James and sometimes John Dewey) holding that there is no truth at which we might aim—only agreement within a community or what works for an individual or what is found to solve a problem. Sometimes Rorty goes as far as claiming that truth and objectivity are nothing more than what our peers will let us get away with saying (1979: 176). He thinks that no inquiry aims at the truth, no methodology is privileged, and nothing about pragmatism speaks for or against democracy.

On the other side of the divide, we have those who think of pragmatism as rejecting an ahistorical, transcendental, or metaphysical theory of truth, but nonetheless being committed to doing justice to the objective dimension of human inquiry—to the fact that those engaged in deliberation and investigation take themselves to be aiming at getting things right, avoiding mistakes, and improving their beliefs, theories, and societies. On this more objective kind of pragmatism, which emanates from Peirce, the fact that our inquiries are historically situated does not entail that they lack objectivity. Neither does the fact that standards of objectivity themselves come into being and evolve. The trail of the human serpent is over everything (to use James' phrase), but (as James himself may or may not have seen) this does not toss us into a sea of arbitrariness, where there is no truth or where truth varies from person to person and culture to culture.

These more objective pragmatists argue that morals and politics, like science, aim at the truth or at getting things right and that the best method for achieving this aim is a method they sometimes call the scientific method (Peirce) or the method of intelligence (Dewey)—what would now be termed deliberative democracy. These pragmatists offer an argument for democracy that appeals to the quality of the decisions supplied by the democratic process. Why should we value decisions that are the products of voting after open debate rather than decisions made through private decision-making and then voting; rather than decisions made through bargaining; rather than decisions made via the elimination of those who disagree with us? We should value them because deliberative democratic inquiry is more likely to give us true or right or justified answers to our questions. As Peirce put it, science, inquiry, and rationality are matters of getting your beliefs in line with experience, evidence, and reasons in an ongoing community project. In our efforts to understand reality "each of us is an insurance company" (*CP* 2: 270). It is foolish to seek one's "own private wisdom" or believe you are "the sole inventor"

of your ideas. To do so is to be like someone who denies needing insurance coverage because he believes he has complete control over all future outcomes without the cooperation of anyone else (*CP* 2: 271).

In contrast, the more subjectivist pragmatists offer an account of politics on which all there is are local practices of inquiry, with nothing to be aimed at and nothing to be said in favor of one method or another.

2. C.S. Peirce (1839–1914) and his Followers

The founder of pragmatism, C.S. Peirce, did not think very hard about morals and politics, and when he did, he often said knee-jerk conservative things. I have argued, however, that his view of truth and inquiry holds real promise for moral and political philosophy (Misak 2000; see also Talisse 2007).

Peirce offers us an anti-foundationalist account of truth and inquiry. Inquiry "is not standing upon the bedrock of fact. It is walking upon a bog, and can only say, this ground seems to hold for the present. Here I will stay till it begins to give way" (*CP* 5: 589). When it gives way, the ground merely shifts, rather than opens up underneath us, as a Cartesian might think. We can doubt one belief and inquire, but we cannot doubt all of our beliefs and inquire. Some beliefs have to be held constant, despite the fact that they too are fallible.

Nonetheless, inquiry aims at the truth and gives rise to genuine norms and standards. Here is the key passage in "The Fixation of Belief," one of the few articles this vastly underemployed and underappreciated philosopher managed to publish. It captures the idea that inquiry is not operating on arbitrary whims. Peirce thought his friend James too often suggested the opposite. A method of inquiry is a "failure" if it:

> makes of inquiry something similar to the development of taste … taste, unfortunately, is always more or less a matter of fashion … [And] I cannot help seeing that … sentiments in their development will be very greatly determined by accidental causes. Now, there are some people, among whom I must suppose that my reader is to be found, who, *when they see that any belief of theirs is determined by any circumstance extraneous to the facts, will from that moment not merely admit in words that that belief is doubtful, but will experience a real doubt of it, so that it ceases to be a belief.*
>
> (*W*3: 253; 1877 emphasis added)

Peirce's argument is that beliefs are such that they resign in the face of recalcitrant experience or in the knowledge that they were put in place by a method that did not take experience seriously. He thinks, though, that recalcitrant experience can take many forms. It can be had in diagrammatic contexts; it can be a clash with what we feel to be right; it can be a clash with what others feel to be right. He distinguishes between two kinds of experience. External facts are simply those that are "ordinarily regarded as external while others are regarded as internal" (W2: 205; 1868). The inner world may exert a comparatively slight compulsion upon us, whereas the outer world is full of irresistible compulsions. Nonetheless, the inner world can also be "unreasonably compulsory" and have "its surprises for us" (*CP* 7: 438). Experience, for Peirce, is a brute, compelling or forceful surprise and that surprise can be that you find out, for instance, how sexist comments that you thought were innocuous have a negative impact on young women.

Inquiry, that is, can be conducted in morals and politics. Peirce says that, as with any kind of inquirer, the ethical deliberator might be hesitant to revise her beliefs and this hesitation can be justified. But it is not always justified: "Like any other field, more than any other [morality] needs improvement, advance … But morality, doctrinaire conservatist [*sic*] that it is, destroys its own vitality by resisting change, and positively insisting, This is eternally right: That is eternally wrong" (*CP* 2: 198). Moral and political judgments, Peirce thinks, are matters for inquiry. They are revisable in light of experience. We take our body of background belief for granted, until the course of experience weighs in against it.

By the very activity of engaging or arguing with others, and by the very activity of believing or asserting something, we commit ourselves to the idea that we aim at getting the right answer. For Peirce, the right answer is that which would forever stand up to reasons and evidence; so it follows from his view that we also commit ourselves to the idea that we must expose our beliefs to all the available and relevant reasons and evidence. If we are to have genuine beliefs that will stand up to the force of experience (if we are to have beliefs which aim at being true or right), then we need to adopt a method that exposes our beliefs to experience. We need to adopt a method that takes the views and experiences of others seriously. We might as well call it the method of democracy. When we aim at getting things right and when we believe and assert, we buy into the idea of deliberation and we buy into the democratic idea that all views and experiences must be part of that deliberation.

3. Oliver Wendell Holmes (1841–1935)

One can detect a strong current of pragmatism running through the legal thought of Holmes: the ideas hammered out in the Metaphysical Club animated his view of law. Holmes is interested in the common law—as Frederic Kellogg puts it, the bottom-up theory of law (2007: 19). In excellent pragmatist fashion, Holmes describes law as starting with cases "and only after a series of determinations on the same subject matter" does it come "by … induction to state the principle" (1870 [1995]: 212). Law is not something that is set in stone or in statute. It is a growing, evolving, ongoing enterprise. It is an enterprise of inquiry, starting from precedent and then driven by experience, conflict, and unanticipated problems. Law, for Holmes, evolves in a fallible way, where doubt, conflict, and disputes about what the law is are resolved under the force of experience.

On this pragmatist conception of legal inquiry, precedents, legislation, or moral principles are not immutable truths. In the first few lines of *The Common Law* (1882), Holmes gives us his pragmatist account of what law is: "The life of the law has not been logic: it has been experience" or "the felt necessities of the time" (1882: 1). The law does not consist of a fixed body of doctrines and syllogisms derived from them but, rather, it is an organic structure that has grown up in response to experience. He says that all theories that consider the law "only from its formal side" are failures (1882: 36–7). Whatever code or set of principles or statutes might be adopted, "[n]ew cases will arise which will elude the most carefully constructed formula" and will have to be reconciled (1870 [1995]: 213).

Holmes thinks that courts are engaged in the business of inquiry. Judges are engaged in a search for the best answer we can come to given the time in which we are living and the circumstances in which we find ourselves. The "secret root from which the law

draws all the juices of life," Holmes argues, is "the consideration of what is expedient for the community concerned" (1882: 35). Laws are good if they fit with the values of the community, as determined by judges who are looking to experience to decide cases.

Here we feel pressing in on us the questions that animate the great debate within pragmatism. Do validity, rationality, truth, and other normative concepts turn on expediency, individual or community will and desire, and hard social facts? Or do they turn on responsiveness to something less subjective? Is the aim of our inquiries to get answers that satisfy current, local, interests or is it something more? In the following passage, it might seem that Holmes thinks that there is nothing more:

> It is perfectly proper to regard and study the law simply as a great anthropological document. It is proper to resort to it to discover what ideals of society have been strong enough to reach that final form of expression, or what have been the changes in dominant ideals from century to century. It is proper to study it as an exercise in the morphology and transformation of human ideas.
>
> (1889 [1952]: 212)

Here he seems to argue that in the absence of any principled or certain way of settling political and legal indeterminacy and conflict, security and stability are best served by taking the people's will as what is right (Dyzenhaus 1997).

But as Kellogg (2007) argues, Holmes does not mean that the values that inform law are simply those that happen to be in force in the community. Indeed, the way Holmes incorporates background belief and values into his conception of legal inquiry is as subtle as Peirce's.

We look to the law to enable us to predict how our behavior will be received. When conflicts arise, judges resolve them in an inquiry not unlike scientific inquiry. Science and law are experience-driven enterprises, with background beliefs and cultural norms having real force. But nonetheless, there are external standards. Motives and arguments are *offered* to judges, for their consideration. Law proceeds via "judgment as to the relative worth and importance of competing legislative grounds." (1897 [1952]: 181). Judges, that is, have to use their judgment to weigh the claims of custom or background belief about what is right and wrong. "The law," Holmes says, is the "witness and the external deposit of our moral life. Its history is the history of the moral development of the race. The practice of it, in spite of popular jests, tends to make good citizens and good men" (1897 [1952]: 170). The history of law is a history of development, on which we improve our views about what is good. It is not simply a record of what happened to be thought good.

We have seen that for Peirce, external standards are found by considering whether our beliefs abide by the force of experience—or whether, instead, they are determined by accidental circumstances. Holmes and others who speak about the legal and political domain have an especially difficult task in giving shape to what counts as non-accidental. It is hard to see how there can be non-accidental standards of law and politics if one resists the temptation, as does Holmes, to appeal to God, to Reason, or to some internal logic. The pitfall for the naturalist or pragmatist is that we can make sense of only one kind of standard—what the community happens to believe.

As keenly as Peirce, Holmes pays attention to what makes a belief *well*-settled. In Holmes' words, "a well-settled legal doctrine embodies the work of many minds, and has been tested in form as well as substance by trained critics whose practical interest

it is to resist it at every step" (1870 [1995]: 212). A well-settled belief is not simply the belief that fits with the spirit of the times. External standards are employed by those many minds and those many trained critics. For instance, one external standard is found by considering what a reasonable or prudent person, taking into account the weight of experience, would think. If a workman on top of a house tosses a beam into a busy street, he should have foreseen that the beam might hit someone. Hence that standard applies to his behavior (Holmes 1882: 55–6). This is the kind of thing that looking to external or non-accidental circumstances amounts to in the legal domain. In the same vein, we find Holmes gesturing at an aim of our legal inquiries:

> The truth is, that the law is always approaching, and never reaching, consistency. It is forever adopting new principles from life at one end, and it always retains old ones from history at the other … It will become entirely consistent only when it ceases to grow … However much we may codify the law into a series of seemingly self-sufficient propositions, those propositions will be but a phase in a continuous growth.
>
> (1882: 36–7)

The aim of law would be reached when disputes no longer arose. Like Peirce, Holmes does not think that this is a likely possibility. Peirce would call it a "regulative assumption"—something that we have to go on in order to make sense of proceeding with our inquiries.

4. John Dewey (1859–1952) and his Followers

John Dewey was deeply engaged in the pressing political questions of his day—debates about communism, reform in immigrant housing and in education, etc. He almost always took the stance of a progressive and liberal reformer. As a student, Dewey spent two years in the new graduate program at Johns Hopkins, where he coincided with Peirce's brief tenure there. He was decidedly not keen on Peirce's technical approach and was on an intellectual collision course with his irascible teacher. Peirce is careful in reviews and in public to commend Dewey as a fellow pragmatist, but in letters and in private he was extraordinarily harsh. He wishes that Dewey was more rigorous (*CP* 8: 239, 8: 240).

Dewey, too, was absorbed by the questions of how we are to think of the world, given that *we* bring so much to it; how we are to understand truth, given that we cannot have unfettered access to the subject matter of our beliefs and theories. He initially thought that the answers to these questions were to be found in Hegel, but as his attempts to hammer out a Hegelian solution faltered, he moved gradually to a more resolute pragmatism.

Like the other pragmatists, Dewey's over-riding mission is to encourage scientific thought in all branches of philosophy. He wants to offer a "unified theory of inquiry"—a single way of thinking about how we resolve problematic situations in science, ethics, politics, and law (*LW* 12: 102; 1938). "The problems of men" must be brought under the sweep of science or inquiry (*MW* 5: 10; 1908). Again with his pragmatist predecessors, Dewey argues that any search for absolute truth is not merely bound to be fruitless, but it in fact isn't what we take ourselves to be seeking. No one investigating a particular problem aims for that. Rather, the investigator aims at "security"—at a reliable solution to the problem at hand (*LW* 4: 3f.; 1929). In ethics and politics, just as in science, we aim to transform a doubtful or problematic situation into one in which "confusion and

uncertainty" are resolved (*MW* 14: 144; 1922). Faith in absolutes is to be replaced by "nature, including humanity, with all its defects and imperfections … as the source of ideals" (*LW* 4: 244; 1929).

Dewey argued that democracy walks hand in hand with the scientific method. Democracy is the use of the experimental method to solve practical problems; it is an application of "cooperative intelligence" or inquiry (*LW* 13: 187; 1939); it is the space in which we can "convince and be convinced by reason" (*MW* 10: 404; 1916). Democracy, he suggests, is a precondition of our practices in every domain of inquiry, from physics to politics. We experiment, reflect, and discuss. This way of inquiring—this scientific method—requires the unimpeded flow of information and the freedom to offer and to criticize hypotheses (*EW* 3: 33; 1895, *LW* 11: 375; 1936).

Although we have seen that Peirce also thinks that these principles are essential to the scientific method, he has two objections to Dewey's thought. The first is to the residual Hegelianism present in his metaphysics. Dewey's account of inquiry takes the situation itself to have the doubt, not the inquirer: "It is the *situation* that has these traits. *We* are doubtful because the situation is inherently doubtful" (*LW* 12: 109; 1938). But how can it be, we might do well to wonder, that a situation itself can be doubtful? Second, Peirce objects to Dewey's idea that knowledge, logic, and reality can be understood by looking at how human beings think. Ernest Nagel, one of Dewey's best graduate students, puts the point well: "There is perhaps only a hair line which divides a sound application of the principle of contextual analysis from a commission of the genetic fallacy"—from the accusation that Dewey's philosophy "confounds questions of validity and logical order with questions of origin and development" (Nagel 1939: 578). Peirce thinks that Dewey steps over that line.

Dewey's critics lined up to make this kind of objection throughout the great span of his writing life and Dewey struggled mightily to try to answer them. One attempt has him arguing that moral problem or situation is thus:

> Which shall he decide for and why? The appeal is to himself; what does *he* really think the desirable end? What makes the supreme appeal to him? What sort of an agent, a person, shall he be? This is the question finally at stake in any genuinely moral situation: what shall the agent *be*?
>
> (*MW* 5: 194; 1908)

But on this view, how am I to make sense of the ideas of accuracy, error, getting a better belief, making a mistake, etc.? Dewey suggests that we can tell when a conclusion is misguided because the result will be internal conflict, discomfort, and "compunctions" (*EW* 4: 298; 1894). We can query our desires, aims, and conceptions of ourselves once we see how they lead to disharmony. Jennifer Welchman articulates the problem with this view: "An act that is 'right' because it is true to myself may still be wrong because the sort of self I am is a wrong or bad self" (1995: 109).

At other times Dewey focuses on the idea of a community and its growth or improvement. He says:

> The problem is to extract the desirable traits of forms of community life which actually exist, and employ them to criticize undesirable features and suggest improvement. Now in any social group whatever, even in a gang of thieves, we find some interest held in common, and we find a certain amount of interaction

> and cooperative intercourse with other groups. From these two traits we derive our standard. How numerous and varied are the interests which are consciously shared? How full and free is the interplay with other forms of association?
>
> (*MW* 9: 89; 1916)

A gang of thieves exhibits a "partial and distorted" "education," as their shared interests are not numerous and varied and they have an inevitably limited harmonious interplay with other groups. Democratic and liberal forms of association are much better, as they maximize our abilities to expand our shared interests and to develop our society's possibilities:

> A democracy is more than a form of government; it is primarily a mode of associated living, of conjoint communicated experience. The extension in space of the number of individuals who participate in an interest so that each has to refer his own action to that of others, and to consider the action of others to give point and direction to his own, is equivalent to the breaking down of those barriers of class, race, and national territory which kept men from perceiving the full import of their activity.
>
> (*MW* 9: 93; 1916)

The aim of democratic deliberation is to "convert strife into harmony, monotony into a variegated scene, and limitation into expansion. The converting is progress, the only progress conceivable or attainable by man. Hence every situation has its own measure and quality of progress, and the need for progress is recurrent, constant" (*MW* 14: 195; 1922). In the 1920 *Reconstruction in Philosophy* he puts it thus:

> [T]he process of growth, of improvement and progress, rather than the static outcome and result, becomes the significant thing. Not health as an end fixed once and for all, but the needed improvement in health—a continual process—is the end and good. The end is no longer a terminus or limit to be reached. It is the active process of transforming the existent situation. Not perfection as a final goal, but the ever-enduring process of perfecting, maturing, refining is the aim in living. … Growth itself is the only moral "end."
>
> (*MW* 12: 181; 1920)

But as Matthew Festenstein says, Dewey's criteria for these aims

> appear to be undefined in his philosophy, or defined only in notoriously vacuous ways: as what is conducive to "growth," to the coordination of activities and interests, or to consummatory experience. Rational or "intelligent" agency, it seems, is viewed as instrumental and goal-directed, but the goals to which it is or should be directed have been left out of the picture of inquiry and practical judgment.
>
> (2008: 89)

Sidney Hook, another of Dewey's students, put it thus: growth is not in itself a good. It can be dangerous—"we sometimes find ourselves wishing not only that something wouldn't grow but that it had not come to be" (1977: 12).

The difficulties do not go away for Dewey. For it is not clear that he has anything to say to those who think that the height of self-realization is to forge a homogenous community, by elimination of the other. What do we say about those who assert that the good life centers around a violent nationalism, for instance? Dewey's often excellent and interesting analyses of democracy are ineffective against those who would argue that they are deeply uninterested in living among, or associating with, a minority class, race or different territorial group and would much prefer to do away with them. Morton White puts the problem clearly. On Dewey's view, the "most obnoxious ends and means" could be defended by a unified community (White 1957: 201). Festenstein articulates it thus: "Whatever Dewey's personal moral or political commitments … his experimentalism is vulnerable to appropriation by whatever social forces are most powerful" (2008: 90).

In a letter to Scudder Klyce in 1915, Dewey responds to the charge:

> [M]y only claim is that the objection brought against it is true. I have not given or tried to give any "solutions." But it doesn[']t seem to have occurred to the objectors that to say that moral life is a s[e]ries of problems and that morality *is* their solution as they arise would naturally preclude me from proffering solutions.
>
> (*CJD* [03522]: 1915.05.06)

What we have is a series of problems and what we do is try to solve them. There is not more to say than that. It is here that Richard Rorty's view that we do not aim at truth or rightness, but merely at solidarity, finds its footing in Dewey.

Dewey sees the objection to this line of thought: if we are merely constantly solving problem after problem, "learning and then relearning the meaning of our active tendencies," then "[d]oes this not reduce moral life to the futile toil of a Sisyphus who is forever rolling a stone uphill only to have it roll back so that he has to repeat his old task?" (*MW* 14: 144; 1922). Dewey answers that moral life is like this only from the perspective of someone looking for fixed definitive answers to our questions. It is not futile from the perspective of one who sees that "continual search and experimentation to discover the meaning of changing activity, keeps activity alive, growing in significance" (*MW* 14: 144–5; 1922). That is, we need to see that there is no certainty to be had and there is no set meaning or purpose to life. Deliberation goes on and on and the aim is "learning the meaning of what we are about and employing that meaning in action" (*MW* 14: 194; 1922).

But the problem of adjudication will not go away. The Nazi or the child molester might find meaning in what we want to say are vile and odious acts. Without some kind of check that goes beyond the individual and beyond the current, local, possibly homogenized community, nothing critical can be said.

5. A Sophisticated Pragmatist Political Philosophy

Much work has been done on trying to take the insights at the heart of pragmatist political philosophy and resolve the problems that faced Dewey. I have made one attempt by leveraging Peirce's view of truth, on which there is something that we are aiming at. Despite the fact that we can never be certain that we have it, we are aiming at belief that would forever stand up to experience and deliberation. Hence, we need to get as

much information as we can and that requires taking the views of others seriously—it requires a democratic method of inquiry in morals and politics. Robert Talisse (2002, 2007) has argued along similar lines, taking Peirce's "Fixation of Belief" to have shown that there are norms intrinsic to belief as such—that in believing we take ourselves to be responding appropriately to the relevant reasons and evidence. From this, he develops a "folk epistemology," a rough characterization of the basic norms that govern our lives as believers and which support a case for democratic politics. Only under democratic social conditions are we are able to make good on the folk epistemic norms. Elizabeth Anderson (1993, 2006a, 2006b) has reinterpreted Dewey's view, developing an account of democracy that takes seriously the idea there is a distinction between what a community or individual happens to value and what is, in fact, valuable.

All of these new pragmatists follow both Peirce and Dewey (and, although less self-consciously, Holmes) in taking ethical, political, and legal inquiry to come under our cognitive capacity. Our beliefs in these domains can be warranted or not warranted, as the case may be. These new pragmatists also follow their predecessors in taking thought experimentation to be crucial. Experiment, Dewey says "is a process of tentative action: we 'try on' one or another of the ends; imagining ourselves actually doing them, going, indeed, in this make-believe action just as far as we can without actually doing them" (*EW* 4: 251; 1894). We can learn something in thought experimentation. We propose potential solutions to the problems that press themselves upon us; we try to predict the consequences of those solutions; and we ask whether our reactions to those consequences would be positive or negative and whether other people's reactions would be positive or negative. We can learn something about ourselves and, if we try on other people's shoes, as it were, we can learn something about those others. It is, of course, best to involve as many different kinds of others in our deliberations in order to really discover what their reactions are. Given that moral and political inquiry is largely about our relations with others, these are important things to learn.

Anderson's work is an excellent example of this new pragmatist attempt to invoke genuine normativity, while staying within the naturalist framework—while staying in what is altogether human ethical and political deliberation. She notes that when I judge something as valuable I judge that it is properly valued, not that I happen to value it on some occasion or another. In ethics and politics and law we are interested in what *is* right or wrong, valuable or worthless, not what *seems to me* or even *what seems to most people* to be right or wrong, valuable or worthless. Nonetheless, "emotional experiences have a special tie to normative force. Feelings of admiration are *evidence* of admirability, as desire is evidence of desirability. Of course, such feelings are defeasible" (Anderson 2006b: 6). We take the experience of finding ourselves resenting *x* or admiring *y* as prima facie evidence for the truth of corresponding ethical judgments. What we learn is what options are attractive, repugnant, etc. to us. Then we rationally weigh and evaluate those feelings, sometimes finding that other reasons defeat them. Our judgments, put in place largely through experience, can be subject to disillusionment through further experience, taking into account the experience of others, reflection, and argument.

That is, the pragmatist's view is that we do not know what we value or what is valuable until we engage in inquiry (see Welchman 1995: 189). As Alan Ryan (1995: 337) puts it, at the heart of Dewey's view, even if he did not always see it, is that our moral judgment is just that: *judgment*. We consider, reflect, get more information, and scrutinize our desires. We ask whether what we desire really is desirable. We open up the gap between the two. We discover things, not just about ourselves, but about what is right

and wrong, through inquiry. This, I submit, is a kind of merging of the Peircean and Deweyan brands of pragmatism and it is the future of pragmatist political philosophy.

Related Topics

Democracy, Discourse Theory

References

Anderson, Elizabeth (1993) *Value in Ethics and Economics*, Cambridge, MA: Harvard University Press.

—— (2006a) "The Epistemology of Democracy," *Episteme* 3: 8–22.

—— (2006b) "Replies to My Critics," *Symposia on Gender, Race and Philosophy*, (www.mit.edu.sgrp), 2/1.

Dewey, John (1871–1952) *The Correspondence of John Dewey*, 3 vols., ed. Larry A. Hickman, Carbondale, IL: Southern Illinois University Press. (Cited as *CJD* volume number: year, letter number.)

—— (1972) *The Early Works of John Dewey, 1882–1898*, 5 vols., ed. Jo Ann Boydston, Carbondale, IL: Southern Illinois University Press. (Cited as *EW* volume number: page; original year of publication.)

—— (1978) *The Middle Works of John Dewey, 1899–1924*, 15 vols., ed. Jo Ann Boydston, Carbondale, IL: Southern Illinois University Press. (Cited as *MW* volume number: page; original year of publication.)

—— (1985) *The Later Works of John Dewey 1925–1953*, 17 vols., ed. Jo Ann Boydston, Carbondale, IL: Southern Illinois University Press. (Cited as *LW* volume number: page; original year of publication.)

Dyzenhaus, David (1997) "Holmes and Carl Schmitt: An Unlikely Pair?" *Brooklyn Law Review* 63: 165–88.

Festenstein, Matthew (2008) "John Dewey: Inquiry, Ethics, and Democracy," in C. Misak, ed., *The Oxford Handbook of American Philosophy*, Oxford: Oxford University Press.

Hacking, Ian (2007) "On Not Being a Pragmatist: Eight Reasons and a Cause," in C. Misak, ed., *New Pragmatists*, Oxford: Oxford University Press.

Holmes, Oliver Wendell Jr. (1870 [1995]) "Codes, and the Arrangement of the Law," in S. Novick, ed., *The Collected Works of Justice Holmes*, Chicago, IL: University of Chicago Press.

—— (1882) *The Common Law*, London: MacMillan.

—— (1889 [1952]) "Law in Science and Science in Law," in *Collected Legal Papers*, New York: Harcourt, Brace, and Company.

—— (1897 [1952]) "The Path of the Law," in *Collected Legal Papers*, New York: Harcourt, Brace, and Company.

—— (1920 [1952]) *Collected Legal Papers*, New York: Harcourt, Brace, and Company.

Hook, Sidney (1959 [1977]) "John Dewey—Philosopher of Growth," reprinted in S. Morgenbesser (ed.), *Dewey and his Critics*, Lancaster: Lancaster Press.

Kellogg, Frederic R. (2007) *Oliver Wendell Holmes, Jr., Legal Theory, and Judicial Restraint*, Cambridge: Cambridge University Press.

Misak, Cheryl (2000) *Truth, Politics, Morality: Pragmatism and Deliberation*, New York: Routledge.

Nagel, Ernest (1939) "Some Leading Principles of Professor Dewey's Logical Theory," *Journal of Philosophy* 36(21): 576–81.

Peirce, Charles Sanders (1931–58) *Collected Papers of Charles Sanders Peirce*, i–vi, ed. C. Hartshorne and P. Weiss; vii and viii, ed. A. Burks, Cambridge, MA: Belknap Press. (Cited as *CP* plus volume and paragraph number.)

—— (1982–2010) *The Writings of Charles S. Peirce: A Chronological Edition*, vols. 1–6 and vol. 8, ed. E. Moore, Bloomington: Indiana University Press. (Cited as *W* volume number: page number; year of original publication.)

Rorty, Richard (1979) *Philosophy and the Mirror of Nature*, Princeton, NJ: Princeton University Press.

Ryan, Alan (1995) *John Dewey and the High Tide of American Liberalism*, New York: W.W. Norton.

Talisse, Robert (2002) "Two Concepts of Inquiry," *Philosophical Writings* 20: 69–81.

—— (2007) *A Pragmatist Philosophy of Democracy*, London: Routledge.

Welchman, Jennifer (1995) *Dewey's Ethical Theory*, Ithaca, NY: Cornell University Press.

White, Morton (1957) *Social Thought in America: The Revolt Against Formalism*, rev. edn., Boston, MA: Beacon Press.

58
POSTMODERNISM AND POLITICS

Todd May

1. What is Postmodernism?

In order to discuss the relation of postmodernism to politics, the main hurdle that has to be cleared stands at the outset. We have to understand what postmodernism is. But if we are to do that, we'll need to drop the idea that there is such a thing as postmodernism which is there to be understood. Postmodernism is not like a quark or a public policy or a number: something that might in some sense be said to exist—or at least whose existence can be argued about—in order to comprehend it. Postmodernism is the name given to a number of movements that arose in a number of different areas. In architecture, for instance, Charles Jencks introduces the term in *The Language of Post-Modern Architecture* (1977) to refer to the eclecticism that supplanted the minimalist approach of high modernism. In painting, it sometimes refers to the idea that innovation has come to an end and that the project of painting is to recycle previous painting strategies with ironic self-reference. This is particularly exemplified in the work of David Salle. In literature, it is unclear whether the term has any use. Sometimes it refers to the breakdown of modernist narrative, exemplified for example in David Foster Wallace's *Infinite Jest*. However, if we are looking for narrative play and breakdown, James Joyce's *Ulysses* will do just as well, alongside works by Virginia Woolf, William Faulkner, and others. And if we call their work postmodernist, then it's not at all clear what modernism is.

Turning to philosophy and theory, matters are even worse. We can certainly classify Jean-François Lyotard's middle period as postmodernist, mostly because he does. Aside from that, however, things get pretty murky. I have colleagues in philosophy who have told me that it refers to the idea of epistemological relativism: that there is no truth and, as a consequence, anything goes epistemically. (In fact, I don't know of any theorist who might be placed in the postmodern tradition who claims that.) Moreover, it might be said that the term "postmodernism" functions chronologically as well as, or more than, thematically. The idea would be that we live in a period that has somehow broken from the previous one, and that the term *postmodernism* can be used to indicate this break. If we do this, then there are at least two more theoretical tendencies, in addition to Lyotard's, that can comfortably be brought under the postmodernist label. The first would be the work of Jean Baudrillard, whose discussion of hyperreality is predicated on the idea that we live in a period that is qualitatively different from previous history, one that must be thought in new terms. The second would be theorists of the rise of transna-

tional capitalism and, in particular, its consequent rampant consumerism. The second tendency focuses on the idea that the globalization of the past several decades has been marked primarily by the spread of capitalist economics and culture, and that this spread is distinctive enough to warrant the recognition of an epochal divide.

Looking at things this way has an important advantage. It gives at least some thematic coherence to the concept of postmodernism by taking postmodernism to refer to the idea of a chronological rupture with the past about which there are debates regarding the thematic character of that rupture. So we will characterize postmodernism as though it were about such a historical cleft, and will isolate three tendencies in postmodernism: Lyotard, Baudrillard, and transnational capitalism/consumerism, primarily represented in the work of Zygmunt Bauman. We will proceed, then, by introducing the thematic elements of their work that bear upon the character of postmodernism, and then ask what political implications, particularly for resistance, arise from their disparate, although not entirely unrelated, views of the character of the postmodern period.

2. Lyotard

Among discussions of postmodernism, Lyotard's *The Postmodern Condition* (1984 [1979]) is by far the most famous. Published in 1979, that single work, more than any other, has introduced the term *postmodernism* to a wider public. It is less often recognized that in the English translation to the work there are two distinct views of postmodernism in play. One, taking up the body of the text, is historical in character. The other, which appears in the Afterword, is more ahistorical. We will focus on the former, since it receives the bulk of Lyotard's attention and has been by far the more influential of the two.

"Simplifying to the extreme," Lyotard writes (1984: xxiv), "I define *postmodernism* as incredulity toward metanarratives." We have entered, Lyotard believes, a period in which we no longer trust the grand narratives or metanarratives that have grounded our existence and especially our knowledge, narratives like the progress of knowledge or increasing freedom and democracy. This, in itself, is neither good nor bad. We might say that it is fraught. Capitalism has taken advantage of this situation in order to thrust upon us an ethic of efficiency or performativity that delimits our possibilities; however, we, in turn, might utilize the loss of overarching narratives to proliferate the kinds of stories we tell ourselves and the kinds of lives we might lead.

Modernism was grounded in two different versions of legitimating metanarratives.

> The subject of the first of these versions is humanity as the hero of liberty. If the social subject is not already the subject of scientific knowledge, it is because that has been forbidden by priests and tyrants. The right to science must be reconquered.
>
> (Lyotard 1984: 31)

This story is a familiar one. It is the Enlightenment story of science emerging victorious against medieval religious superstition. One of the functions of this story is to preserve the right of the state to promote knowledge by supporting scientific endeavor, a promotion that has, among its effects, the marginalization of other kinds of knowledge. The second version is a Hegelian one. "The subject of knowledge is not the

people, but the speculative spirit" (Lyotard 1984: 33). In this version of legitimation, knowledge comes to its own self-realization through its unfolding. People are involved in, but not fundamentally the subject of, this self-realization. History brings knowledge to its own self-realization through the activities of people, but it is the journey of knowledge, not the emancipation of people from ignorance, that is the centerpiece of the narrative.

These metanarratives have gradually eroded in the period after the Second World War. By the time he writes *The Postmodern Condition*, Lyotard is convinced that neither of them holds sway. There are several reasons for this, including for instance the rise of the welfare state, which, by keeping the possibility of communism at bay, has turned people over to an individualism that militates against the collective nature of the progress of knowledge. We now find ourselves in a situation in which the grounding metanarratives of modernism no longer have a grip on us, a situation that opens us to new possibilities and new dangers.

The primary danger of the postmodern condition lies in the domination of capitalism and its criterion of performativity. Knowledge that receives support is knowledge that increases human performance, particularly in regard to the development of capitalism itself.

> Capitalism solves the problem of research funding in its own way: directly by financing research departments in private companies, in which demands for performativity and recommercialization orient research first and foremost toward technological "applications": and indirectly by creating private, state, or mixed-sector research foundations that grant program subsidies to university departments, research laboratories, and independent research groups.
>
> (Lyotard 1984: 45)

Although the indirect funding does not require immediate results, it uses universities in order to fund research that sooner or later will contribute to the productivity of the more applied research that happens in private companies. In the US, we have seen this model become even more prominent in recent years as the decrease in federal funding of universities requires them to rely more and more on private alternatives. If anything, the basic research that used to take place at universities is being eclipsed in favor of research that is more immediately linked to performative requirements.

The hegemony of capitalism and its ethic of performativity do not concern solely the university. They constitute a more general aspect of the postmodern condition. As Lyotard notes, in slightly different language, in *The Differend*, "Working conditions in a capitalist system all result from the hegemony of the economic genre, in which the issue is to gain time" (1988: 176). However, there is another side to postmodernism, one that militates against the reduction to performativity. If the loss of metanarratives opens the space to a capitalist insistence on performativity without an overarching justifying narrative, it also opens the space to other narratives and thus other ways of seeing things and living that have been marginalized during the period of modernism. In *The Postmodern Condition*, with its focus on knowledge, Lyotard points to a movement in science he calls *paralogy*, which is founded on the idea that science can no longer be considered within a stable logical system of inferences. He sees paralogy arising in the emergence of quantum physics. In contrast to the stability of science required for performativity:

> Quantum theory and microphysics require a far more radical revision of the idea of a continuous and predictable path. The quest for precision is not limited by its cost, but by the very nature of matter. It is not true that uncertainty (lack of control) decreases as accuracy goes up: it goes up as well.
>
> (Lyotard 1984: 56)

Paralogical science is, in an important way, closer to the postmodern condition than capitalist performativity. Even though neither requires a grand narrative or metanarrative, the latter strives to minimize novelty, except those novelties that contribute to performance. It seeks to oppress anything that does not contribute to performance. Paralogy, by contrast, conforms to the open space created by the loss of the modernist metanarratives by allowing, and even encouraging, results that fall outside the borders of any system. It challenges the prescriptive elements of any scientific (or generally epistemic) set of rules.

> The function of the differential or imaginative or paralogical activity of the current pragmatics of science is to point out these metaprescriptives (science's "presuppositions") and to petition the players to accept different ones. The only legitimation that can make this kind of request admissible is that it will generate new ideas, new statements.
>
> (Lyotard 1984: 65)

If the postmodern condition is centrally a condition defined by the loss of the two overarching narratives that define modernism, then, at least epistemically, the generation of new ideas and new statements reflects the spirit and possibilities of this new condition better than the capitalist insistence on delimiting novelty to performativity.

3. Baudrillard

For Jean Baudrillard, the postmodern situation is defined not so much by an openness to new conditions of reality as by the absence of reality altogether. If, for Lyotard, the loss of the modern metanarratives places us in a condition in which we can see and thus re-make the world anew, for Baudrillard we can do little more than see something that is no longer really even the world. For Baudrillard, the central fact of the postmodern situation is not the loss of any particular narrative. It is the loss, rather, of the referent of our words and our thoughts, the loss of representation altogether. This loss arises from the rise of media, in the strict sense of that term. We might say that our lives have become so deeply mediated that we have lost contact with the referents of what it is that is being mediated. We live no longer in the real, but in what Baudrillard calls the *hyperreal*, which emerges (Baudrillard 1983a: 3) as "the product of an irradiating synthesis of combinatory models in a hyperspace without atmosphere."

The hyperreal pretends to a reality that does not really exist except through images. It is a mediated relation to a reality that is no longer there. Imagine it this way. When one watches the evening news on television, one expects that the camera images one is seeing reflect the reality on the ground. The camera is assumed to be simply, and literally, a lens onto what is happening. But suppose that weren't true. Suppose that there were nothing more than images without reality, not because of some conspiracy to mislead us, but rather because it is the images, not any reality reflected by them, that

is the true "medium" of our existence. This is the hyperreal. It is defined by Baudrillard as a matter of simulation rather than dissimulation. As he puts the point (1983a: 5), "To dissimulate is to feign not to have what one has. To simulate is to feign to have what one hasn't." In the hyperreal, reality is simulated, on the basis of images.

One might object here that reality has not gone away. When a camera takes a picture of a bit of reality, the reality is indeed there. Or, in opposition to the title of one of Baudrillard's books (1995), the Gulf War did indeed take place. When Baudrillard argues that the simulation of the hyperreal has substituted itself for the real, however, he is not arguing that there is nothing on the other side of the camera or another medium. He is arguing that whatever is there has become irrelevant. More specifically, it is irrelevant to those who are immersed in the medium itself, the viewers. For those of us whose lives largely consist in exposure to and interaction with television, computers, video games, virtual reality, iPhones and iTunes, what is portrayed or represented through these media matters not at all. It is the representation that matters. If the representation were wholly fictitious, it would not change our experience a jot. There might well be a reality out there that is being depicted in some way or another by the media that is the ether of our experience, but this reality does not bear upon us. We live in the hyperreal.

This is the point of Baudrillard's striking formulation about Disneyland, which he calls (1983a: 23) "a perfect model of all the entangled orders of simulation." Disneyland, he tells us,

> is there to conceal the fact that it is the "real" country, all of "real" America, which *is* Disneyland … Disneyland is presented as imaginary in order to make us believe that the rest is real, when in fact all of Los Angeles and the America surrounding it are no longer real, but of the order of the hyperreal and simulation.
>
> (1983a: 25)

The contrast Baudrillard is drawing, although controversial in its formulation, does not pit something with ontological reality (Disneyland) against a fiction (the rest of America). His argument is rather more subtle. It is that when we experience Disneyland we tell ourselves that this is a fictional arena, one to be contrasted with the reality that we experience outside its gates. However, we are mistaken. Outside the gates of Disneyland we do not experience reality but instead live in the hyperreal. The comfort Disneyland provides is to hide this fact from us in offering us the illusion of an outside reality by announcing its own fictional character.

In Baudrillard's eyes, then, what characterizes postmodernism is the loss of reality in favor of the hyperreal and simulation. We are no longer in an age of representation, nor even, as Marx would have it, an age of the ideological mystification of the represented. "Whereas representation tries to absorb simulation by interpreting it as false representation, simulation envelops the whole edifice of representation as itself a simulacrum" (Baudrillard 1983a: 11). As with Lyotard, Baudrillard sees our situation as a new one, but his characterization of it relies not upon the narratives within which we situate ourselves but upon the technology through which we experience what used to be the world. If we turn to our third perspective on postmodernism, the rise of transnational capitalism and the dominance of consumerism, we find a model that intersects with, but is not reducible to, either of these.

4. Transnational Capitalism

The rise of transnational capitalism has been much discussed. Masao Miyoshi insists that we must distinguish the older multinational corporation from the emerging transnational corporation, pointing out (1993: 736) that

> a multinational corporation (MNC) is one that is headquartered in a nation, operating in a number of countries … [and is] finally tied to the home nation. A truly transnational corporation, on the other hand, might no longer be tied to its nation of origin but is adrift and mobile, ready to settle anywhere and exploit any state including its own, as long as the affiliation serves its own interest.

The rise of transnational corporations is intertwined with a decline in the power of nation-states, and thus with a decline in the ability of national citizens to effect political change. Transnational corporations answer to no national entity. If a national (or local) political entity makes things inconvenient for them, they simply relocate. As a result, political entities seek to serve the interests of powerful, mobile, transnational corporations, which are thus immune to political pressure. This, of course, is consonant with the ideology of neoliberalism.

This rise of transnational corporations is often entwined with an emerging consumerism, although it is not necessarily so entwined. (David Harvey (2007) and Naomi Klein (2008) offer trenchant analyses of the rise of transnational capitalism that do not refer to a corresponding consumerism.) The link between the two is not hard to see. On the one hand, citizens often cannot affect corporate operations. On the other, the dominance of corporations leads to the dominance of an economic discursive "genre," to use Lyotard's term, one that has spawned the consumption of goods as a dominant activity, replacing any attempt to create a public space and citizenry or to make corporations responsive to human needs or values. The postmodern character of this consumerism has been articulated perhaps most doggedly by Zygmunt Bauman. (There are, of course, many other works that treat current consumerism, often from a less postmodern angle. See, for instance, Barber (2007) or, in a less popular vein, Lane (2000).)

In a series of books, Bauman has described what he calls a *liquid* modernity.

> What all these features of liquids amount to, in simple language is that liquids, unlike solids, cannot easily hold their shape. Fluids, so to speak, neither fix space nor time … they pass around some obstacles, dissolve some others and bore or soak their way through others still … These are fitting reasons to consider "fluidity" or "liquidity" as fitting metaphors when we wish to grasp the nature of the present, in many ways *novel*, phase in the history of modernity.
>
> (Bauman 2000: 2)

This liquid modernity is centered on a transient consumerism rather than the stable, productive society of the earlier phase of modernity (Bauman 2000: 76): "postmodern society engages its members primarily in their capacity as consumers rather than producers."

In *Consuming Life*, Bauman defines consumerism as

> a type of social arrangement that results from recycling mundane, permanent and so to speak "regime-neutral" human values, wants, desires and longings into the *principal propelling and operating forces of society*, a force that coordinates systematic reproduction, social integration, social stratification and the formation of human individuals, as well as playing a major role in the processes of individual and group self-identification in the selection and pursuit of individual life policies.
>
> (Bauman 2007: 28)

The recycling is what gives consumerism its liquid character. Under the regime of consumerism, people are encouraged to move from desire to desire, focusing not on producing or even on owning but rather on the activity of consuming itself. For Bauman, the previous period, the one described by Foucault for instance in *Discipline and Punish*, was a modernity rooted in production and stability. In order for a society of production to sustain itself, its forms needed to be stable and predictable. Discipline, as described by Foucault, was a method for rendering people predictable.

By contrast, a society of consumers does not concern itself with the creation of stolid, productive individuals but, rather, with restless consumers whose desires, once satisfied, become irrelevant to them.

> Consumer society thrives as long as it manages to render the *non-satisfaction* of its members (and so, in its own terms, their unhappiness) *perpetual*. The explicit method of achieving such an effect is to denigrate and devalue consumer products shortly after they have been hyped into the universe of consumers' desires.
>
> (Bauman 2007: 47)

In Bauman's view, consumerism defines not only our activities but also our sense of who we are and our relationships. In *Liquid Love*, he describes the way the liquidity of consumerist society leads us to invest in relationships that are characterized not by engagement and vulnerability but, instead, by momentary satisfaction and expendability.

> A relationship, the expert will tell you, is an investment like all the others: you put in time, money, efforts that you could have turned to other aims but did not, hoping that you were doing the right thing and that what you've lost or refrained from otherwise enjoying would be in due course repaid—with profit. You buy stocks and hold them as long as they promise to grow in value, and promptly sell them when the profits begin to fall or when other stocks promise a higher income (the trick is not to overlook the moment when that happens).
>
> (Bauman 2003: 13)

As Bauman points out, in consumerist society, not only are products the objects of consumption, consumers themselves, i.e. we, become commodities. Just as the products we buy seek to stand out one from another in order to be noticed for purchase, so we seek to stand out to ourselves and to one another. We become advertisements for ourselves and for the products we buy in much the same way that those products become advertisements for themselves. "The task of the consumers, therefore, and the principal motive prompting them to engage in incessant consumer activity, is the task of lifting

themselves out of that grey and flat invisibility and insubstantiality, making themselves stand out from the mass of indistinguishable objects" (Bauman 2007: 12). Consumerism makes commodities of us all, turning us into the very liquidity that is the ether of our lives.

5. Three Theories and their Implications for Politics

We have, then, three theories of the postmodern. They are distinct, and yet not entirely unrelated to one another. All of them, of course, claim that we are in a qualitatively distinct period from that of modernism. Moreover, this qualitative distinction is linked in some way to capitalism or capitalist development. For Lyotard, capitalism militates against metanarratives; for Baudrillard, it creates the technological emergence of contemporary media; and for Bauman, it underlies the consumerism of the present age. Moreover, there are a series of specific intersections among these theories. The spectatorial character of simulation intersects with consumerism; in fact, in some of Baudrillard's earlier writings (Baudrillard 1975, 1981), he argues that Marx's productivist orientation must be replaced with a more consumerist view. The loss of metanarratives also helps undergird consumerism. If we no longer see ourselves as part of a grand narrative of human progress of some sort or another, this opens the door to engagement in, and identification with, the smaller pleasures of consumption. Or again, the absence of a metanarrative can both foster and be fostered by simulation. Without such a narrative, we are more likely to be subject to the effects of simulation, and, conversely, those effects undercut the formation of such a narrative.

However, the point to be insisted upon for our purposes is that all of these postmodern views are politically inflected. If we are to ask about the relation of these views to politics, we can see that they already are political. In Lyotard's case, the relation is manifest. The loss of metanarratives opens a struggle between the performativity of capitalism and the creation of other narratives, among them that of scientific paralogy. As we will see momentarily, paralogy is only one among the different narratives that struggle against the domination of performativity. In Baudrillard's case, the politics is not as clearly on the surface, but neither is it far from it. Baudrillard's work stems from a critique of Marxist productivism, of the view that who we are is defined by our place in the order of capitalist production, i.e. workers or capitalists. He thus sees us in a new relation to power and politics. This comes out most clearly in his book *Forget Foucault*, where he writes:

> What Foucault does not see is that power is never there and that its institution, like the institution of spatial perspective versus "real" space in the Renaissance, is only a simulation of perspective—it is no more reality than economic accumulation—and what a tremendous trap that is.
>
> (Baudrillard 1987: 41)

If each of these theories of postmodernism is politically inflected, one might wonder how it is that we are to act politically in our new postmodern situation. Are we simply subject to the changes of the postmodern situation, or are there specific strategies that each theorist commends in the face of the salient characteristics upon which each has focused? As it turns out, all three theorists have what might be called theories, or at least suggestive views, of resistance. For Lyotard, these involve linguistic struggle and crea-

tion, for Baudrillard they are a matter of what he calls *seduction*, and for Bauman they involve the creation of interpersonal solidarity and, unsurprisingly, solidity.

The linguistic struggle Lyotard discusses emerges not only in his discussion of paralogy in *The Postmodern Condition*, but in a more rigorous way in his later book *The Differend*. The subtitle of that book is "phrases in dispute." For Lyotard, there are different modes or what he calls *genres* of discourse. When one utters a phrase, that phrase does not dictate what phrase is to follow. "It is necessary to link, but the mode of linkage is never necessary" (Lyotard 1988: 29). A genre does not consist in a single phrase but, rather, is the framework within which phrases occur. Therefore, political struggle can occur around the genres. This idea is nascent in the discussion we've had about *The Postmodern Condition*. Capitalism dictates a genre of performativity, or, as Lyotard puts it in *The Differend*, "Capitalism gives political hegemony to the economic genre" (1988: 141). Thus political struggle is struggle over genres of discourse: which ones will be allowed, which ones are oppressed, what opportunities one has to engage in different ones. The hegemony of a particular genre suppresses the possibility of other genres: genres of narration, for instance, or obligation. (We should note that the domination of a genre does not make it a metanarrative. A metanarrative seeks to justify other narratives or genres. A dominant genre does not justify; it simply dictates the kinds of phrases that can be linked to other phrases.)

Politics, then, consists in keeping open the space of different genres, resisting the reduction to a single or a few genres. Politics "is not a genre, it is the multiplicity of genres, the diversity of ends, and par excellence the question of linkage" (Lyotard 1988: 138). In the absence of a metanarrative, the field of linkage, i.e. the way phrases can be linked to other phrases, is open. It is not dictated by an overarching story of justification. However, this postmodern condition does not preclude the possibility of the domination of some genre or genres. Capitalism has allowed the economic genre of performativity to dominate the postmodern epoch. The task of political resistance, then, is to open up the field for other linkages, other genres. It is not to substitute another genre for the economic one, for instance to substitute paralogy for capitalism. Rather, it is to ensure that there can be many genres in play, and that people are not prohibited from moving among those genres in their linkages. To prohibit them in this way would be to limit not only what they can say, but how they can think and live. The postmodern condition, if it is to be liberating, must free us from the bonds of a dominant genre that would replace the metanarrative of the modern period.

For Baudrillard, political struggle is not as much a matter of language. In fact, it is not as much a matter of struggle. To struggle, in the classic political sense, is nostalgic. It is to believe that there is a reality over which one can contest. In our period of hyperreality, politics cannot consist in anything like a struggle for or over something that is in dispute. It must have an entirely different structure. When Baudrillard discusses these matters, the structure he allots to it is that of seduction. Seduction, in Baudrillard's eyes, is associated with a certain form of the feminine. Unsurprisingly, it operates at the level of signs rather than reality. Seduction offers symbols in a play that does not seize anything, but rather entices without having to seize: "seduction as an ironic, alternative form, one that breaks the referentiality of sex and provides a space, not of desire, but of play and defiance" (Baudrillard 1990: 21).

Seduction does not have an end-point. Rather, it is a process, one that is moldable and even reversible. What makes it seduction is simply that it is a play without stakes. Stakes are the illusion of reality. Seduction refuses that illusion, and plays solely on the register of symbols without reference. In politics, it is the game of pretending to obey

in order to seduce those who think they have power, rather than obeying because one believes. Baudrillard marks this contrast in distinguishing Law from Rule.

> The Law is part of the world of representation … It is a text, and falls under the influence of meaning and referentiality. By contrast, the Rule has no subject, and the form of utterance is of little consequence: one does not decipher the rules, nor derive pleasure from their comprehension—only their observance matters, and the resulting giddiness.
>
> (Baudrillard 1990: 132)

One can see this strategy of seduction operating in Baudrillard's text *In the Shadow of the Silent Majorities*. The silent majority—or the masses, depending on one's political orientation—are brought to bear as a justifier for the political strategy of a particular elite. The difficulty, for this elite, is to make the silent majority or the masses speak, to give it voice. Baudrillard argues, however, that the silence of the majority is not its failure or inability to act; it is precisely its action. The silent majority does not need a voice. Its silence is its way of acting. In its silence, the majority or the masses defuse the messages of politicians who still find themselves caught within the illusion of reality. "The mass is not a place of negativity or explosion, it is a place of absorption and implosion" (Baudrillard 1983b: 22). This is seductive, drawing upon the energy of politicians without struggling with or against them.

Bauman's view of resistance is one that appears in theoretical gestures in his work. Because of the pervasiveness of liquid modernity, he seems sober about the chances for struggle against the consumerism that has come to define us. But he is not entirely despairing.

> What is most conspicuously missing in the economic calculus of the theorists … is the huge area of what A.H. Halsey called the "moral economy"—family sharing of goods and services, neighbourly help, friends' cooperation: all the motives, impulses and acts from which human bonds and lasting commitments are plaited.
>
> (2003: 69)

The key terms here are *bonds* and *lasting*. In a situation that is defined by transient individual choices of consumer products and consumer display, resistance comes in the form of the creation of solid political and interpersonal bonds, bonds that can withstand the corrosion of a liquid modernity.

Because of this, Bauman believes that the approaches of both Foucault's analysis of disciplinary power and the Critical Theorists' critique of modern society have missed crucial and emancipatory aspects of earlier modernity. Foucault, in his work on the prisons, focuses on the constraining power of disciplinary action. However, the very rigidity he associates with discipline is inseparable from its potential for liberation.

> The end of the Panopticon augurs *the end of the era of mutual engagement*: between supervisors and the supervised, capital and labour, leaders and their followers, armies at war. The prime technique of power is now escape, slippage, elision and avoidance, the effective rejection of any territorial confinement with its cumbersome corollaries of order-building.
>
> (Bauman 2000: 11)

Whereas Foucault seems to see an advance of techniques of control in the rise of disciplinary power, Bauman, without denying this, also sees the possibility of resistance: the enemy is recognized and known. Therefore, anything that builds stability into the system also offers the opportunity to undermine the powers that create and sustain that stability. What makes resistance against postmodern consumerism so difficult is precisely the instability of its liquid character.

In order to create or re-create this stability, it is necessary, as Bauman notes, to build bonds that are lasting. These bonds need not only be interpersonal ones. They can also be political. This is why Bauman inverts the view of the Critical Theorists that the private must be preserved from the encroachment of the public. Whereas the Critical Theorists were concerned about the ways in which the power of the economic and political elites has infiltrated individual life, and thus sought to protect individual creation and freedom from this infiltration, the political task now must move in the opposite direction: toward the emergence of a public solidarity.

> *Any true liberation calls today for more, not less, of the "public sphere" and "public power"*. It is now the public sphere which badly needs defence against the invading private—though, paradoxically, in order to enhance, not to cut down, individual liberty.
>
> (Bauman 2000: 51)

For Bauman, the project of struggle against liquid modernity requires nourishing both the interpersonal and political bonds between people, the creation of solidarity and citizenship. Without these, it is impossible to resist the liquefaction of consumerism and its consequent isolation of people into individuals consuming goods in a pointillist time, disengaged from meaningful relations with others.

Related Topics

Pluralism, Natural Law and Rights Theory, Needs and Distributive Justice, Nationalism, Discourse Theory

References

Barber, Benjamin (2007) *Consumed: How Markets Corrupt Children, Infantilize Adults, and Swallow Citizens Whole*, New York: Norton.

Baudrillard, Jean (1975) [1973] *The Mirror of Production*, trans. Mark Poster, St. Louis: Telos Press.

—— (1981) [1972] *For a Critique of the Political Economy of the Sign*, trans. Charles Levin, St. Louis: Telos Press.

—— (1983a) *Simulations*, trans. Paul Foss, Paul Patton, and Philip Beitchman, New York: Semiotext(e).

—— (1983b) *In the Shadow of the Silent Majorities and Other Essays*, trans. Paul Foss, Paul Patton, and John Johnston, New York: Semiotext(e).

—— (1987) [1977] *Forget Foucault*, trans. Nicole Dufresne, New York: Semiotext(e).

—— (1990) [1979] *Seduction*, trans. Brian Singer, New York: St. Martin's Press.

—— (1995) [1991] *The Gulf War Did Not Take Place*, trans. Paul Patton, Bloomington: Indiana University Press.

Bauman, Zygmunt (2000) *Liquid Modernity*, Cambridge: Polity Press.

—— (2003) *Liquid Love: On the Frailty of Human Bonds*, Cambridge: Polity Press.

—— (2007) *Consuming Life*, Cambridge: Polity Press.

Harvey, David (2007) *A Brief History of Neoliberalism*, Oxford: Oxford University Press.

Jencks, Charles (1977) *The Language of Post-Modern Architecture*, New York: Rizzoli.

Klein, Naomi (2008) *The Shock Doctrine: The Rise of Disaster Capitalism*, New York: Picador.
Lane, Robert (2000) *The Loss of Happiness in Market Democracies*, New Haven, CT: Yale University Press.
Lyotard, Jean-François (1984) [1979] *The Postmodern Condition: A Report on Knowledge*, trans. Geoff Bennington and Brian Massumi, Minneapolis: University of Minnesota Press.
—— (1988) [1983] *The Differend: Phrases in Dispute*, trans. Georges Van Den Abbeele, Minneapolis: University of Minnesota Press.
Miyoshi, Masao (1993) "A Borderless World? From Colonialism to Transnationalism and the Decline of the Nation-State," *Critical Inquiry* 19.

59
SOCIAL CHOICE THEORY

John A. Weymark

1. Arrow's Theorem

In Kenneth Arrow's seminal *Social Choice and Individual Values* (Arrow 1951), collective decision-making is viewed as a problem in preference aggregation: the objective is to determine a collective (social) ranking of a set of alternatives based on individual preferences. This procedure for aggregating preferences, a *social welfare function*, is designed before the preferences are known so that it will be applicable whatever these preferences turn out to be. In this respect, a social welfare function operates like the rules for conducting an election: the procedure for determining a winner based on the ballots cast is specified before the voting takes place.

Prior to Arrow's book, the formal analysis of collective decision-making studied particular voting procedures, such as majority rule. Much of this analysis consisted of identifying problems with existing voting rules and of designing new rules to overcome them. For example, pairwise majority is subject to *Condorcet's Paradox*. There are three candidates for an election (x, y, and z) and three voters. Voter 1 prefers x to y to z, voter 2 prefers y to z to x, and voter 3 prefers z to x to y. Using majority voting a pair at a time results in a cycle with x beating y, y beating z, and z beating x. There is no clear-cut winner.

Arrow's most fundamental contribution was to change the focus from the analysis of specific collective decision-making procedures to the identification of the properties that any good preference aggregation procedure should exhibit. Once these properties have been identified, the task is then to identify what aggregation rules satisfy them. Unfortunately, the properties that Arrow thought reasonable are incompatible, as shown in his famous impossibility theorem.

Arrow's analysis is formal, abstract, and axiomatic. Individuals are characterized by their preferences, the set of alternatives is some set whose members are unspecified, and the properties of aggregation rules are formalized as axioms. Alternatives can be anything: quantities of a public good (i.e. a good that is non-rival in consumption), consumption and production plans for an economy, candidates for an election, etc.

A *preference* is a binary relation R on the set of alternatives X, with xRy interpreted as meaning that alternative x is weakly preferred to alternative y. The corresponding strict preference relation is P ("strictly preferred to"). R is *reflexive* if any alternative is weakly preferred to itself; *complete* if for any two distinct alternatives, at least one of them is weakly preferred to the other; and *transitive* if whenever one alternative is weakly preferred to a second and the second is weakly preferred to a third, then the first is weakly preferred to the third. A preference is an *ordering* if it satisfies these three "rationality" properties. The set of all orderings on X is $\mathcal{R}$.

There are a fixed finite number of individuals, n, with $n \geq 2$. Each individual i has a preference ordering R_i on X. A *preference profile* $\mathbf{R} = (R_1, \ldots, R_n)$ is a list of individual preferences, one for each person. The set of all possible preference profiles is $\mathcal{R}^n$, the n-fold Cartesian product of $\mathcal{R}$. Preference profiles are a priori restricted to be in some *preference domain* $\mathcal{D}$ contained in $\mathcal{R}^n$. A *social welfare function* f assigns a social preference on X to each preference profile in $\mathcal{D}$.

The Arrovian axioms are those used in the second edition of Arrow's book. The first axiom was incorporated into the definition of a social welfare function, rather than stated separately by Arrow.

> *Social Ordering*. For each preference profile $\mathbf{R}$ in the preference domain, the corresponding social preference R is an ordering.

Social Ordering is Arrow's *collective rationality* axiom. One rationale for it is that social preferences should exhibit the same degree of rationality as individual preferences. By requiring social preferences to be orderings, cycles as in the Condorcet Paradox are precluded.

If there are no a priori restrictions on the individual preferences, then f should produce a social ordering for every conceivable preference profile.

> *Unrestricted Domain*. The preference domain $\mathcal{D}$ is $\mathcal{R}^n$.

If everybody agrees that one alternative is strictly preferred to a second, it seems reasonable for the first alternative to be socially preferred to the second, a principle due to Vilfredo Pareto.

> *Weak Pareto*. For any pair of alternatives x and y and any preference profile $\mathbf{R}$ in the preference domain, if xP_iy for every individual i, then xPy.

When deciding how to socially rank a pair of alternatives, Arrow argued that all that should matter is what they are and how individuals rank them. Everything else, including the rest of the individual preference orderings, is irrelevant. Descriptions of the alternatives might matter if the pair being considered differ only in some feature that one individual should have a right to choose (e.g. the color of his bedroom walls) or one of them is given special status (e.g. by a constitution).

> *Independence of Irrelevant Alternatives*. If two preference profiles $\mathbf{R}$ and $\mathbf{R}'$ in the preference domain coincide on a pair of alternatives, then the corresponding social preferences R and R' also coincide on this pair.

It is often mistakenly thought that this axiom precludes the presence or absence of some third alternative from influencing the social ranking of any pair of alternatives as, for example, might happen if an additional candidate contests an election. In Arrow's framework, the set of alternatives is fixed, so this situation does not arise.

Arrow's final axiom ensures that nobody has undue influence in determining social preferences. Someone is a *dictator* if the social preference agrees with this person's preference on any pair of alternatives for which he expresses a strict preference.

Nondictatorship. There is no dictator.

Arrow's Theorem shows that these axioms are incompatible if there are at least three alternatives. With only two alternatives, majority rule satisfies all five axioms. In this case, the requirement that social preferences be transitive is vacuous. With an infinite society (e.g. all future generations are being considered), the axioms are compatible, but the only social welfare functions that satisfy them exhibit dictatorial features.

Arrow's Theorem can be reformulated in choice-theoretic terms. A social choice is made from an *agenda* of feasible alternatives, a subset of X. Prior to designing the choice procedure, the actual agenda is only known to be in the *agenda domain* $\mathcal{A}$. For each preference profile $\mathbf{R}$ in the preference domain and each agenda A in the agenda domain, a *social choice correspondence* C chooses a nonempty subset of the agenda, the *choice set* $C(\mathbf{R}, A)$, all members of which are regarded as being equally good. The agenda domain is assumed to include all the two- and three-alternative agendas. This domain assumption is incompatible with choice-theoretic analogs of the Arrovian social welfare function axioms.

2. Weakening the Arrow Axioms

The Arrow axioms are logically independent and any four of them are consistent. There is an extensive literature that considers the implications of dropping or weakening some of the Arrow axioms. Except as described below, this has not been a promising way to escape from the nihilism of Arrow's Theorem.

The desirability of each Arrovian axiom has been questioned, at least in some applications of the theory. In many contexts, it is natural to suppose that some preferences profiles will not arise, in which case Unrestricted Domain is inappropriate. Amartya Sen has suggested that Weak Pareto is not innocuous, arguing that the reasons why individuals hold their preferences, not just what these preferences are, matters. For example, in a two-person society, if the only difference between x and y is that person 1 tortures person 2 in x and both prefer x to y because the first person is a sadist and the second is a masochist, then one might not want to respect their common preference for x over y. Donald Saari has argued that Independence of Irrelevant Alternatives effectively prevents a social welfare function from taking account of the transitivity of individual preferences because by ranking alternatives pairwise, the linkages across pairs provided by transitivity is lost. Nondictatorship is inappropriate if the objective is to rank alternatives according to their technical merit and only one individual has the relevant expertise.

Justifying Social Ordering in terms of endowing society with the same degree of rationality as individuals can be regarded as illegitimately anthropomorphizing society. Nevertheless, it is desirable for social preferences to be *acyclic* (i.e. to have no strict preference cycle over a finite number of alternatives). If Social Ordering is weakened by only requiring social preferences to be reflexive, complete, and acyclic, then the axioms are consistent if there are at least as many alternatives as individuals, as illustrated by the voting rule used by the UN Security Council prior to 1965.

The choice-theoretic Arrovian axioms can also be questioned. A natural weakening of the choice-theoretic collective rationality assumption is *Path Independence*, which requires that for any profile $\mathbf{R}$, the alternatives chosen from any agenda $A = A^1 \cup A^2$ would also be chosen if we first separately choose from A^1 and A^2 and then make the

final choice from among these choices. With Path Independence, nobody can benefit by manipulating the order in which alternatives are considered.

Sequential voting applied to the alternatives and preference profile used in Condorcet's Paradox is subject to agenda manipulation. Majority rule is first used to determine a winner from one pair of alternatives with the first-round winner then being compared to the third alternative. Depending on the order in which alternatives are compared, any outcome may result. For example, x is ultimately chosen if the first vote is between y and z. Thus, whoever determines the voting order can ensure that his favored outcome is chosen. This scenario arises in legislative and committee settings in which the standard voting rules result in three alternatives being considered: the *status quo* x, a motion y, and an amendment z. Had z been the motion and y the amendment, a different outcome may result.

3. Domain Restrictions

For the kinds of social choice problems encountered in economic and political applications, the set of alternatives is not some abstract set but, rather, something concrete. Furthermore, individual preferences and agendas exhibit considerable structure. As a consequence, there are natural domain restrictions and, hence, Arrow's Theorem is not directly applicable.

X is *one-dimensional* if the alternatives can be identified with points on the real line $\Re$. For example, X is the set of nonnegative numbers when each alternative specifies the expenditure on some public good and it is a set of m numbers if there are m candidates for an election, with higher numbers assigned to more right-wing candidates. In such applications, it is natural to suppose that individual preferences are single-peaked. A preference R is *single-peaked* if there is a unique best alternative in X, the *peak*, and alternatives on the same side of the peak are worse the farther they are from it.

With single-peaked preferences, Duncan Black has shown that pairwise majority rule results in a social preference that is an ordering if the number of individuals is odd. With an even number of individuals, the social strict preferences are transitive, but the social indifference relation need not be.

Pairwise majority rule, and many other social choice procedures, treat individuals symmetrically and alternatives neutrally.

> *Anonymity*. For any two preference profiles $\mathbf{R}$ and $\mathbf{R}'$ in the preference domain for which $\mathbf{R}'$ is obtained from $\mathbf{R}$ by permuting the individual preferences, the corresponding social preferences R and R' coincide.
>
> *Neutrality*. For any two preference profiles $\mathbf{R}$ and $\mathbf{R}'$ in the preference domain and any four alternatives w, x, y, and z, if wR_ix if and only if yR'_iz for every individual i, then wRx if and only if $yR'z$.

By setting $w = y$ and $x = z$, it follows that Neutrality implies Independence of Irrelevant Alternatives. With Anonymity, social preferences are invariant to permutations of the preferences among the individuals. With Neutrality, the names of alternatives are irrelevant—only the pattern of preferences (person by person) matters. Anonymity is inappropriate if someone should have the right to decide between two particular alternatives. Similarly, Neutrality is inappropriate if any alternative, such as the *status quo*, should be given special treatment.

A domain of preference profiles is *Arrow inconsistent* if no social welfare function satisfies the non-domain Arrovian axioms on it. When the set of alternatives is multidimensional, the preference domains encountered in economic and political applications are typically Arrow inconsistent. For example, if the set of alternatives is the set of all possible allocations of $m \geq 2$ divisible public goods and the preference domain consists of all profiles of Euclidean spatial preferences, then the domain is Arrow inconsistent. With a *Euclidean spatial preference*, there is a best alternative x, with y strictly preferred to z if and only if it is closer to x. A Euclidean spatial preference is a multidimensional generalization of a single-peaked preference that is often used in political models. The preference domain is also Arrow inconsistent when there are two or more public or private goods and preferences exhibit the standard assumptions made in economic models.

With a social choice correspondence, it is also possible to restrict the agendas from which choices are made. For example, in an exchange economy with two or more divisible private goods, an agenda consists of all the ways that these goods can be allocated to the individuals given the aggregate amounts of the goods available. By varying these resource endowments, different agendas are obtained. Martin Bailey has observed that Arrow's non-domain choice-theoretic axioms are consistent in exchange economies if preferences satisfy the standard economic assumptions. For example, these axioms are satisfied if the aggregate endowment is first divided equally and then the individuals trade at competitive equilibrium prices. By restricting both the preference and agenda domains, a mix of Arrovian possibility and impossibility theorems has been obtained.

4. Social Welfare Functionals

Amartya Sen has argued that the information used by a social welfare function is inadequate for comparing alternatives in terms of their contribution to social welfare. Specifically, the definition of a social welfare function prevents the social ranking from taking account of non-preference welfare information. As a consequence, classical social decision rules used in social and political philosophy such as utilitarianism or leximin (utility) are ruled out from the outset. These rules can be formalized using Sen's concept of a social welfare functional.

In Sen's framework, the well-being of individual i is measured by a *utility function* U_i that assigns a number to each alternative in X. A *social welfare functional* F determines a social preference ordering of X for each profile of utility functions $\mathbf{U} = (U_1, \ldots, U_n)$ in some domain $\mathcal{D}$ of *profiles of utility functions*. With the *utilitarian* rule, x is socially strictly preferred to y if and only if the sum of individual utilities with x is larger than the corresponding sum with y. With the *leximin* rule, this social ranking obtains if and only if the smallest utility that is not the same for both alternatives is larger for x than with y. The leximin social welfare ordering underlies John Rawls' difference principle, but with utility substituting for his index of primary goods.

Welfarism is the consequentialist principle that requires that only utility consequences matter when socially ranking alternatives. Given a profile $\mathbf{U}$, the *utility consequences* of alternative x is the vector (ordered list) of n numbers $U(x) = (U_1(x), \ldots, U_n(x))$ indicating what each person's utility is with x. The set of all such vectors is the n-dimensional Euclidean space $\Re^n$. A *social welfare ordering* is an ordering R^* of $\Re^n$. A social welfare functional F is *welfarist* if there exists a social welfare ordering R^* for which for every profile $\mathbf{U}$ in $\mathcal{D}$ and every pair of alternatives x and y in X, the social ranking of x and y is

given by the ranking of the utility vectors $U(x)$ and $U(y)$ by R^*. Hence, in socially ranking two alternatives, the only information used is the utilities obtained with them, and not their physical descriptions or the utility functions used to generate these utilities. Utilitarianism and leximin, as well as many other social decision rules, are welfarist.

Welfarism is characterized by three axioms for F.

> *Unrestricted Domain*. $\mathcal{D}$ is the set of all possible profiles of utility functions.
>
> *Pareto Indifference*. For any pair of alternatives x and y and any profile of utility functions $\mathbf{U}$ in the domain, x is socially indifferent to y if $U(x) = U(y)$.
>
> *Independence of Irrelevant Alternatives*. For any two alternatives x and y and any two profiles of utility functions $\mathbf{U}$ and $\mathbf{U}'$ in the domain, if $U(x) = U'(x)$ and $U(y) = U'(y)$, then the social ranking of x and y coincide for these profiles.

Given Unrestricted Domain, a social welfare functional is welfarist if and only if it satisfies Pareto Indifference and Independence of Irrelevant Alternatives. With Pareto Indifference, the descriptions of alternatives are irrelevant, whereas with Independence of Irrelevant Alternatives, the utility functions used to generate the utility consequences are irrelevant.

The ability to make intrapersonal and interpersonal utility comparisons may be limited, which places constraints on what social welfare functionals can be employed. Informational restrictions are usually modeled by specifying what transformations to the utility functions preserve the meaningful utility information. A utility function U_i is *ordinal* if the only meaningful statements are those that are preserved by applying an increasing transform $g_i(u)$ to it. That is, the utility $U_i(x)$ is replaced by $V_i(x) = g_i(U_i(x))$. With an ordinal utility function, it is meaningful to make intrapersonal comparisons of utility levels. If the same increasing transform must be applied to each person's utility function, then utility levels are also interpersonally comparable. U_i is *cardinal* if the only meaningful statements are those that are preserved by applying an affine transform $g_i(u) = a_i + b_i u$, where a_i and b_i are constants with $b_i > 0$. With such transforms, both utility levels and differences in utility are intrapersonally comparable. If the b_i are the same for everyone, then utility differences are interpersonally comparable and if the a_i are also the same, then so are utility levels.

U_i implicitly defines a preference R_i by setting xR_iy if and only if $U_i(x) \geq U_i(y)$. Thus, if utility is ordinal and interpersonally noncomparable, Sen's framework reduces to that of Arrow because only individual preference information is meaningful.

Social welfare functionals have been used to provide axiomatizations of a number of different social decision rules, including utilitarianism and leximin. Utilitarianism is only meaningful if utility differences are interpersonally comparable. With this informational assumption, Claude d'Aspemont and Louis Gevers have axiomatized utilitarianism by supplementing the welfarism axioms with the social welfare functional analogs of Weak Pareto and Anonymity. Leximin, on the other hand, is only meaningful if utility levels are interpersonally comparable. Peter Hammond has axiomatized leximin by combining the welfarism axioms with Anonymity, Strong Pareto (any change that makes someone better off with hurting anyone else is a social improvement), and an equity axiom that regards any change that reduces the inequality in the utilities of any two individuals as being a social improvement.

In order to address problems in population ethics, Sen's framework has been extended to allow for different population sizes by Charles Blackorby and David Donaldson.

Alternatives are now complete histories of the world, with possibly different people alive in different histories. For each history, only the individuals who ever live in it have utilities. Utilities are lifetime utilities, with a utility of zero representing a life that is just worth living. Welfarism now requires that the social ranking of alternatives be determined by a social welfare ordering of all possible vectors of individual utilities for all possible population sizes.

Familiar variable-population social welfare orderings are *total* and *average utilitarianism*, which compare alternatives according to their utility sums and their average utilities, respectively. Total utilitarianism is subject to Derek Parfit's *repugnant conclusion*: for any alternative in which everybody has the same arbitrarily high utility, there is a socially preferred alternative with a larger population in which everybody has lives barely worth living. Average utilitarianism has the drawback that it recommends that an individual should not be brought into existence if by doing so average utility decreases, even if this person is extremely well off and nobody else is affected.

To overcome these problems, Blackorby and Donaldson have proposed using a *critical-level utilitarian* rule, which ranks alternatives by first subtracting a positive constant—the *critical level*—from each person's utility before summing them. With this rule, adding a person with the critical level of utility to any society is a matter of social indifference. Several extensions of this basic approach have been developed by Blackorby and Donaldson in conjunction with Walter Bossert.

5. Harsanyi on Utilitarianism

In the early 1950s, John Harsanyi used expected utility theory to provide decision-theoretic foundations for utilitarianism. For the set of alternatives $X = \{x_1, \ldots, x_m\}$, a *lottery* is a probability distribution $p = (p_1, \ldots, p_m)$ that yields outcome x_j with probability p_j. Receiving x_j for certain is equivalent to facing the lottery in which $p_j = 1$. A preference R on the set of lotteries L on X is *represented* by the utility function U on L if pRq if and only if $U(p) \geq U(q)$ for all lotteries p and q. R is an *expected utility preference* if it can be represented by a utility function for which the utility of a lottery is the sum of the utilities for each outcome in X weighted by the probabilities of their occurrence. Expected utility preferences are characterized by a set of axioms that Harsanyi regarded as being compelling rationality requirements.

For Harsanyi, welfare judgments are impersonal preferences expressed by a moral agent who ranks lotteries on social alternatives based on a sympathetic and impartial concern for each member of society. He modeled such preferences using what is now known as a *veil of ignorance*. The moral agent engages in a thought experiment in which he imagines having an equal chance of being each person in society, complete with that person's preferences and objective circumstances. In this way, the problem of ranking social alternatives is transformed into a problem of individual decision-making under uncertainty in which the outcome once the uncertainty has been resolved is an *identity-outcome pair* consisting of being some person i with some risk-free social alternative x. Assuming that the preferences of each individual and the moral agent satisfy the expected utility axioms, Harsanyi argued that the moral agent must rank alternatives using the average utilitarian rule. In developing his principles of justice, John Rawls employed a thicker veil in which it is not known what the preferences are outside the veil or who, in fact, exists.

Harsanyi's utilitarian conclusions have been challenged by Amartya Sen using an informal argument that was later formalized by John Weymark (1991). Only increasing

affine transforms preserve the expected utility form of a utility function. If these are the only permissible transforms of the utility functions, which is what Harsanyi mistakenly thought to be the case, then the moral agent's preferences over identity-outcome lotteries would supply the interpersonal comparisons of utility differences required for utilitarianism to be meaningful. However, an expected utility preference need not be represented by a utility function with the expected utility form; any increasing transform of it represents the same preferences. This ordinality precludes making comparisons of utility differences, thereby undermining Harsanyi's utilitarian conclusions.

Harsanyi offered an alternative justification for a weighted form of utilitarianism based on aggregating a single profile of preferences over lotteries. It is also subject to the Sen-Weymark critique.

6. Strategy-Proofness

A social choice function f chooses one alternative from the set of alternatives X for each preference profile in the domain $\mathcal{D}$. The *range* of f is the set of alternatives that are chosen for some profile in $\mathcal{D}$.

If preferences are private information, individuals must be provided with incentives to truthfully report their preferences so that the desired outcome $f(\mathsf{R})$ is chosen when the true preference profile is R. A social choice function is *strategy-proof* (nonmanipulable) if nobody can ever obtain a preferred outcome by reporting a false preference. A social choice function that takes account of the fine structure of individual preferences is typically manipulable. On many domains, strategy-proofness implies the *tops-only property*—what is chosen only depends on the individuals' most-preferred alternatives on the range.

Point voting procedures such as the Borda rule are easily manipulated. The *Borda rule* for m alternatives when nobody is indifferent between distinct alternatives has each person assign m points to his most-preferred alternative, $m - 1$ points to his second-best alternative, and so on, with the chosen alternative being the one with the most points (ties are broken according to some fixed rule that does not depend on preferences). Suppose that the preferences of three individuals for five alternatives are given in descending order by w, x, y, z, v for person 1; x, w, y, v, z for person 2, and z, y, x, w, v for person 3. For this profile, x is chosen. Person 1 can manipulate the outcome so that w is chosen, which he prefers to x, by falsely reporting the preference in which x and v are interchanged in his ordering.

An alternative is *Pareto optimal* if there is no other alternative that is unanimously preferred to it. *Pareto Optimality* is the requirement that only Pareto optimal alternatives are chosen. Allan Gibbard and Mark Satterthwaite have independently shown that if a social choice function satisfies Unrestricted Domain and Pareto Optimality, then it must be dictatorial when there are three or more alternatives.

Positive results are obtained if X is one-dimensional and preferences are single-peaked. For example, with an odd number of individuals, Black's *median voter theorem* shows that on the domain $\mathcal{S}^n$ of all single-peaked preference profiles, the median preference peak defeats every other alternative in a pairwise majority vote and that choosing the median peak is strategy-proof. For any positive integer n, Hervé Moulin has identified all of the social choice functions on the domain $\mathcal{S}^n$ that are strategy-proof, treat individuals symmetrically, and always picks Pareto optimal alternatives. When the alternatives are multidimensional, strategy-proof rules that choose Pareto optimal alternatives are

typically dictatorial unless the range is one-dimensional. Because of the contributions of, among others, Salvador Barberà and his collaborators, much is now known about the implications of strategy-proofness on restricted preference domains.

7. Rights

Welfarist approaches to social choice preclude taking account of individual rights when making collective decisions. As Sen's *liberal paradox* demonstrates, some kinds of rights are inconsistent with the Pareto principle and a weak collective rationality requirement.

Sen illustrated this conflict with an example of two individuals deciding who should read an erotic novel. In alternative x, the prudish person 1 reads it; in y, it is read by lascivious person 2; and in z, nobody reads it. The prude prefers z to x to y because he thinks that nobody should read this smut, but if someone must, better him than the impressionable person 2. Person 2 prefers x to y to z because he thinks that it would be best if the prude expanded his horizons by reading the book, but if this is not possible, he would like to read it. If an individual has the right to determine the social ranking between alternatives that differ only in whether he reads the book, then socially y is preferred to z and z is preferred to x. If unanimous preferences are also respected, then x is preferred to y. Hence, the social preferences are cyclic.

Sen's version of the liberal requirement that everybody has some rights is given by the following axiom for a social welfare function.

> *Liberalism*. For each individual, there exist distinct alternatives for which the social preference agrees with this person's preference whenever it is strict.

Sen's Theorem shows that no social welfare function can satisfy Unrestricted Domain, Weak Pareto, and Liberalism if social preferences are required to be acyclic. Sen's response to this result is to abandon the Pareto principle. In terms of his example, he does not think that unanimity should be respected because the preferences are held for inappropriate reasons; they are meddlesome.

Allan Gibbard has shown that Weak Pareto can be dispensed with in Sen's Theorem if Liberalism is strengthened by letting each individual determine the social ranking on *any* pair of alternatives that differ only in some features that lie within his private sphere (e.g. in what books he reads). Thus, the fundamental conflict is between liberal rights and having an unrestricted preference domain, not with Paretianism. Gibbard's solution to this conflict is to either only give an individual these kinds of rights when his preferences for the alternatives in his private sphere do not depend on what anybody else gets or to allow individuals to waive some of their rights.

When reformulated as a problem of choosing one of the available social alternatives, the kinds of rights that Gibbard and Sen have considered give individuals the right to exclude some alternatives from being chosen. Robert Nozick does not believe that this is the correct way to model rights. He regards an individual as having the right to choose the features of a social alternative that lie within his private sphere, but not to make any decision about the complete social alternative. Unfortunately, as shown by Wulf Gaertner, Prasanta Pattanaik, and Kotaro Suzumura, Nozick's approach does not circumvent the problems that Gibbard and Sen have identified.

8. Judgment Aggregation

Judgment aggregation is concerned with the problem of making consistent collective judgments on a set of logically connected propositions based on the relevant group members' judgments on them. The analysis of judgment aggregation originated with Lewis Kornhauser and Lawrence Sager's *doctrinal paradox*. A three-judge court considers three propositions:

p: the defendant contracted not to do x,
q: the defendant did x,
r: the defendant is in breach of contract.

Legally, the conclusion r is valid if and only if the premises p and q are both true. Suppose that only judges 1 and 2 think that p is true, whereas only judge 1 thinks that q is. Thus, only judge 1 thinks that the defendant is guilty. The defendant is not in breach of contract if liability is determined by a majority vote on the conclusion. This outcome is reversed if majority votes are instead taken on the premises because then p and $\neg q$ are judged to be true and, hence, $\neg r$ is not. Thus, the propositions $\{p, q, r \leftrightarrow (p \wedge q), \neg r\}$ are inconsistent even though the judges each have consistent judgments.

More generally, let $\mathbf{L}$ be a *set of propositions* endowed with the *negation operator* $\neg$ and some concept of consistency. An *agenda* A is a nonempty subset of $\mathbf{L}$ that is closed under negation. A *judgment set* is a subset J of A of propositions that are believed to be true. J is *consistent* if it is a logically consistent set of propositions and *complete* if it contains one member of each proposition-negation pair in A. For a group of $n \geq 2$ individuals, a *judgment aggregation rule* G assigns a collective judgment set $G(\mathbf{J})$ to each *profile* $\mathbf{J} = (J_1, \ldots, J_n)$ of individual judgment sets in some domain $\mathcal{D}$. Consider the following axioms for G.

Unrestricted Domain. The domain $\mathcal{D}$ consists of all possible profiles of consistent and complete individual judgment sets on A.

Collective Rationality. For any profile $\mathbf{J}$ in $\mathcal{D}$, $G(\mathbf{J})$ is a complete and consistent collective judgment set.

Anonymity. For any two profiles $\mathbf{J}$ and $\mathbf{J}'$ in $\mathcal{D}$ for which $\mathbf{J}'$ is a permutation of $\mathbf{J}$, $G(\mathbf{J}) = G(\mathbf{J}')$.

Systematicity. For any two profiles $\mathbf{J}$ and $\mathbf{J}'$ in $\mathcal{D}$ and any two propositions p and q in A, if for each individual i, p is in J_i if and only if q is in J_i', then p is in $G(\mathbf{J})$ if and only if q is in $G(\mathbf{J}')$.

Systematicity combines an *independence* condition that requires the collective judgment on a proposition to only depend on the individual judgments on it and a *neutrality* condition that the aggregation procedure must not take account of the names of propositions. Christian List and Philip Pettit have shown that these four axioms are inconsistent if A contains two distinct atomic propositions p and q and either $p \wedge q$, $p \vee q$, or $p \rightarrow q$. There are many generalizations of this impossibility theorem. A number of possibility theorems have also been obtained by relaxing these axioms.

9. Other Issues

The literature on axiomatic models of resource allocation developed by Hervé Moulin, William Thomson, and many others employs a number of axioms that can only be defined using the special structure provided by economic models. By dispensing with the

Arrovian independence axiom in either its social welfare function or social choice correspondence formulation, many allocation mechanisms with nice properties have been characterized. This framework has been used to examine the foundations of egalitarian theories and to investigate the implications of different fairness criteria. For example, Ronald Dworkin's idea that individuals should be compensated for inequalities due to factors beyond their control, but not for choices for which they are responsible, has been formalized and explored in a number of different contexts by Marc Fleurbaey, François Maniquet, and John Roemer.

Fleurbaey (2007) has argued that it is not necessary to resort to interpersonal utility comparisons to avoid Arrovian impossibilities. Some fairness principles rely on comparisons that involve what Arrow would regard as being irrelevant alternatives. By weakening Arrow's independence axiom so as to permit such comparisons, possibility theorems can be obtained using only individual preferences to determine social preferences. Because of their use of both an independence condition and fairness criteria, these results provide a link between Arrovian social choice on economic domains and axiomatic models of resource allocation.

Michel Balinski and Rida Laraki (2010) have proposed an alternative approach to collective decision-making in which individuals assign grades to alternatives. The possible grades (e.g. excellent, good, etc.) are assumed to have commonly understood meanings. Grades allow individuals to evaluate outcomes, not just to rank them. Balinski and Laraki advocate a specific way of determining an overall grade from the individual grades, the *smallest median*. This grade is the median value of the grades when this median is unique and the smaller of the two possible medians otherwise. The overall grades can be used to determine a choice from a set of alternatives or to rank order them. The usefulness of this approach was tested in a pilot project conducted during the 2007 French Presidential elections.

Judgment aggregation raises epistemic considerations, but it is not supposed that there is necessarily a correct judgment. In contrast, a jury engages in a collective decision whose objective is to make a correct decision about the guilt of the defendant. This raises the problem of how well different decision procedures track the truth. The *Condorcet jury theorem* shows that larger juries are more likely to make the correct decision using majority rule when there are only two possible choices if jurors make their decisions independently and have a better than even chance of making the correct decision.

The authority and legitimacy of the jury system does not rely on juries always making the correct decision but, rather, on its tendency to make correct decisions. David Estlund (2008) believes that the authority and legitimacy of democratic procedures also depend on their tendency to make good decisions, subject to an acceptability requirement that prevents delegating authority to Platonic philosopher kings. He supposes that there are objectively good public decisions—political truths—which the political process should be designed to identify. This view of collective decision-making is the polar opposite to the traditional social choice approach in which there are no truth-based standards for making good decisions. In practice, polities deal with both of these kinds of decisions.[1]

Related Topics

Utilitarianism and Consequentialism; Rational Choice Theory

Note

1 Citations to the literature discussed here that are not in the References can be found in the Further Reading.

References

Arrow, K. J. (1951) *Social Choice and Individual Values*, New York: Wiley. Second Edition, 1963.

Balinski, M. and R. Laraki (2010) *Majority Judgment: Measuring, Ranking, and Electing*, Cambridge, MA: MIT Press.

Estlund, D. M. (2008) *Democratic Authority: A Philosophical Framework*, Princeton, NJ: Princeton University Press.

Fleurbaey, M. (2007) "Social Choice and Just Institutions: New Perspectives," *Economics and Philosophy* 23: 15–43.

Weymark, J. A. (1991) "A Reconsideration of the Harsanyi-Sen Debate on Utilitarianism," in J. Elster and J. E. Roemer (eds.), *Interpersonal Comparisons of Well-Being*, Cambridge: Cambridge University Press: 255–320.

Further Reading

Good textbooks are W. Gaertner, *A Primer in Social Choice Theory*, Second Edition (Oxford: Oxford University Press, 2009); J. E. Roemer, *Theories of Distributive Justice* (Cambridge, MA: Harvard University Press, 1996); and A. K. Sen, *Collective Choice and Social Welfare* (San Francisco, CA: Holden-Day, 1970). Detailed surveys can be found in K. J. Arrow, A. K. Sen, and K. Suzumura (eds.), *Handbook of Social Choice and Welfare* (Amsterdam, North-Holland; Vol. 1, 2002; Vol. 2, 2011). W. Bossert and J. A. Weymark, "Utility in Social Choice," in S. Barberà, P. J. Hammond, and C. Seidl (eds.), *Handbook of Utility Theory: Vol. 2, Extensions* (Boston, MA: Kluwer, 2004), 1099–1177, provides a comprehensive survey of social welfare functionals. The canonical reference on population issues is C. Blackorby, W. Bossert, and D. Donaldson, *Population Issues in Social Choice Theory, Welfare Economics, and Ethics* (Cambridge: Cambridge University Press, 2005). A good introduction to judgment aggregation is provided by C. List, "The Theory of Judgment Aggregation: An Introductory Review," *Synthese* (2012) 187: 179–207. A. Sen, *Choice, Welfare and Measurement* (Cambridge, MA: MIT Press) collects Sen's most important articles on social choice theory. D. G. Saari, *Decisions and Elections: Explaining the Unexpected* (Cambridge: Cambridge University Press, 2001) offers a novel interpretation of voting paradoxes as well as his arguments in support of point voting.

60
RATIONAL CHOICE THEORY

Peter Vanderschraaf

Moral and political philosophy is motivated in large part by attempts to rationally justify the choices individuals make and the institutions that can, and often do, govern them. But what counts as a rational justification of a choice or an institution? *Rational choice theory*, also known as *normative decision theory*, is the branch of philosophy and social science that seeks to provide a systematic account of how agents should choose from among alternative options in given contexts. As the names suggest, decision theory has a descriptive as well as a normative side. Descriptive decision theory seeks to explain and to predict how agents *do* choose. Decision theory has ancient roots. Plato and Aristotle, and later Hobbes and Hume, are but a few of the giants of moral and political philosophy who discuss and apply important elements of normative and descriptive decision theory. But decision theory became a field in its own right only in the twentieth century with the development of the first mathematically rigorous accounts of rational choice and a body of experimental studies that tested these accounts.

Since the 1950s decision theory, and in particular rational choice theory, has played an important role in the analysis of the social contract, understood in a generic sense as a system of norms and other institutions that can regulate the conduct of a society's members. In this chapter, I review two main ways social and political philosophy and rational choice theory interact. In Section 1, I discuss alternative proposals for deriving part of the social contract, its principles of distributive justice, by considering what a single rational agent chooses in special circumstances. In Section 2, I discuss the analysis of norms through the lens of *game theory*, where multiple agents choose and jointly produce social outcomes. I keep the focus of this chapter on rational choice theory, reflecting the fact that so much of social philosophy is normative in orientation. However, I do not ignore descriptive decision theory, since that part of decision theory explores the limits of the alternatives actual individuals can be expected to follow.

1. The Social Contract and Individual Rational Choice

In the 1950s, John Harsanyi (1953, 1955) and John Rawls (1958) independently proposed principles of distributive justice that each defended as the products of rational choice. While Harsanyi and Rawls differed sharply over the specific principles they championed, both argued for their principles from an approach in the *impartial spectator*

tradition of Adam Smith (1759/1982). To illustrate their approach, I will use a story inspired by the plot of the 1938 film *If I Were King*: King Louis has grown annoyed at complaints raised against his regime. Louis orders François Villon, a commoner subject who is also one of Louis' most outspoken critics, to propose a new rule for distributing the benefits and burdens of the realm among the subjects. Louis promises to enforce any rule Villon proposes satisfying one condition: Villon's proposed rule must result in a ranking of the subjects, where rank 1 refers to the subject having the highest social position in terms of benefits and burdens, rank 2 refers to the subject having the second highest social position, and so on, and Louis will assign one of these ranks to each subject only after Villon proposes the new rule. Villon realizes at once that his decision will determine his own social position along with his fellow subjects' social positions. But Villon also realizes Louis has effectively prevented Villon from knowing *which* social position he or anyone else will be in before he chooses.

Villon must choose a crucial part of a social contract, a rule of distributive justice, while he is in effect constrained by a *veil of ignorance* preventing him from knowing in advance anyone's place in the resulting distribution. Harsanyi and Rawls argue that principles regulating distribution chosen by rational agents restricted by the appropriate veil of ignorance are principles of justice because the veil ensures an impartial choice. The individual rational choice approach to this sort of problem has an agent like Villon apply extensions of rules for decision that a rational agent might use in more ordinary circumstances. For a simple example of this approach, suppose Villon can choose one of four rules, each of which produces a different distribution across five ranked subjects. Figure 1 summarizes these alternatives.

The Figure 1 matrix entries represent an index of absolute shares of goods that determine social position. In a typical single-agent decision problem, the agent chooses from among a set of alternative *acts*, aiming to achieve a desired member of a set of *outcomes* determined by the acts and relevant *states*. In this example, the acts are the

Rules for Distribution	Subjects				
	s_1	s_2	s_3	s_4	s_5
A_1	28	25	20	15	12
A_2	35	30	25	15	10
A_3	30	29	28	27	6
A_4	25	22	19	16	13

Figure 1 Distribution Problem with Absolute Shares

members of the set $\{A_1, A_2, A_3, A_4\}$ of rules and the states are the members of the set $\{s_1, s_2, s_3, s_4, s_5\}$ of ranked subjects. Villon must choose some $A \in \{A_1, A_2, A_3, A_4\}$ without knowing in advance the state $s \in \{s_1, s_2, s_3, s_4, s_5\}$ that determines his own position.

Contemporary rational choice theory is based upon the idea that an agent can, and should, decide in a manner consistent with this agent's preferences. According to one approach to rational choice, Villon can start by assigning a numerical index or *ordinal utility function* to the alternative outcomes that strictly increases according to his preference ranking over these outcomes. Villon always strictly prefers a greater over a lesser quantity of the goods that determine social rank. So the Figure 1 matrix summarizes an ordinal utility function for Villon, though this is clearly only one of many such functions derivable from Villon's preferences. Given no information regarding probabilities of the various outcomes, the Figure 1 matrix characterizes a decision problem under *ignorance* or *uncertainty* (Luce and Raiffa 1957). Decision theorists have proposed a variety of rules for decisions under uncertainty. One of the most important of these rules, and one that has also become influential in political philosophy, is the *maximin* rule: Identify the minimum ordinal utility possible given each act, then choose the act yielding the maximum of these minimums. Decision theorists who favor maximin do so in large part because one who follows this rule achieves at least one's best guaranteed outcome given one's choice. Maximin choosers "play it safe." In the Figure 1 Distribution Problem, A_4 is the maximin choice, and produces the distribution yielding the lowest ranked subject s_5 her best possible outcome. In early presentations of his theory of *justice as fairness* (1958, 1971), Rawls stressed his belief that a person constrained by a veil of ignorance would be *risk averse*, and would choose principles of justice as if society were designed so that his enemy assigns him his social position. The *difference principle* of Rawls' theory assigns goods that determine social position in a manner consistent with maximin.

An ordinal utility function fails to capture information crucially important to most decision problems, namely, the relative desirability of the alternative outcomes. The probabilities of outcomes are also crucially important to most decision problems. A long rational choice tradition maintains one should choose an option yielding the greatest *expected value*, defined in terms of relative desirability and likelihood of outcomes. One can state a principle of *Bayesian rationality* precisely in terms of the agent's *subjective probability distribution* over the outcomes, reflecting her degrees of belief, and her *cardinal utilities* over the outcomes, reflecting her relative intensity of preferences. A Bayesian determines the expected cardinal utility of each alternative act given her probabilities over the outcomes, then chooses an act having the maximum expected utility. Given the relevant probabilities and cardinal utilities, a decision problem becomes a problem of *risk* rather than uncertainty. In the twentieth century, various authors, including Ramsey (1926/1931), von Neumann and Morgenstern (1944), Savage (1954) and Anscombe and Aumann (1963), put Bayesian rationality on a rigorous footing by proving *representation theorems* showing that an agent with sufficiently rich and coherent preferences on the *lotteries* over the possible outcomes can derive cardinal utilities and even subjective probabilities over these outcomes. Today economists tend to use "utility" synonymously with "cardinal utility," and either assume agents choose according to Bayesian rationality or explain why these agents adopt some other standard.

Suppose Villon considers analyzing his problem as one of risk rather than uncertainty. Then he needs probabilities over the states in $\{s_1, s_2, s_3, s_4, s_5\}$ and utilities over the outcomes. But Louis has effectively created a veil of ignorance rendering Villon an impartial spectator with respect to the Distribution Problem. To derive his utilities *as*

impartial spectator, Villon must somehow take into account the preferences of all the subjects. Villon needs a method for aggregating the subjects' preferences into a single *social welfare function*. Harsanyi (1955) proved a representation theorem establishing such a method. Harsanyi began by assuming each individual Agent *i* in a given society of *n* agents has preferences sufficiently coherent and rich to derive her own utilities over the outcomes, represented as a function $u_i(.)$ where $u_i(x)$ is Agent *i*'s cardinal utility at outcome *x*. Harsanyi next assumed that some agent assuming the role of impartial spectator has *social preferences* that satisfy the same richness and coherence conditions the individuals' preferences satisfy. Such social preferences imply the existence of a social welfare function $W(.)$ where $W(x)$ is the utility of the impartial spectator at outcome *x*. But Harsanyi went one step further, and required that the social preferences respect a *Pareto condition*: Given alternatives *x* and *y*, if $u_i(x) = u_i(y)$ for each Agent *i*, then $W(x) = W(y)$, and if $u_i(x) \geq u_i(y)$ for each Agent *i* and $u_i(x) > u_i(y)$ for some Agent *i*, then $W(x) > W(y)$. Harsanyi noted that the Pareto condition in effect sets a moral constraint upon the impartial spectator's preferences, but maintained this is a weak and uncontroversial constraint (1955, 1977). Harsanyi proved that the resulting social welfare function is a weighted sum of the individual utility functions, that is, for a given outcome *x*,

$$W(x) = \alpha_1 u(x) + \ldots + \alpha_n u_n(x), \text{ where } \alpha_i > 0 \text{ for each Agent } i.$$

Harsanyi also showed that the weights are equal, that is, $\alpha_1 = \ldots = \alpha_n$, if the impartial spectator can express the individuals' utility functions in equal units on the basis of *interpersonal utility comparisons*. From Harsanyi's premises, Bayesian rationality applied across the various preferences of a society's members becomes a form of utilitarianism. For example, suppose the utility function of each subject *i* is characterized by $u_i(X) = log(X)$ where *X* is the number of shares, reflecting a diminishing marginal utility for additional shares. After comparing utilities and rescaling, Villon has the alternate values for the individual utilities and the social welfare function summarized by the Figure 2 matrix.

		Subjects					
		s_1	s_2	s_3	s_4	s_5	$W(\boldsymbol{x}) = \frac{1}{5} \cdot \sum_{i=1}^{5} u_i(\boldsymbol{x})$
Rules for Distribution	A_1	3.3322	3.2189	2.9957	2.7081	2.4849	2.9480
	A_2	3.5553	3.4012	3.2189	2.7081	2.3026	3.0372
	A_3	3.4012	3.3673	3.3322	3.2958	1.7918	3.0377
	A_4	3.2189	3.0190	2.9444	2.7726	2.5649	2.9184

Figure 2 Distribution Problem with Utilities

The equal weights of the sum defining the social welfare function reflect the famous slogan of classical utilitarianism: "Everybody to count for one, nobody for more than one." One can alternatively interpret the weights as defining probabilities over outcomes. If Villon regards each alternative rank Louis might assign him equally likely, then the weight for each of the five subjects is 1/5. Viewed this way, Villon's Bayesian choice is A_3, the utilitarian choice. The particular functional form of the individual utilities in this example is not essential. The Bayesian choice results in utilitarianism in a Harsanyi representation of the social welfare function for any set of individual utility functions that strictly increase according to each individual's strict preferences over outcomes.

Harsanyi's preference utilitarianism and Rawls' difference principle remain the best known principles for distributive justice associated with individual rational choice, and their veil of ignorance approach raises questions that have helped spur a body of work that continues to grow in several directions. Here I will briefly discuss some of the issues and work motivated by three of these questions: (1) By what yardstick should a social distribution be measured? (2) How should a rational agent choose behind a veil of ignorance? (3) How are members of actual societies bound by the result of a certain individual rational choice?

Harsanyi builds his utilitarianism upon largely unrestricted individual preferences, while Rawls uses an index of absolute shares of primary social goods as the basis of his difference principle. A number of authors propose alternate social yardsticks they argue are more appropriate for analyzing distributive justice. To give just two examples, Dworkin (1981a, b) defines a just distribution solely in terms of certain *resources*, roughly those things people are not responsible for, while Sen (1985, 1992) defines distributive justice in terms of *capabilities* enabling individuals to *function* in various ways. The object of rational choice will, of course, vary both according to the goods at stake and the conditions of the choice setting. Dworkin, for instance, uses the device of a hypothetical auction somewhat similar to a choice under a veil of ignorance to argue for a principle of *resource egalitarianism*. Rawls (1971) argues at length that agents constrained by the veil of ignorance of his theory would be unable to apply orthodox Bayesian rationality, because Rawls' veil deprives agents of knowledge of the particulars of their society and also of any of their personal characteristics. Harsanyi (1975) and others, including Barry (1989), Binmore (1994) and Bailey (1997), argue that Rawls' case for the difference principle is unconvincing. These authors contend that behind a veil of ignorance as "thick" as Rawls constructs it, gambling on the results of utilitarianism is at least as rational, if not more rational, than gambling on the results of maximin. But even if one accepts the idea of preference utilitarianism in principle, serious conceptual problems remain. Resnik (1987) points out that Harsanyi's utilitarianism presupposes the space of social outcomes has structural properties that might not obtain in many practical applications. And preference utilitarianism is predicated upon interpersonal utility comparisons, perhaps the most controversial idea in the whole history of utilitarianism. Debates over the legitimacy of interpersonal utility comparisons have created something of an industry in philosophy, although recent and remarkable work inspired by Harsanyi himself on utilities constructed from *extended* or *empathetic preferences* (Binmore 1989, 1994, 1998; Fleurbaey *et al.* 2008) might finally swing the tide in favor of the idea.

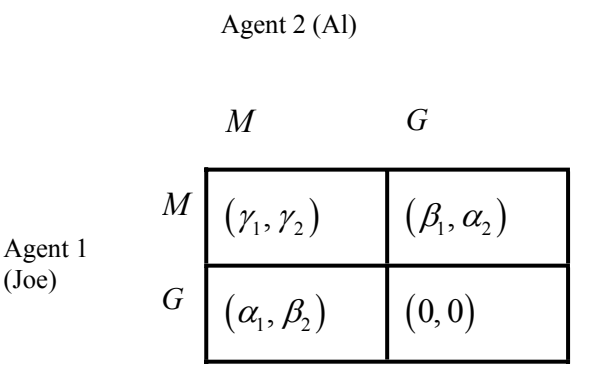
Finally, one might question how the individual rational choice approach to distributive justice relates to actual societies. Villon can regard his deliberations as worthwhile, secure in the knowledge that Louis will enforce his choice. But why would members of

a real society with no preexisting enforcement mechanism follow the principles chosen behind some hypothetical veil of ignorance? Raising this question echoes part of Hume's famous critique of the social contract (1742/2006), where he argues that we are not bound by the terms of some contract none of us in fact participated in creating. Interestingly, empirical work suggests that actual individuals might find the principles chosen by an ideally "rational" agent constrained by a veil of ignorance unacceptable. In a striking series of experiments, Frohlich, Oppenheimer and Eavey (1987, 1992) assigned groups of students the problem of agreeing upon a rule for distributive justice subject to the constraint that no one knew her future position in the distribution. Figure 1 summarizes the distributions in one of these experiments, where shares were interpreted as thousands of dollars. In these experiments neither maximin nor utilitarianism gained much support, and the predominant rule chosen was A_2, which can be described as: maximize the average subject to a floor constraint. Even if one disregards such empirical findings, the problem remains that the outcome of the rational choice varies according to background conditions. One might maintain that for the given story, Villon's rational choice is utilitarianism, but Binmore (1994, 1998) points out that maximin becomes the rational choice if one changes the story so that any subject can veto a choice made behind the veil of ignorance and repeat the process herself. Which stories should philosophers use to underwrite rational choice analyses of justice? In short, principles of individual rational choice are a valuable tool for analyzing the social contract, but they are only a tool. Philosophers who use particular principles of rational choice need to explain why they use these tools and why real individuals should follow the choices these tools generate.

2. The Social Contract and Game Theory

Game theory is the part of decision theory that considers how the decisions of multiple agents can interact so as to produce various social outcomes. In 1954, Richard Braithwaite presented a lecture at the University of Cambridge entitled *Theory of Games as a Tool for the Moral Philosopher* that launched a new tradition of moral and political philosophy that incorporates elements of game theory. Von Neumann and Morgenstern (1944) and Nash (1950a, 1950b, 1951, 1953) established game theory as an important branch of the social sciences in the mid-twentieth century. However, they had intellectual precursors such as Zermelo (1913) and Borel (1921, 1924), who gave mathematical analyses of certain game theoretic problems, and even pre-twentieth-century philosophers such as Hobbes, Hume, Locke and Rousseau, who had important informal proto-game theoretic insights in some of their works. I will use another story, inspired by the plot of another film, from the 1960s, *The Apartment*, to illustrate parts of the game theoretic approach to analyzing norms: Joe Dobisch and Al Kirkeby lead active social lives. Each dates women he cannot bring to his home, and can enjoy the company of one of his romantic partners in private only in a certain apartment rented by C. C. Baxter, a fellow employee of Al's and Joe's company. Baxter leaves his apartment for part of each evening so that Joe or Al may bring a date there, but Baxter leaves it to Al and Joe to decide between themselves who uses the apartment during the appointed time when Baxter is away. This presents Al and Joe with a serious problem. Each most wants to use the apartment undisturbed by the other as often as possible. If on a given evening neither tries to use the apartment, both their social lives suffer. But if both try to use the apartment on the same evening, they fight, and neither prefers fighting over "giving in" so that the other can have an uninterrupted private meeting.

Al's and Joe's predicament is an especially simple example of a set of individuals in *circumstances of justice* (Rawls 1971; Barry 1989). Such individuals must decide collectively how to distribute some relatively scarce good or goods among themselves, knowing not everyone can receive as much as he wants. Here evening time for exclusive use of the apartment is the limited resource at stake. The Figure 3 matrix summarizes the apartment problem for a given evening.



M = claim nothing, G = claim all

$\alpha_i > \beta_i > 0$, $\alpha_i > \gamma_i$, $i \in \{1, 2\}$

Figure 3 Division

This matrix characterizes a game in *strategic form*, where M and G are the agents' *pure strategies*. The first (second) coordinate of each ordered pair or *payoff vector* is Agent 1's (Agent 2's) payoff. The payoffs are cardinal utilities reflecting the relative desirability of the possible outcomes for each agent. It is important to note that while each agent in a strategic form game has his own payoff function, game theory does not in general presuppose interpersonal utility comparisons across the agents. But in order to predict any outcome in a game, one must ascribe some knowledge and some standard of rationality to the agent or agents who are to determine the outcome. Game theorists typically assume at minimum that the agents engaged in a game know the payoff structure of the game and are Bayesian rational. Indeed, game theorists generally use "rational" synonymously with "Bayesian rational." As we shall see shortly, these assumptions alone are often necessary but not sufficient for predicting the outcomes of games.

In the Figure 3 Division game, a rational agent follows his part of (M, G) if he expects the other to also follow his part of (M, G), and follows his part of (G, M) if he expects the other to follow his part of (G, M). (M, G) and (G, M) are *Nash equilibria* (1950a, 1951), that is, at either of these outcomes each agent follows a best response in terms of expected payoff given that the other follows this outcome. Noncooperative games in general can have any finite number of agents, each having a set of pure strategies not necessarily the same as those of her counterpart agents. Nash showed that if each agent in a game has finitely many pure strategies, then this game has at least one Nash equilibrium if agents can *mix* strategies by pegging their pure strategies on probabilistically independent random devices. But even in games as simple as Division, it is no easy

matter predicting which outcome rational agents will follow. Joe and Al can coordinate on either (M, G) or (G, M), but Joe strictly prefers (G, M) while Al strictly prefers (M, G). Division is a *conflictual coordination game* where the agents have conflicting preferences over a variety of available Nash equilibria. Vanderschraaf (2006) argues that in general a community of individuals is in circumstances of justice precisely when their interactions have the structure of an appropriate conflictual coordination game.

Nash based his equilibrium concept upon mixed strategies. Lewis (1969) and Aumann (1974, 1987) explored the consequences when agents follow *contingency strategies* predicated upon certain clues in their environment. Contingency strategies include mixed strategies as a special case, since agents' contingency strategies can be either probabilistically independent or correlated. One set of contingency strategies Joe and Al can follow in Division is to flip a coin that lands "heads-up" with probability λ, then follow (G, M) if the coin lands "heads-up" and (M, G) if the coin lands "tails-up." If either Joe or Al deviates unilaterally from this "coin-flip" strategy system, his expected payoff is lower than his expected payoff if he conforms. Since conforming is each agent's best response given that the other conforms, this system characterizes a *correlated equilibrium* (Aumann 1974, 1987). Suppose Al and Joe know that both are rational, that both know the payoff structure of the game, and that both will follow the "coin-flip" strategy system described here. If each can also infer the consequences of this mutual knowledge, they have *common knowledge* (Lewis 1969) that they will follow this equilibrium, and no other. An equilibrium of a game with multiple equilibria that agents follow given that they have this kind of common knowledge is a *convention* (Lewis 1969; Cubitt and Sugden 2003). Several contemporary authors, including especially Gauthier (1979, 1986), Sugden (1986/2004), Taylor (1987), Binmore (1994, 1998) and Skyrms (1996), present accounts of moral norms as distinguished conventions of corresponding conflictual coordination games. Gauthier, Sugden and Taylor were among the first to argue that analyzing moral norms in terms of game theoretic conventions continues and extends the moral conventionalist tradition of Grotius, Hobbes and Hume.

For a concrete example, suppose that, in the Figure 3 game, Joe's payoffs are defined by $\alpha_1 = 6$, $\beta_1 = 3$ and $\gamma_1 = 1$ and Al's payoffs are defined by $\alpha_2 = 8$, $\beta_2 = 1$ and $\gamma_2 = -1$. Then their game is summarized by the Figure 4 matrix.

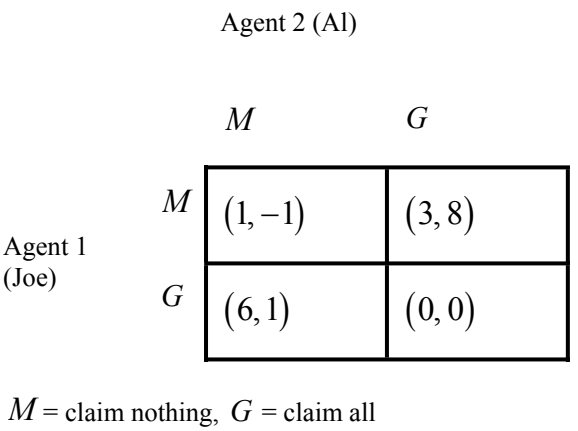


Figure 4 Braithwaite Game

This game is structurally equivalent to the game Braithwaite uses as a motivating example in his Cambridge lecture. Here neither Joe nor Al prefers that the apartment goes unused over the other using the apartment undisturbed, and Joe, but not Al, regards a fight as the worst possible outcome. Which outcome should they follow? Any "solution" to a noncooperative game should be an equilibrium, since by definition only at an equilibrium is each agent's chosen strategy a rational response to the other agents' chosen strategies. But this restriction evidently casts too wide a net. The agents in a conflictual coordination game can alternate between each agent's most preferred outcome according to any probability distribution, so any such game has infinitely many different correlated equilibria. Agents like Joe and Al need some rule or system of rules for *equilibrium selection*. Upon reflection Al and Joe are likely to conclude that not just any equilibrium will serve as a solution to their division problem. They might agree that the solution they ultimately follow should be *fair*. But given the asymmetries in this game, Al and Joe might have some difficulty agreeing upon proper criteria of fairness.

Braithwaite found inspiration in Nash's then recent analysis of the *bargaining problem* (Nash 1950b, 1953). Nash envisioned the resolution of division problems as a two-stage process. In stage one, agents determine what outcome ensues should they fail to agree upon some division of the good at stake. After establishing this *nonagreement point*, in stage two the agents select some allocation of the good from a set of allocations at which each does at least as well as he would at the nonagreement point. Nash, and Braithwaite after him, proceeded from the premise that at the nonagreement point, no agent would concede any of the good to another agent. In the Figure 4 game, this premise implies that (G, G) is the nonagreement point. Nash (1953) argued that in principle one can analyze any bargaining problem either *axiomatically* by considering which allocations satisfy certain formal desiderata, or from the perspective of agents who settle upon an allocation via some *bargaining process*. Nash thought that the axiomatic and bargaining approaches should yield consistent results. According to Nash's own proposed axiomatic solution, when payoffs are scaled so that each agent receives 0 at the nonagreement point, the good is divided so as to maximize the product of individual payoffs. In the Figure 4 Braithwaite game, at the Nash solution Joe and Al follow the (M, G) outcome 13/14 of the time and the (G, M) outcome 1/14 of the time, and achieve the expected payoff vector

$$\frac{1}{14}\cdot(6,1)+\frac{13}{14}\cdot(3,8)\approx(3.214,7.5).$$

Game theorists and philosophers have challenged Nash's analysis of the bargaining problem on several fronts and have proposed a variety of competing axiomatic theories of bargaining. For instance, Lucas (1959) and Gauthier (1986) define the nonagreement point in terms of strategies they contend each player would willingly follow in the absence of a binding agreement rather than upon threats to claim all of the good at stake, partly on the grounds that to follow such a threat strategy could prove costly to oneself in the end. Braithwaite, and later Gauthier (1986), defend a solution that for the 2-agent case is equivalent to a solution formalized by Kalai and Smorodinsky (1975). At the solution Braithwaite and Gauthier favor, after scaling payoffs so that all receive 0 at the nonagreement point, the ratio of the payoff an agent receives to his payoff if he gets all the resource is the same for each agent. Braithwaite's own proposed solution to the Figure 4 game has the agents follow (M, G) for 27/43 of the time and (G, M)

for 16/43 of the time, reaching the expected payoff vector (4.116, 5.395). Sen (1970) proposes a *leximin* rule closely related to the maximin criterion that Rawlsian distributive justice satisfies, where one first maximizes the utility of the worst-off individual or individuals, then among these maximizers, maximizes the utility of the next worst-off individual or individuals, and so on. Applied to the Braithwaite game where again (G, G) defines the nonagreement point, at the leximin solution the agents follow (M, G) for 16/23 of the time and (G, M) for 7/23 of the time, reaching the expected payoff vector (3.913, 5.870). The Nash, Kalai-Smorodinsky and leximin solutions to the Figure 4 game are depicted graphically in Figure 5.

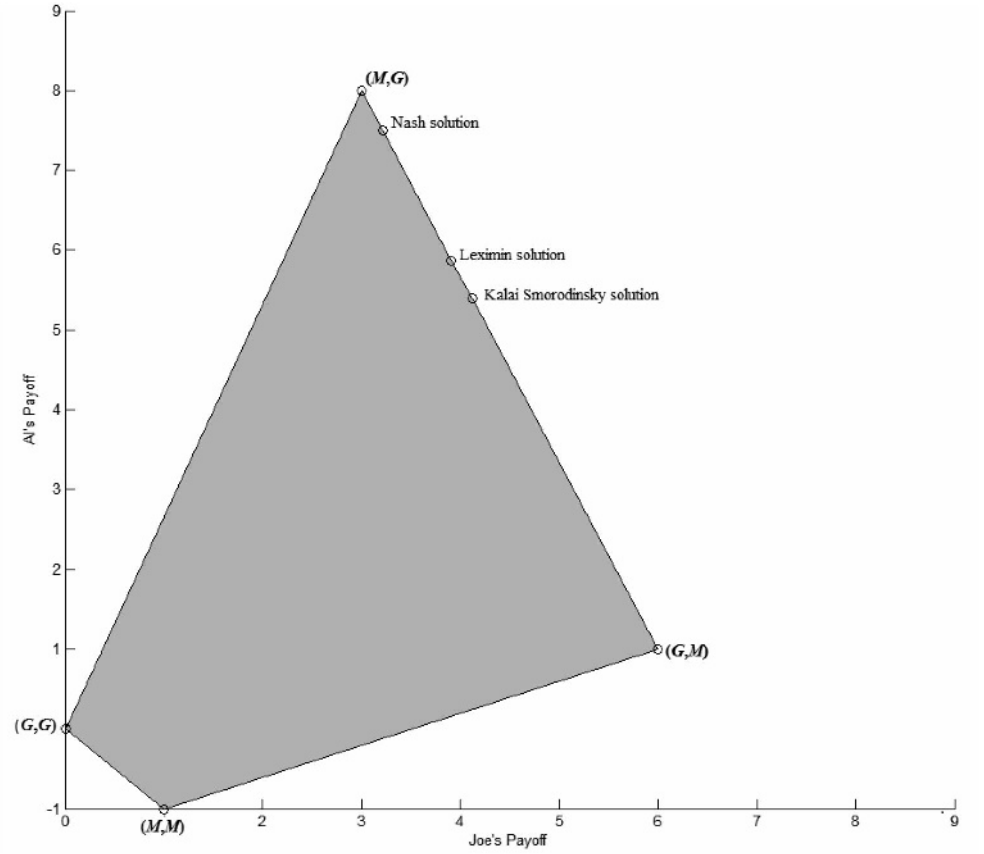


Figure 5 Three Axiomatic Solutions to the Braithwaite Bargaining Problem

Axiomatic bargaining theory is, today, a discipline in its own right, and whole families of alternative solutions to bargaining problems are discussed in works such as Roth (1979) and Thomson and Lensberg (1989). Bargaining analysis has made relatively few inroads into philosophy since Braithwaite's work, partly because the axioms that underwrite competing solutions are open to debate. For example, the axioms of the Kalai-Smorodinsky and the leximin solutions, but not those of the Nash solution, reflect the intuition that each individual in the group is to receive a certain minimum share, whereas the axioms of the Nash solution, but not those of the Kalai-Smorodinsky and

leximin solutions, require that one's share does not vary either according to one's choice of utility scale or according to the number of times the division rule is applied across subsets of the group. Game theorists also have made only limited progress towards completing the *Nash program* of relating the axiomatic and the bargaining approaches to division programs. Some of the central results of the *alternating offers* models of bargaining (Stähl 1972; Rubinstein 1982; Muthoo 1999) show that a pair of rational agents who exchange proposals for allocation successively will converge to the Nash solution if they know the amount of the good they can share diminishes at a certain *discount rate* with each offer. Recently, several authors have shown that *boundedly rational* agents who play a 2-agent bargaining problem repeatedly according to a dynamic adjustment process can converge into either the Nash or the Kalai-Smorodinsky solution depending upon the specifics of the dynamics (Young 1998; Binmore *et al.* 2003; Robles 2008). The ultimate impact of bargaining theory upon social philosophy will depend in part upon how the Nash program continues to progress.

Rules for allocating shares of divisible goods justly are part of a moral code, whose requirements philosophers often refer to as precepts of *natural law*. Several of the giants of the early modern natural law tradition, including especially Grotius, Hobbes and Hume, presented natural law precepts in a conditional form that one can summarize as follows, where φ stands for a certain form of conduct: One must φ (or refrain from φ with respect to others if, and only if, one expects the others to φ (or to refrain from φ) with respect to oneself. Some interpret early modern natural law theorists such as Hobbes and Hume as presenting conventionalist accounts of large parts of the natural law (Gauthier 1979; Hampton 1986; Kavka 1986; Sobel 2009). In the spirit of the early modern tradition, one can use game theory to formulate conventionalist natural law precepts. For another concrete example, suppose that in the Figure 3 game $\alpha_1 = \alpha_2 = 7$, $\beta_1 = \beta_2 = 2$ and $\gamma_1 = \gamma_2 = 6$. Figure 6 summarizes the resulting *Hawk–Dove* game.

Given the Apartment Problem interpretation, this game reflects a scenario where Al and Joe each prefer the apartment going unused on a given evening over the other using the apartment undisturbed, and where a fight is the worst outcome for both. Another closely related interpretation of Hawk-Dove is a situation where each agent wishes to

		Agent 2 (Al)	
		M	*G*
Agent 1 (Joe)	*M*	(6, 6)	(2, 7)
	G	(7, 2)	(0, 0)

M = claim nothing, *G* = claim all

Figure 6 Hawk–Dove

claim ownership of some indivisible good, and one of the agents becomes the owner by choosing G when the other chooses M. In this interpretation, each agent would rather follow (M, M), where neither claims the good so that each is free to use it, than give in to the other. In their accounts of natural law, Hobbes and Hume give specific requirements that determine ownership of indivisible goods in several specific situations. Both Hobbes and Hume argue that a previously unowned and indivisible good becomes the property of the first to possess this good (*Leviathan* 15: 27, 28; *Treatise* 3.2.3: 6–7). Hobbes adds that in some cases, presumably cases where a former owner has died and left no will, ownership of the contested good should be conveyed by primogeniture (*Leviathan* 15: 27, 28). The Hobbesian inheritance rule corresponds to a correlated equilibrium convention where Joe and Al follow (G, M) if Joe has priority and (M, G) if Al has priority according to primogeniture. The "finders keepers" rule corresponds to a different correlated equilibrium convention where, if Al and Joe each knows he is the first to possess the good at stake one third of the time, then they follow (G, M) when Joe knows he is first to possess, (M, G) when Al knows he is first to possess, and (M, M) the rest of the time.

As these examples illustrate, the specifics of conventionalist natural law precepts can vary somewhat according to context. But these natural law conventions set strict limits on how much more some members of society may benefit from the convention than others. Each member of society can expect to "win" some of the time even in cases where the benefits of the convention are asymmetrical. The primogeniture convention is one such asymmetrical natural law convention, because the expected benefits are biased in favor of the eldest male heirs. A solution to the bargaining problem can also characterize a possibly asymmetric natural law convention. A natural law convention is *ideal* when in fact each member of society expects to benefit equally if all follow the convention, as is the case for the "finders keepers" convention described here. In short, this conventionalist approach defines the requirements of the natural law as correlated equilibrium conventions that are nonarbitrary because they conform to a version of the "Golden Rule" prohibiting one from exploiting others (Vanderschraaf 2011).

In his Cambridge lecture, Braithwaite speculated that game theory might in time transform moral philosophy, much as statistics had transformed the social sciences. Such a transformation has yet to occur—Braithwaite thought it might take centuries—but game theory is now an important tool in social and political philosophy. One great advantage of the game theoretic approach is that this approach explicitly incorporates certain differences across individuals and analyzes how individuals interact in light of these differences. Game theory is well suited for analyzing a social contract chosen by individuals as they are, rather than how they might be in some counterfactual primordial state, such as the veil of ignorance states Harsanyi and Rawls envision. Game theory, and in particular the theory of *repeated games*, sheds new light upon the ancient problem of reconciling individuals' own interests with moral requirements. Given an analysis of norms as equilibria, one can argue that conforming to a moral norm serves one's interests in the long run, even if the prospects of an immediate gain occasionally tempt one to violate the norm (Taylor 1976; Skyrms 1998). Social scientists are now using game theory to explore to what extent, if any, moral norms are sustained by the altruistic interests of the people who follow them (Gintis *et al.* 2005). Moral norms can be interpreted as devices for equilibrium selection having considerable explanatory power (Binmore 1994, 1998). Given the huge number of available equilibria in even the simplest conflictual coordination games, it is a remarkable fact that people coordinate

successfully so often both in laboratory experiments and in real life on equilibria that characterize norms we consider parts of a just social contract. For example, in many symmetrical division problems, people tend to claim equal shares of the good at stake (Kagel and Roth 1995; Camerer 2003). They follow a special case of Aristotle's principle of distributive justice, namely, that equals are to receive equal shares (*Politics* 1282b18–22). Some argue that the norms of the social contract are best explained as the endpoints of *cultural evolution*, or more formally as the equilibria of appropriate games that emerge via the pressures of social dynamics (Sugden 1986/2004; Skyrms 1996, 2004). Philosophers have barely begun to integrate the relevant empirical findings and insights of evolutionary analysis into game theoretic accounts of the social contract. Likewise, philosophers are only starting to investigate large parts of game theory that may spur the transformation in thought Braithwaite foresaw. Much of *cooperative game theory*, where agents can form binding coalitions, and *mechanism design*, where one begins by first considering a desired outcome and then constructs an appropriate game with the aim of realizing this outcome, is unexplored territory for philosophers who analyze the social contract. In time some of the few who have appreciated the significance to moral and political philosophy of mechanism design (Roemer 1994, 1996; Binmore 1998) and of some of the concepts of cooperative game theory (Young 1994; Moulin 2003) may be viewed as the forerunners of a much larger and very important program.

Related Topics

Contractarianism, Utilitarianism and Consequentialism, The Difference Principle, Social Evolution, Social Choice Theory

References

Anscombe, F. J. and Aumann, R. (1963) "A Definition of Subjective Probability," *Annals of Mathematical Statistics* 34: 199–205.

Aristotle (1998) *Politics*, trans. C. D. C. Reeve, Indianapolis: Hackett Publishing Company.

Aumann, R. (1974) "Subjectivity and Correlation in Randomized Strategies," *Journal of Mathematical Economics* 1: 67–96.

—— (1987) "Correlated Equilibrium as an Expression of Bayesian Rationality," *Econometrica* 55: 1–18.

Bailey, J. W. (1997) *Utilitarianism, Institutions and Justice*, New York and Oxford: Oxford University Press.

Barry, B. (1989) *Theories of Justice*, Berkeley and Los Angeles: University of California Press.

Binmore, K. (1989) "Social Contract I: Harsanyi and Rawls," *Economic Journal* 99: 84–102.

—— (1994) *Game Theory and The Social Contract, Volume I: Playing Fair*, Cambridge, MA: MIT Press.

—— (1998) *Game Theory and The Social Contract, Volume II: Just Playing*, Cambridge, MA: MIT Press.

——, Samuelson, L. and Young, H. P. (2003) "Equilibrium Selection in Bargaining Models," *Games and Economic Behavior* 45: 296–328.

Borel, E. (1921) "La theorie du jeu les equations integrales a noyau symetrique," *Comptes Rendus del Academie* 173: 1304–1308.

—— (1924) "Sur les jeux oú interviennent le hasard et l'habileté des joueurs," in E. Borel, *Théorie des Probabilités*, Paris: Librairie Scientifique.

Braithwaite, R. (1955) *Theory of Games as a Tool for the Moral Philosopher*, Cambridge: Cambridge University Press.

Camerer, C. (2003) *Behavioral Game Theory: Experiments in Strategic Interaction*, Princeton, NJ: Princeton University Press.

Cubitt, R. and Sugden, R. (2003) "Common Knowledge, Salience and Convention: A Reconstruction of David Lewis' Game Theory," *Economics and Philosophy* 19: 175–210.

Dworkin, R. (1981a) "What is Equality? Part 1: Equality of Resources," *Philosophy and Public Affairs* 10: 185–246.

—— (1981b) "What is Equality? Part 2: Equality of Welfare," *Philosophy and Public Affairs* 10: 283–345.
Fleurbaey, M., Salles, M. and Weymark, J. (2008) *Justice, Political Liberalism, and Utilitarianism: Themes from Harsanyi and Rawls*, Cambridge: Cambridge University Press.
Frohlich, N., Oppenheimer, J. A. and Eavey, C. L. (1987) "Laboratory Results on Rawls' Distributive Justice," *British Journal of Political Science* 17: 1–21.
Frohlich, N. and Oppenheimer, J. A. (1992) *Choosing Justice: An Experimental Approach to Ethical Theory*, Berkeley: University of California Press.
Gauthier, D. (1979) "Thomas Hobbes: Moral Theorist," *Journal of Philosophy* 76, 547–559.
—— (1986) *Morals By Agreement*, Oxford: Clarendon Press.
Gintis, H., Bowles, S., Boyd, R. and Fehr, E., eds. (2005) *Moral Sentiments and Material Interests: The Foundations of Cooperation in Economic Life*, Cambridge, MA: MIT Press.
Hampton, J. (1986) *Hobbes and the Social Contract Tradition*, Cambridge: Cambridge University Press.
Harsanyi, J. (1953) "Cardinal Utility in Welfare Economics and the Theory of Risk Taking," *Journal of Political Economy* 61: 343–345.
—— (1955) "Cardinal Utility, Individualistic Ethics, and Interpersonal Comparisons of Utility," *Journal of Political Economy* 63: 309–321.
—— (1975) "Can the Maximin Principle Serve as a Basis for Morality? A Critique of John Rawls's Theory," *American Political Science Review* 69: 594–606.
—— (1977) *Rational Behavior and Equilibrium in Games and Social Situations*, Cambridge: Cambridge University Press.
Hobbes, T. (1651/1991) *Leviathan*, ed. Richard Tuck, Cambridge: Cambridge University Press.
Hume, D. (1740/2000) *A Treatise of Human Nature*, ed. David Fate Norton and Mary J. Norton, Oxford: Oxford University Press.
—— (1742/2006) "Of the Original Contract," in *Moral Philosophy*, ed. G. Sayre-McCord, Indianapolis: Hackett Publishing Company.
Kagel, J. and Roth, A. (1995) *The Handbook of Experimental Economics*, Princeton, NJ: Princeton University Press.
Kalai, E. and Smorodinsky, M. (1975) "Other Solutions to Nash's Bargaining Problem," *Econometrica* 16: 29–56.
Kavka, G. (1986) *Hobbesian Moral and Political Theory*, Princeton, NJ: Princeton University Press.
Lewis, D. (1969) *Convention: A Philosophical Study*, Cambridge, MA: Harvard University Press.
Luce, R. D. and Raiffa, H. (1957) *Games and Decisions: Introduction and Critical Survey*, New York: John Wiley and Sons.
Lucas, J. R. (1959) "Moralists and Gamesmen," *Philosophy* 34: 1–11.
Moulin, H. (2003) *Fair Division and Collective Welfare*, Cambridge, MA: MIT Press.
Muthoo, A. (1999) *Bargaining Theory with Applications*, Cambridge: Cambridge University Press.
Nash, J. (1950a) "Equilibrium Points in *n*-Person Games," *Proceedings of the National Academy of Sciences of the United States* 36: 48–49.
—— (1950b) "The Bargaining Problem," *Econometrica* 18: 155–162.
—— (1951) "Non-Cooperative Games," *Annals of Mathematics* 54: 286–295.
—— (1953) "Two-Person Cooperative Games," *Econometrica* 21: 128–140.
Ramsey, F. (1926/1931) "Truth and Probability," in *The Foundations of Mathematics and Other Essays*, ed. R. B. Braithwaite, New York: Harcourt Brace, 156–198.
Rawls, J. (1958) "Justice as Fairness," *Philosophical Review* 67: 164–194.
—— (1971) *A Theory of Justice*, Cambridge, MA: Harvard University Press.
Resnik, M. (1987) *Choices: An Introduction to Decision Theory*, Minneapolis: University of Minnesota Press.
Robles, J. (2008) "Evolution, Bargaining, and Time Preferences," *Economic Theory* 35: 19–36.
Roemer, J. (1994) *Egalitarian Perspectives: Essays in Philosophical Economics*, Cambridge: Cambridge University Press.
—— (1996) *Theories of Distributive Justice*, Cambridge, MA: Harvard University Press.
Roth, A. (1979) *Axiomatic Models of Bargaining*, Berlin: Springer Verlag.
Rubinstein, A. (1982) "Perfect Equilibrium in a Bargaining Model," *Econometrica* 50: 97–109.
Savage, L. J. (1954) *The Foundations of Statistics*, New York: John Wiley and Sons.
Sen, A. (1970) *Collective Choice and Social Welfare*, San Francisco, CA: Holden-Day, Inc.
—— (1985) *Commodities and Capabilities*, New York: Elsevier Science Publishing Company.
—— (1992) *Inequality Re-Examined*, Cambridge, MA: Harvard University Press.
Skyrms, B. (1996) *Evolution of the Social Contract*, Cambridge: Cambridge University Press.

—— (1998) "The Shadow of the Future," in *Rational Commitment and Social Justice: Essays for Gregory Kavka*, ed. J. L. Coleman and C. W. Morris, Cambridge: Cambridge University Press, 12–21.
—— (2004) *The Stag Hunt and the Evolution of Social Structure*, Cambridge: Cambridge University Press.
Smith, A. (1759/1982) *The Theory of Moral Sentiments*, Indianapolis: Liberty Fund, Inc.
Sobel, J. H. (2009) *Walls and Vaults: A Natural Science of Morals*, Hoboken, NJ: John Wiley and Sons, Inc.
Stähl, I. (1972) *Bargaining Theory*, Stockholm: Stockholm Economics Research Institute, Stockholm School of Economics.
Sugden, R. (1986/2004) *The Economics of Rights, Co-operation and Welfare*, 2nd edn., Houndsmills: Palgrave Macmillan.
Taylor, M. (1987) *The Possibility of Cooperation*, Cambridge: Cambridge University Press.
Thomson, W. and Lensberg, T. (1989) *Axiomatic Theory of Bargaining With a Variable Number of Agents*, Cambridge: Cambridge University Press.
Vanderschraaf, P. (2006) "The Circumstances of Justice," *Politics, Philosophy and Economics* 5: 321–351.
—— (2011) "Convention and Natural Law," unpublished manuscript.
Von Neumann, J. and Morgenstern, O. (1944) *Theory of Games and Economic Behavior*, Princeton, NJ: Princeton University Press.
Young, H. P. (1994) *Equity in Theory and Practice*, Princeton, NJ: Princeton University Press.
—— (1998) *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*, Princeton, NJ: Princeton University Press.
Zermelo, E. (1913) "Uber eine Anwendung der Mengenlehre auf die theorie des Schachspiels," *Proceedings, Fifth International Congress of Mathematicians* 2: 501–504.

Further Reading

G. F. Gaus gives an excellent and mostly informal discussion of some of the most important conceptual issues regarding the use of decision and game theory in social philosophy in *On Philosophy, Politics and Economics* (Belmont, CA: Wadsworth, 2008). M. Peterson gives a fine and elementary overview of the main parts of normative decision theory in *An Introduction to Decision Theory* (Cambridge: Cambridge University Press, 2009). Peterson includes chapters on game theory and social choice theory, complete proofs of the von Neumann-Morgenstern and the Harsanyi Representation Theorems, and some interesting critiques of the foundations of orthodox Bayesian decision theory that complement Gaus' discussion well. H. P. Young presents a self-contained, clear yet mathematically sophisticated survey of a variety of formal approaches to problems of fair division in *Equity in Theory and Practice* (Princeton, NJ: Princeton University Press, 1994). Young's book contains a wealth of fascinating examples and accompanying analyses ranging from rules for dividing a contested garment to progressive taxation schemes to *n*-player bargaining problems. D. Gauthier's *Morals By Agreement* (Oxford: Oxford University Press, 1986) and R. Sugden's *Economics of Rights, Cooperation and Welfare*, 2nd edn. (Houndsmills: Palgrave Macmillan, 2004/1986) are pioneering studies that analyze the social contract using elements of game theory. Gauthier uses game theory to develop an account of morality as the product of strategic rational choice. Sugden incorporates elements of evolutionary game theory to provide an explanation of the origins of norms and institutions. K. Binmore gives a magisterial theory of the social contract developed with game theory in his two-volume work with titles *Playing Fair* (Cambridge, MA: MIT Press, 1994) and *Just Playing* (Cambridge, MA: MIT Press, 1998). *Playing Fair* focuses primarily upon decision theoretic and game theoretic preliminaries, and includes much provocative discussion of other authors' earlier attempts to apply decision theory to moral and political philosophy. *Just Playing* concentrates more on Binmore's own original theory, and readers with a background in decision and game theory can read this book independently of *Playing Fair*. Binmore summarizes the main parts of his theory in the much shorter work *Natural Justice* (Oxford: Oxford University Press, 2005).

61
DISCOURSE THEORY

William Rehg

For the purposes of this chapter, "discourse theory" refers to two broad trends in social-political thought, distinguished by their opposing philosophical commitments. The postmodern discourse theory of thinkers such as Ernesto Laclau and Chantal Mouffe draws on poststructuralist and Marxist thought, emphasizes the role of power and contest in politics, and tends to be skeptical of consensus-oriented views of democratic deliberation. The rationalist discourse theory of Jürgen Habermas, Karl-Otto Apel, and other deliberative democratic theorists, by contrast, links classical modern political philosophy with the American pragmatist tradition—Peirce, Dewey, and Mead—and thus takes a more optimistic view of the normative powers of rational discourse for shaping legitimate political outcomes. This chapter focuses primarily on Habermas's discourse theory—the most comprehensive synthesis of the rationalist approach, and a useful framework for surveying the relevant issues, including the challenges posed by postmodern discourse theorists.

Habermas's philosophical passions were formed by intellectual and political currents in the postwar Germany of the 1950s and 1960s. The complicity of German philosophy in the rise of National Socialism, along with the growing democratic sensibility in the new Federal Republic, sparked an enduring interest in critical democratic theory. Meanwhile, wide-ranging philosophical explorations eventually led him to pragmatism and philosophy of language. Although one can see elements of a discourse theory in Habermas's earlier work (e.g. 1971; see McCarthy 1978), his mature position first took shape in the 1970s, and only reached its full expression in his 1992 treatise on democracy and law (*Between Facts and Norms*, 1996), which drew on his two-volume *Theory of Communicative Action* (1984/87; cf. 1979).

Habermas's mature research program rests on the coherence of four broad initiatives: a reconstruction of sociocultural evolution as a collective learning process; a developmental psychology and theory of personal identity-formation; a communicative theory of societal integration; and a deliberative conception of legal-political legitimacy. An idea of social enlightenment through rational discourse animates the whole. In what follows we focus on his analysis of discourse and its application to democratic theory.

1. Communicative Action

Habermas's turn to a discourse-theoretic framework in the 1970s was bound up with the problem of social order (2001a): how is society possible? This question requires an account of the ongoing mutual intelligibility and overall stability of social life. His

answer draws on a number of different sociological theories, but he gives primacy to communicative action, which he sharply distinguishes from strategic action.

Strategic actors are not so much interested in mutual understanding as in achieving the individual goals they each bring to the situation. Actor *A* will thus appeal to *B*'s desires and fears so as to motivate the behavior on *B*'s part that is required for *A*'s success. As reasons motivating *B*'s cooperation, *B*'s desires and fears are only contingently related to *A*'s goals. *B* cooperates with *A*, in other words, not because *B* finds *A*'s project inherently interesting or worthy, but because of what *B* gets out of the bargain: avoiding some threat that *A* can make or obtaining something *A* has promised (which may be of inherent interest to *B* but for *A* is only a means of motivating *B*).

Habermas recognizes that strategic action plays an important role in modern forms of societal integration, particularly in markets and bureaucratic administrations. However, he insists that strategic action ultimately depends on communicative forms of social integration. In communicative action, or what Habermas later came to call "strong communicative action" (1998b: ch. 7), speakers coordinate their action and pursuit of individual (or joint) goals on the basis of a shared understanding that the goals are inherently reasonable or merit-worthy. Whereas strategic action succeeds insofar as the actors achieve their individual goals, communicative action succeeds insofar as the actors freely agree that their goals (or shared goal) are reasonable, that they merit cooperative behavior. In communicative action, actors mobilize the potential for rationality given with ordinary language and its inherent telos of rationally motivated agreement (Habermas 1984: 99).

To spell out this rational potential of speech, Habermas enlists speech-act theory, which has its roots in Wittgenstein's later work and was developed by J. L. Austin and John Searle. Communicative interaction unfolds through the exchange of "speech acts," the use of language in which speakers perform social actions, understood as rule-governed moves in linguistically mediated interaction. In doing so, they establish a social relationship in which each recognizes the other as an accountable subject. Taking the speech-act offer-and-response as the starting point for his analysis of social interaction, Habermas goes on to advance a provocative thesis about its rational potential: the mutual intelligibility of a speech act depends on the actors' grasp of the kinds of reasons that make it acceptable (1998b: 232). If this is so, then speech acts are *inherently* open to potential criticism and justification on the basis of good reasons, and when we accept someone's speech-act offer, we tacitly assume that good reasons support it.

More precisely, speech acts involve "validity claims" that are open to criticism and defense. As situated communicative actions, speech acts typically presuppose three "world relations," each of which can be made explicit as a validity claim: the external world shared by the interlocutors, to which truth claims refer; their inner subjective worlds, to which sincerity claims refer; and the social relationship constituted by their interaction—for its part situated in a wider social world—to which rightness claims refer. Speech acts normally make only one validity claim explicit: a speaker states a fact about the objective world, expresses a desire, criticizes a choice as morally wrong, and so on. But speech acts can usually be criticized and defended along each of the three basic validity dimensions. For example, when a child asks his parents for a second piece of cake, they might reject it for making a false assumption about the objective world (e.g. the cake is gone), as insincere (he only wants to outdo his sister), or as wrong (giving him the cake would not be fair).

2. A Complex Theory of Argumentative Cogency

At the core of Habermas's theory of communicative action lies the thesis that ongoing social interaction requires actors continually to make and accept validity claims on the supposition that good reasons could, if necessary, be supplied to justify the claims. The validity basis of speech thus grounds the rationality of communicative interaction, and with it the rationality of social integration. In drawing attention to validity claims, Habermas moves beyond the narrow focus on truth and the fact-stating use of language, which dominated philosophy and the social sciences at mid-twentieth century. Moreover, he builds into social interaction an implicit reference to argumentation or discourse, the "reflective form" of communication in which actors put aside the pressures of action and take time to critically discuss a validity claim that has become problematic. Because the different kinds of validity claim require somewhat different types of reasons for their justification, we can distinguish different types of corresponding discourse. But first some remarks on the idea of discourse itself.

Habermas's understanding of discourse is inspired by initiatives in argumentation theory that aimed to recover the classical Greco-Roman understanding of arguments as embedded in social practices of debate and persuasion. Consequently, to understand the features in virtue of which arguments deserve acceptance, one must articulate the presuppositions that structure such practices and the internal goods at which they aim, namely the construction and evaluation of cogent arguments that make the rationality of a position manifest (Johnson 2000). To see how argumentation achieves this, one must understand not only the relevant logical standards, but also the dialectical and rhetorical demands of argument-making in context: how to assess an argument in relation to objections and competing arguments, and how to frame it for the particular audience.

To develop this analysis of argumentative practices, Habermas aligns the three classical perspectives—logic, dialectic, and rhetoric—with a more recent triadic distinction between product, procedure, and process. The argument itself, as a package of reasons supporting a conclusion, constitutes the *product* of argument-making practices, and is assessed according to its logical properties (not only formal validity, but also informal standards of substantive plausibility, adequate support, etc.). That product, however, typically issues from some kind of method or *procedure*, more or less formalized according to the social or institutional context, in which participants test possible reasons and arguments for their dialectical strength against objections. At the procedural level, arguers work within rules, whether tacit or explicit, strict or loose, governing the proper flow of discourse as a dialogical exchange (e.g. criticisms should be relevant to the topic, challenges call for replies). Such procedures always involve social *processes* in which participants with a particular set of background assumptions and psychological makeup attempt to convince each other (or an audience) of the better argument.

But which—or whose—standards count in the evaluation of cogency? Argument-making practices and their norms differ according to the social, cultural, and institutional context. Beyond very abstract standards, what makes a scientific argument logically cogent as a product differs from what makes an ethical or legal argument logically cogent; one can say the same for the different procedures that govern scientific argumentation, health-care ethics boards, law courts, and so on. And sometimes the relevant logical and dialectical standards are themselves subject to debate.

Because Habermas is keen to avoid a merely conventionalist or relativist understanding of discourse, he claims that these contextual variations are rational insofar as they

do not undermine or distort two culturally invariant structural features of discourse: the system of validity claims, and the necessary conditions of rational process. If all language use inherently rests on certain basic types of validity claim, and if each type of claim requires a different kind of discourse for its vindication, then domain-specific differences in justificatory aims, standards of relevance, and the like can be traced back to the validity claim(s) at issue in the domain. One can thereby account for, and critically scrutinize, a level of differences that legitimately vary with institutional context but not with culture. (This analysis assumes the cultures have passed through learning processes that freed the dynamics of communicative action from authoritarian confinement; see Owen (2002).) For example, it should not surprise us that scientific discourses focused on empirical truth employ somewhat different procedures, and different standards for assessing logical strength, than do criminal courts, where truth, rightness, and sincerity come together in arguments over the defendant's guilt. But within each domain, cultural variations in standards should not violate argumentative demands built into the type of validity at stake.

In looking to conditions of rational process as an invariant constraint, Habermas assumes that competent participants in argumentation operate with an implicit sense of the difference between genuinely rational processes of persuasion, in which one responds to arguments solely on their merits, and illicit forms of winning agreement through bullying, disinformation, rhetorical subterfuge, and the like (1984: 25). The rational force of arguments ultimately rests on this normative distinction that arguers themselves bring to their practice. This distinction, in turn, points to formal process conditions that characterize genuinely rational persuasion—and that participants allegedly presuppose in their practices of argumentation.

Habermas generally takes the "approximate" satisfaction of four conditions as the most crucial: (a) the discourse includes everyone able to make a relevant contribution, (b) the participants have equal voice, (c) they mean what they say, and thus are not hampered by internal psychological sources of unfreedom, and (d) the discourse is free of any external institutional constraints or coercive pressures arising from the structure of interaction (cf. 1990: 89; 2008: 82). In articulating these conditions, Habermas does not consider himself to be imposing normative standards on arguers from the outside but, rather, to be spelling out "unavoidable pragmatic presuppositions" that arguers themselves must adopt, at least tacitly, whenever they regard their argument-making as rational. Strictly speaking, these presuppositions are "counterfactual idealizations": they posit a horizon, as it were, a set of optimal conditions toward which genuine discourse aims, though we can never guarantee that our discourse has adequately satisfied their demands. Nonetheless, we must presume we have not obviously violated these conditions in ways that patently undermine the reasonableness of the outcome (Habermas 2003: 108). Should violations become apparent, then the outcome is open to reasonable doubt.

Habermas's idealizing account of rational discourse remains controversial. His confidence in his account rests partly on its intuitive appeal and partly on the self-refuting nature of statements that deny idealizing presuppositions (1990: 89–91). (Consider, for example, the following: "We rationally persuaded ourselves of claim X by excluding some people with relevant information on the issue"; "we all rationally agreed to Y by disallowing a relevant objection.") But one might also point out that institutionalized procedures often reflect these idealizing presuppositions, and thus offer evidence for their hold on us.

In any case, the logical and dialectical quality of discourse ultimately depends on the discursive process. This is because in real contexts of argumentation, such as science, policymaking, and law, the quality of an argument cannot be secured simply from the logical or dialectical perspectives. Few real arguments rest on self-evident premises from which one can straightforwardly deduce conclusions. Rather, one must critically test the quality of the supporting reasons in light of the available information and plausible challenges, and often one must weigh competing considerations. The quality of this procedural testing depends in turn on the social-psychological conditions under which the procedure is carried out—in other words, it depends on the quality of the rhetorical process of persuasion.

The pragmatic presuppositions of reasonable process thus provide a universal normative framework within which all discourse must take place in order to count as an "exercise in manifest rationality" (Johnson 2000: 1). Disputes about which validity claim is at issue, criticisms of existing procedures, revisions in substantive elements of argument evaluation—all such disputes affect how participants will judge the cogency of arguments, but in each case the standards of evaluation themselves are in question. What cannot be put in question, according to Habermas, are the ideal conditions of rational process. These are not immune, however, to contextual specification, according to the validity claim at issue—more on which below.

3. Types of Discourse

In Habermas's discourse theory, the vindication of validity claims varies with the type of claim according to (a) the relevant audience, (b) the expected level of consensus, and (c) the kind of reasons and arguments that support the claim (the relevant reasons). In fact, Habermas considers only two validity claims—empirical-theoretical truth and moral rightness—to admit of discourse in the strict sense, which is to say: discourse in which we could reach universal consensus on both the reasons and conclusion, were we able to pursue the matter under sufficiently ideal conditions for a sufficient length of time. By contrast, sincerity claims about a speaker's inner world of emotions, desires, and the like, are tested not with arguments in discourse but by comparing them with the speaker's overt behavior and choices; at issue is not only the possibility of deliberate deception but also that of self-deception. Insofar as sincerity becomes a topic of critical discussion, it tends to occur in therapeutic settings, or in what Habermas calls "therapeutic critique" (1984: 23).

More recently, Habermas has distinguished a number of validity claims beyond the basic three introduced above. Each type depends for its vindication on a corresponding form of discourse (1993: ch. 1; 2008: ch. 3). Decisions about means to an end call for *pragmatic discourses*, in which actors test what we might call an "efficacy claim," weighing the costs and benefits of alternative options and projecting likely future consequences in light of their knowledge of the empirical and social world. Significant personal decisions (e.g. the choice of vocation, marriage) ultimately pose the question of who one is and wants to be; the answer raises an "authenticity claim" that must be justified in an *ethical-existential* or *clinical discourse* that scrutinizes both the actor's self-knowledge (thus, sincerity) and her grasp of shared values. This kind of discourse has a collective analogue, which Habermas calls *ethical-political discourse*, in which a group or polity reflects on its history, traditions, and values in an attempt to identify collective goals that express "an authentic, collective self-understanding" (1996: 108, 95–7).

Ethical discourses differ from *moral discourses* in which participants test the universal validity of rightness claims—norms and obligations for treating persons with due concern and respect (1990: ch. 1; 1998a: ch. 1). As we shall see, this last distinction plays a crucial role in Habermas's political theory. In pluralistic societies, the idea of rational consensus can plausibly orient political deliberation only if citizens with different conceptions of personal authenticity, associated with different subpolitical ethical values, share a set of moral principles with both procedural and substantive implications, that is, implications for fair procedures, basic rights, and just outcomes ("the right"), that constrain the pursuit of self-interest, authenticity, and subpolitical values ("the good").

Idealized process conditions govern the rationality of each of these discourse types, but the latter, in turn, further specify those conditions. This is most obvious with the inclusivity condition, the requirement that all those capable of making a relevant contribution must be included. Because moral rightness claims, as Habermas understands them, have to do with the treatment of persons in general, moral discourses that focus on the justification of basic moral obligations and norms have a potentially universal audience. (Presumably, discourses that apply moral norms to concrete situations should include primarily those involved in the situation; see Habermas 1993: 35–9; Günther 1993). Scientific discourses about the truth of some theory or hypothesis normally must include those with the requisite disciplinary and topical expertise. In pragmatic discourses, experts are the most obvious contributors; but the fact that policy choices affect different groups of people means that these groups too have a legitimate place at the table. Ethical-existential discourses would apparently answer, at least in the first instance, to those who know the particular person and understand the relevant values and culture. And ethical-political discourses must include all those in the group whose values and goals are at issue.

Each of these discourse types also has implications for how much consensus we should expect. Participants in moral and empirical (scientific) discourses aim at a consensus of all participants, as noted above. But that might be too high an expectation for ethical discourses, given the highly interpretive and fluid character of values; for these discourses, it suffices that the participants succeed in eliminating implausible authenticity claims.

Finally, the relevant reasons also differ to some extent according to discourse type. Habermas has been clearest on this point for moral discourse. According to his moral universalization principle (U), if a norm is valid, then participants must be able, in principle, to accept the foreseeable consequences and side-effects of the general observance of that norm for the interests and value-orientations of each affected person (1998a: 42). The relevant reasons, then, include arguments that identify and justify (or criticize) likely effects of the norm on each person's (and group's) pursuit of goods (i.e. interests and values). Scientific discourses, by contrast, turn largely on experimental reports, mathematical arguments, and theoretical considerations. We should also expect that different types of discourse might employ different standards for ascertaining the burden of proof, identifying which position has presumption in its favor, and so on.

Does Habermas's system of types help us understand and evaluate actual discourses according to ideas of reasonableness that both participants and critical theorists allegedly share? Some of the most persistent criticisms of Habermas's discourse theory, even from sympathetic readers, aim at the ideals and idealizations on which it so heavily relies: not only the sharp demarcations between types of validity and discourses, but also the idealized conditions of rational discourse and optimistic assumptions about the

possibility of consensus. A good test for whether his discourse theory can meet such criticisms lies in his analysis of political discourse, where he attempts to bring ideals down to earth.

4. Discourse and Political Legitimacy

Habermas's discourse theory of democracy and law sets forth a normative model of political legitimacy in modern constitutional regimes (1996). Governments are stable in the long run only if citizens believe they are legitimate and thus deserve support. According to Habermas, such beliefs rest on normative intuitions that discourse theory articulates. In articulating those intuitions, Habermas hopes to reconcile two traditions in modern political philosophy: the liberal tradition, which grounds legitimate government in the protection of individual liberty and the rule of law (e.g. Locke), and the civic-republican tradition, which grounds legitimacy in participatory self-governance (e.g. Rousseau) (Habermas 1998a: chs. 9–10). If his theory succeeds, then it provides a framework that is at once normative—and thus a basis for critical social analysis—and grounded in shared intuitions about legitimacy, thus not imposed from on high.

To succeed, however, Habermas must define the sociohistorical context in which the requisite intuitions about legitimacy take effect. Constitutional political unions presuppose a collective decision "to legitimately regulate the common life of the group by means of positive law" (Habermas 1998c: 417; emphasis removed). But just what makes those regulations legitimate? In modern pluralistic societies, one cannot ground the legitimacy of laws in a common religious or metaphysical worldview. But neither can one reduce legitimacy to de facto compliance, for unjust regimes with sufficient majority support can meet that standard. Nor does legitimacy reduce to mere legality, that is, enactment according to a pre-established procedure—for what gives the procedure itself authority?

Habermas's discourse theory offers an answer: laws are legitimate insofar as they are enacted according to a procedure whose inclusive and discursive properties warrant the presumption that its outcomes are reasonable for all citizens to accept (2008: 103). A suitably deliberative democratic process, Habermas claims, meets the demands on legitimacy set forth above, for its discursive features capture normative intuitions that cut across different worldviews, and thus give democratic procedures widespread authority that is not merely a matter of majority rule. Though the latter may be required to close discussion, its contribution to legitimation depends on the quality of that discussion as an *inclusive* process of *learning* (Habermas 2009: 145–9). Habermas thus dubs his view an "epistemic proceduralism": discourse theory provides a set of ideal standards, independent of actual deliberation, which in effect define what counts as a "correct" solution (or solution set). The inclusive character of the democratic process not only contributes to its reasonableness, but also ensures that solutions are genuinely collective—an idea consonant with civic-republican intuitions.

To capture this discursive contribution to legitimacy, Habermas has formulated a broad principle of practical discourse, the "Discourse Principle" (D), which covers the justification of any social norm—moral, ethical, institutional, bureaucratic, legal: "Just those action norms are valid to which all possibly affected persons could agree as participants in rational discourses" (1996: 107). Because laws are norms that regulate actions, they fall under this principle. But because laws are created by a positive act of legislation and hold only within a polity, they have force only within the legal framework from

which they issue. Both the laws themselves and the discourses from which they emerge must have a "legal form," which presupposes a constitution and system of rights that protect individual liberty and the rule of law—the liberal intuition. Thus, legitimacy is not defined by (D) alone but rather by a *democratic principle*, which in effect specifies (D) for legal-political discourse in liberal constitutional democracies: "Only those laws may claim legitimacy that can meet with the assent of all citizens in a discursive process of legislation that has in turn been legally constituted" (1996: 110; translation modified).

The democratic principle sets a very high bar for legitimate laws: the consensus of all citizens. In fact, Habermas's idea that the democratic process warrants a presumption of reasonableness—his solution to the problem of legitimacy described above—suggests that consensus ("assent of all") is better understood as a *defeasible presumption* that an outcome is not *un*reasonable. Thus citizens who disagree with an outcome can, nonetheless, regard it as "rationally acceptable." This interpretation clarifies the role of discourse theory in critical social analysis: if the design and conduct of the actual democratic process sufficiently satisfy discursive idealizations, then citizens are warranted in taking outcomes as legitimate. Ultimately the critic's goal is to help citizens and officials hold themselves accountable to the normative presuppositions they themselves must (tacitly) accept as participants in discourse and as democratic citizens.

5. Discourse Theory and Critical Analysis

The critic's task presents a daunting analytic and empirical challenge, however. Political deliberation is doubly complex. On the one hand, the mix of the ideal types of discourse we saw above means that deliberation, understood as rational discourse, is *internally complex* (1996: 151–68). On the other hand, deliberation is *socially complex*, involving a mix of social actors and venues. The debate over health-care reform in the United States is a case in point. Internally, the debate involved moral questions (e.g. do all citizens morally deserve access? What method of paying for reform is fair?); ethical-political self-understanding (which system for delivering health care best accords with American values—single-payer, market-based, etc.?); pragmatic questions (how can the nation afford universal access? Which system is more efficient?); and legal considerations (which proposals are consistent with constitutional norms and precedents?). Socially, the debate involved ordinary citizens, insured and uninsured; insurance companies; health-care providers; pharmaceuticals; businesses large and small; it took place not only in Congress and other public forums, but also in a wide range of informal everyday settings. Presumably, a Habermasian analysis of the debate would have to examine how well the actual deliberation included these various groups and satisfied the different normative conditions of each kind of ideal discourse—in a way that recognized the constraints that moral principles place on the pursuit of more particular values and goods.

To conduct this analysis, one must understand the relationship between three broad arenas of political communication: institutionalized discourse, everyday communication, and the public sphere (see Habermas 2009; cf. 1996: ch. 8). Legal regulations come about by acts of public lawmaking, but those regulations are legitimate only if the "institutionalized discourse" of lawmakers and public officials is not steered exclusively by bureaucrats, corporate lobbyists, and other special interests, but rather responds to, and further refines, inputs from the informal and decentered public discourse in civil

society, in which citizens make their concerns known and engage in open, inclusive debate. This "everyday communication in civil society" is inevitably dispersed across a variety of forums: face-to-face interactions at home and work; newspapers, talk shows, and internet sites; voluntary associations and citizen lobby groups, and so on. Everyday political talk is linked with institutional discourse by an intermediate arena, the political public sphere comprised of networks of "media-based mass communication," in which journalists and prominent public figures (politicians, experts, intellectuals, etc.) address the citizenry. The public sphere mediates the process of public deliberation insofar as the media broadcasts opinions and arguments, works up opinion polls, and so on—and offers ordinary citizens opportunities to respond in blogs, online chat sites, and the like.

Actors in all three arenas, including corporate lobbyists and special interest groups, contribute to the information and arguments that make their way through the networks that constitute the public sphere. "Considered" or "reflected" public opinions emerge insofar as this process of communication becomes reflexive, such that citizens react to—take an "additional stance" toward—the prominent opinions put forth through the mass media. As products of a reflective stance that weighs the available arguments and information, considered public opinions represent the most reasonable alternative views on controversial matters. They thus supply lawmakers with "parameters for the range of possible decisions which the public of voters would accept as legitimate," and they orient voters in deciding between competing party platforms and candidates (2009: 166f., 162–7).

Given this complexity, the key critical question for assessing the legitimacy of an outcome is whether the democratic process has succeeded in producing "relevant and sufficiently reflexive" public opinions that then have an effect on lawmaking (2009: 162f., 167–9). To answer this question, one must (a) clarify the validity claims and modes of discourse at issue, asking whether the various arguments, perspectives, values, and interests of different groups have found adequate expression in the media-based "elite discourse" in the public sphere; (b) ask whether everyday communication in civil society has been sufficiently inclusive and reasonable, empowering citizens of all stripes to make their views known and to respond intelligently to opinions put forth in the public media; and (c) ask whether or not the media itself is sufficiently independent of forces that would undermine its critical mediating role in one way or another (2009: 172–81). One must then ask (d) whether these public opinions have penetrated institutionalized discourses without distortions engineered by power-holders of different sorts.

The role of compromise becomes especially important at this last point, when legislators, administrators, and lobbyists work out legislative proposals. Habermas acknowledges that deliberation involves not only discourse, but also bargaining processes steered by power (1996: 165–7). The latter necessarily supplement even the most reasonable political discourses, given that human beings have particular interests that moral, ethical, pragmatic, and legal reasons cannot fully reconcile (or eliminate). The critical question is whether or not compromises issued from fair bargaining conditions, such that the more powerful actors did not manipulate the outcome.

6. Mouffe's Critique of Rationalist Democratic Theory

It should come as no surprise that such an ambitious theory has encountered heavy criticism. A sticking point, even for sympathetic readers, is Habermas's ideal of rational consensus (McCarthy 1998; Warnke 1999; Benhabib 1992). Even as a counterfactual

idealization, this assumption seems to underestimate the extent to which subpolitical ethical values (the "good") shape the moral-political standards that Habermas regards as universally valid (the "right"). If Habermas cannot maintain the distinction between the right and the good, and the priority of the former over the latter, then consensual idealizations lose their grip on pluralistic democracies. We should then expect both the fairness of procedures and the justice of outcomes to vary with subgroup ethos.

Chantal Mouffe has provided one of the starkest formulations of this challenge. She rejects Habermas's rationalism as profoundly mistaken, based on dangerous illusions about the possibilities of overcoming the antagonisms endemic to political life (Mouffe 2000a, b). Although Mouffe shares Habermas's concern to defend liberal democracy, her "radical and plural" theory of democracy rests on postmodernist assumptions deeply opposed to Habermas's. Drawing on Derrida, she understands discourse not as argumentation but as a system of social meaning that is constituted by the alternatives it excludes (the "constitutive outside"); following Freud, she emphasizes the human passion for group identifications that separate "us" from "them" (Mouffe 2005; Laclau and Mouffe 1985; Torfing 1999). Although Habermas sees the need for "constitutional patriotism," a commitment to constitutional principles and institutions, he conceives this as a universalist mode of identification that enables citizens to subordinate their particular subpolitical attachments for the sake of agreement across differences (Habermas 1996: 500, 513f.; 2001b: 74). But what if political participation draws its energy precisely from our differences—from we/they oppositions that constitute the collective identities of various groups?

Although we cannot tackle the deeper oppositions here, we can examine Mouffe's more focused Wittgensteinian objection that the idea of an impartial discourse fails even in principle, and thus distorts our understanding of real democracies. Her argument assumes that (1) identification with a particular way of life rests, in the end, on passion rather than reason, and so one should not expect rational argument to adjudicate intergroup differences, and (2) no universal moral perspective exists that would allow citizens to rise above such identifications. Habermas somewhat accepts the first assumption (with his idea of ethical discourse), but he rejects the second. Having accepted both assumptions, Mouffe concludes that political decisions—which generally end discussion short of full consensus—always involve some degree of exclusion of, and violence toward, the group(s) identified with the rejected alternative. The challenge for democratic theory, then, is to show how liberal democratic institutions can foster the transformation of destructive antagonism into "agonism," an ongoing power struggle over the interpretation of democratic principles and their institutional realization that is settled without violence, according to procedures and values that support the mutual recognition of differences as legitimate (Mouffe 2000a, b).

Mouffe's depiction of political life powerfully resonates with some rather obvious phenomena in contemporary politics, where slogans, passion, and power so often overwhelm rational arguments. But she makes a stronger claim than this common fact about politics requires. On her view, rational arguments have no force in adjudicating political differences that turn on competing interpretations of liberal democratic values. To be sure, she seems to allow for rhetorical processes of persuasion (2000a: 97–8), and she recognizes the importance of coalitions between groups that participate in a "conflictual consensus" based on the "articulation" of common aims (2007). But that concession still allows her to insist that "free and unconstrained" consensus is a "conceptual impossibility," because human beings are always passionately embedded in particular ways of life that make impartial judgment impossible (2000a: 98).

There are at least two problems with Mouffe's objection. First, passionate identifications make rational consensus conceptually impossible only if one assumes that passion is necessarily opposed to reason. Although some of Habermas's process idealizations might seem to pit passion against reason, these idealizations are better understood as a critical heuristic that has us ask whether this or that *identifiable* passion in a group is a source of unfreedom that undermines discourse. Depending on the social context, the institutional history, the particular issue and goals of the discourse, passions might or might not prove problematic—indeed, they could prove constructive. Testing political outcomes for legitimacy thus calls for substantive analysis of the concrete context of discourse.

Second, Mouffe seems to assume that all participants enter discourse *already unchangeably committed* to one side or the other of the question at issue. That assumption misses the complex nature of participation in political deliberation. For many questions, we can distinguish two different kinds of participant: advocates passionately committed to one side or the other, and a critical audience of initially uncommitted spectators. The fact that the first group is incapable of impartial judgment does not mean that the second group, often the majority, is not. Thus, the critical audience would count as collectively impartial insofar as members' diverse particular identifications (which possibly inform their judgments on the issue) are not *equivalent* to the identifications that prevent advocates from considering both sides. If such impartiality is possible, then Habermas's ideals have a foothold on political reality.

That said, Mouffe's objection suggests a more realistic view of rational consensus-formation across contexts and groups. Any plausible conception of political discourse must acknowledge that political debates involve competing visions of a good society. Since these visions are, to some extent, rooted in citizens' concrete experiences of their different ways of life, effective political argumentation faces the rhetorical challenge of communicating the force or "passion" of those experiences across local contexts, to those for whom they are unfamiliar. To meet that challenge, citizens often rely on first-person narrative, graphic modes of argument (art, film), and dramatic displays. Such methods represent rhetorical modes of argumentation, or rhetorical supplements to logical arguments. If we think of political discourse this way, then the idea of rational argumentation and consensus across contexts is conceptually impossible only if one believes there is *no difference* between acceptable and unacceptable ways of fashioning such supplements. But that seems implausible: some narratives and dramatic presentations lead the audience into a deeper understanding of a position, whereas others manipulate by concealing as much as they reveal.

Acknowledgments

The author thanks Garth Hallett and Thomas McCarthy for feedback on earlier drafts of this chapter.

Related Topics

Continental Political Philosophy, French Political Thought in the Twentieth-Century, Liberalism, Republicanism, Contractualism and Political Liberalism, Authority and Legitimacy, Democracy, The Pragmatist Project in Political Philosophy, Postmodernism and Politics

References

Benhabib, S. (1992) *Situating the Self*, New York: Routledge.
Günther, K. (1993) *The Sense of Appropriateness*, J. Farrell (trans.), Albany: SUNY Press.
Habermas, J. (1971) *Knowledge and Human Interests*, J. J. Shapiro (trans.), Boston, MA: Beacon.
—— (1979) *Communication and the Evolution of Society*, T. McCarthy (trans.), Boston, MA: Beacon.
—— (1984, 1987) *The Theory of Communicative Action*, 2 vols., T. McCarthy (trans.), Boston, MA: Beacon.
—— (1990) *Moral Consciousness and Communicative Action*, C. Lenhardt and S. W. Nicholsen (trans.), Cambridge, MA: MIT Press.
—— (1993) *Justification and Application*, C. Cronin (trans.), Cambridge, MA: MIT Press.
—— (1996) *Between Facts and Norms*, W. Rehg (trans.), Cambridge, MA: MIT Press.
—— (1998a) *The Inclusion of the Other*, C. Cronin and P. De Greiff (eds.), Cambridge, MA: MIT Press.
—— (1998b) *On the Pragmatics of Communication*, M. Cooke (ed.), Cambridge, MA: MIT Press.
—— (1998c) "Reply to Symposium Participants, Benjamin N. Cardozo School of Law," W. Rehg (trans.). In M. Rosenfeld and A. Arato (eds), *Habermas on Law and Democracy*, Berkeley: University of California Press, pp. 381–452.
—— (2001a) *On the Pragmatics of Social Interaction*, B. Fultner (trans.), Cambridge, MA: MIT Press.
—— (2001b) *The Postnational Constellation*, M. Pensky (trans.), Cambridge, MA: MIT Press.
—— (2003) *Truth and Justification*, B. Fultner (trans.), Cambridge, MA: MIT Press.
—— (2008) *Between Naturalism and Religion*, C. Cronin (trans.), Cambridge: Polity.
—— (2009) "Political Communication in Media Society: Does Democracy Still Have an Epistemic Dimension? The Impact of Normative Theory on Empirical Research," *Europe: The Faltering Project*, C. Cronin (trans.), Cambridge: Polity, pp. 138–83.
Johnson, R. H. (2000) *Manifest Rationality*, Mahwah, NJ: Erlbaum.
Laclau, E. and C. Mouffe (1985) *Hegemony and Socialist Strategy*, London: Verso.
McCarthy, T. (1978) *The Critical Theory of Jürgen Habermas*, Cambridge, MA: MIT Press.
—— (1998) "Legitimacy and Diversity: Dialectical Reflections on Analytic Distinctions." In M. Rosenfeld and A. Arato (eds.), *Habermas on Law and Democracy*, Berkeley: University of California Press, pp. 115–53.
Mouffe, C. (2000a) *The Democratic Paradox*, London: Verso.
—— (2000b) "Which Ethics for Democracy?" In M. Garber, B. Hanssen, and R. L. Walkowitz (eds.) *The Turn to Ethics*, London: Routledge, pp. 85–94.
—— (2005) *On the Political*, London: Routledge.
—— (2007) "Articulated Power Relations-Markus Miessen in Conversation with Chantal Mouffe." London, 2006. Accessed July 8, 2010 from http://www.westminster.ac.uk/__data/assets/pdf_file/0004/6457/csd_mouffe_interview_with_miessen_ArticulatedPowerRelations.pdf.
Owen, D. S. (2002) *Between Reason and History*, Albany: SUNY Press.
Torfing, J. (1999) *New Theories of Discourse: Laclau, Mouffe, Žižek*, Oxford: Blackwell.
Warnke, G. (1999) *Legitimate Differences*, Berkeley: University of California Press.

Further Reading

Bohman, J. and W. Rehg (eds.) (1997) *Deliberative Democracy: Essays on Reason and Politics*, Cambridge, MA: MIT Press. (Essays that represent major perspectives on rationalist democratic theory.)
Findlayson, J. G. (2005) *Habermas: A Very Short Introduction*, Oxford: Oxford University Press. (Very accessible undergraduate-level supplement.)
Fultner, B. (ed.) (2011) *Jürgen Habermas: Key Concepts*, Durham: Acumen. (High-quality but accessible introductions to central aspects of Habermas's philosophy.)
Ingram, D. (2010) *Habermas: Introduction and Analysis*, Ithaca, NY: Cornell University Press. (A detailed, but clear and comprehensive introduction suitable for graduate students.)

Part VIII

ISSUES IN SOCIAL AND POLITICAL PHILOSOPHY

62
EDUCATION

Harry Brighouse

1. Introduction

Education has, until recently, been neglected somewhat by analytical political and moral philosophers. Yet it is impossible fully to evaluate an education system, or a proposal for reform, without making judgments about what values ought to guide educational institutions. Political and moral philosophy, then, have an essential role to play in responsible thinking about educational institutions. Three relevant categories of value can be discerned; values concerning how educational opportunities should be distributed across the population, values concerning what content those opportunities should have, and values concerning how to manage tradeoffs between the values guiding education and other important social values when they are in conflict.

This chapter will discuss distributive values; how educational opportunities should be distributed, and will only be concerned with the education of children. The discussion will be easier to follow if the reader keeps in mind a very rough account of what values should guide the content of education. It will assume a mildly perfectionist account of education's aims; that it should be directed at equipping children to flourish, to have worthwhile lives, within the social structures they will inhabit as adults. In order to equip children to flourish as adults, educators need to do four things: prepare them to be able to make their own, independent judgments about how to live their lives; prepare them with the skills and understanding they will need in order to manage themselves within the labor market conditions that they face; equip them with the knowledge, habits, and skills that they need in order to have successful personal relationships; and prepare them to be competent democratic citizens.

This piece offers no defense of the above account, nor does it offer a substantive account of flourishing. (See Brighouse (2005) for a defense of the aims of education; see also Gutmann (1999) for a landmark discussion.) The purpose of mentioning it is just to provide a placeholder for the content of the opportunities distribution which is under consideration. While *some* accounts of the aims of education might confound what is said in the following sections about the distributive principles, much will stand or fall whatever plausible accounts of educational aims are adopted.

Education is not identical with schooling. Schooling plays an important role in socializing children and in developing their cognitive skills; it appears to contribute considerably to the development of their human capital conceived both narrowly as economic productivity and, more broadly, as their ability to contribute to society and succeed in their personal lives (see Goldin and Katz 2010). But how parents treat children and the social environment in which children are raised also have considerable influence, partly

in how they affect children's interaction with schooling, and partly independently (Lareau 2011; Rothstein 2004). Children who are seriously abused or neglected are at much greater risk than others of various bad outcomes, including poor mental health, suicide, substantial periods without work, involvement in the criminal justice system, and teen pregnancy (Waldfogel 2005). But children raised in poorer homes also face somewhat higher risk of those bad outcomes than children raised in wealthier homes, and children of a single parent (including a parent who repartners) are at greater risk than those raised by both original parents (McLanahan and Sandefur 1994). Parents, grandparents, neighbors, peers, broadcast media, marketers, and other influences, as well as schools, impart education.

This complicates discussions of distributive justice in education. Most contemporary literature, while authors are cognizant of the distinction between education and schooling, implicitly concerns the distribution of schooling rather than education. (Jencks (1988), which frames much subsequent literature, explicitly asks whether schooling should compensate for background educational disadvantages.) The assumption is that how schools should treat children (given the values that should govern them) must depend to some extent on how other influences operate, both because schools supplement the education children receive elsewhere and because the school behavior interacts with that of other agents and agencies to produce whatever effects it has on the child.

A second, and related, complication concerns the relationship between ideal and non-ideal theorizing. Ideal theorizing develops principles of justice for a just society; one in which the institutions are aligned with the ideals, and most of the members of society are, for the most part, disposed to act, and capable of acting, as the principles demand. Non-ideal theorizing, by contrast, makes the assumption that the society under scrutiny is not fully just and, furthermore, will not be made fully just, and asks what principles should guide (either individual or collective) agents aiming to make the society more just. This distinction is rarely made explicit in discussions of justice in education; most proceed as if the conclusions apply in both ideal and non-ideal theorizing. But it could be that the principles that should guide the action of someone designing an education system within a fully just society would be different from those guiding reformers in an unjust society. One of the principles considered later is meritocratic educational equality, which requires that educational prospects be unaffected by social class background. But the case for meritocratic educational equality often rests on the assumption that individuals are competing for the unequally distributed goods to which education provides access; if one believed those goods should be distributed equally, or one inhabited a society in which education had little or no influence over who attained them, that case for meritocratic educational equality might not have any force.

Sections 2–5 explore four alternative proposed principles for the distribution of schooling, and the Section 6 briefly explores the possibility of combining some of them with each other, or with others. Sections 7 and 8 look at how someone committed to one or other of these principles might consider potential tradeoffs in conflicts with other important values.

2. Educational Equality

One popular view, which serves as the baseline of comparison, is educational equality—the view that in some sense of good, everyone should have an equally good edu-

cation. Even politicians and theorists who would reject egalitarianism *tout court* often invoke some sort of educational egalitarianism in justification of policy initiatives. The 2001 No Child Left Behind law in the US, for example, aims at reducing the "achievement gap," where that is understood as the gap between proportions of children from higher and lower socioeconomic classes, or of racial or of ethnic groups, reaching some threshold of competence in various basic academic tasks. Some British policymakers have insisted on a goal of completely de-linking academic performance from social class background. In both cases the concern is that socioeconomic disadvantage leads to inequality of educational achievement. In most developed countries further efforts are taken to ensure that children with disabilities in general, and learning-related disabilities in particular, have opportunities to learn in excess of those they would have if they were accorded only the same level of resources as other children.

Inequalities of both educational resource expenditures and educational achievement vary considerably even among wealthy capitalist societies, as well as over time within them. Take the United States. Children from the bottom quintile of the income distribution enter kindergarten with much lower math and reading scores than children from the top quintile of the income distribution; and these differences predict differences in high school graduation, college entrance, and college graduation (Lee and Burkham 2004). The fact that socioeconomic class gaps predict achievement and attainment gaps has been true in the US since records began; but there is now evidence associating increases in educational gaps with increases in income inequality; the gap in average achievement between children from the bottom and the top income deciles increased considerably in the 1970 to 2005 period during which income inequality increased (Reardon 2011). That the gap varies across countries and over time suggests that it is, to some degree, possible to influence its size through public policy.

But if one sought to do this, what principle of educational equality should one be guided by? (See Jencks (1988) for an exploration of the difficulties of specifying the content of the principle of educational equality, and Swift (2004) for an example of the meritocratic version of the principle.) The designation is multiply ambiguous. Which version of the principle should be adopted, if any, depends on the moral considerations that animate the more general thought. Whereas many involved in public policy discussions bemoan, for example, the fact that, as a result of the low levels of education among children from disadvantaged backgrounds, society forgoes the availability of considerable human capital, this consideration does not directly support a principle of educational *equality*: as we shall see later, a principle of optimizing human capital development is, in principle, consistent with considerable inequality of educational opportunities and of outcomes. The most common consideration evoked directly in favor of educational equality is the idea that, in an environment in which socially produced rewards are distributed highly unequally, and in which education is a key influence on who does well in competition for these rewards, it is unfair if some face a lower probability of attaining those rewards than others because they had a lower quality education. Unlike demands to equalize overall condition, then, educational egalitarianism is closely associated with meritocracy; the idea that inequalities of outcomes are justified as long as the competition for those outcomes is fair and rewards some combination of talent and effort. The principle of educational equality does the work of ensuring that, despite unequal social starting points, children have equal opportunities to develop the talents that the competitions are structured to reward. So, whereas there might be nothing wrong, according to the educational egalitarian, in having a wide wages gap, there is something wrong if

some children have much better chances of getting the jobs to which high wages are attached because they were given superior chances to develop their talents.

Unsurprisingly, then, the dominant version of the principle is meritocratic educational equality which states, consistently with the motivation set out in the previous paragraph, that:

> An individual's prospects for educational achievement should be a function only of that individual's effort and talent, not of his or her social class background.

This principle, or something like it, lies behind a good deal of contemporary rhetoric. But, as stated, it faces several challenges, of which three bear closer investigation.

The first is that it is unstable. In singling out social background as an unacceptable source of influence on outcomes the principle arbitrarily favors the talented, who merit no more credit for their natural advantages than the well-born do for their social advantages. Why should the naturally talented have special access to unequally distributed rewards?

A second challenge objects that the means that would be needed to realize the principle are unacceptable because they would undermine other values. For example, some people think that prohibiting, or imposing punitive taxes on, elite private schooling would violate parental liberty. Perfectly realizing the principle would probably require even more intrusive measures; interfering with the ordinary child-rearing practices of middle-class parents who prepare their children to take good advantage of the opportunities presented in school (such as teaching them to read at home, reading them bedtime stories, and teaching them middle-class manners).

In fact, observing a conflict between two values in particular circumstances does not establish that either principle is wrong. Even quite radical educational egalitarians (see below) tend to agree that when the principle comes into conflict with ordinary child-rearing practices that lie at the heart of family life, it should give way to the value of the family. But this does not render it inert. Those radicals will usually maintain that although parental liberty is important, it is not so important that it requires us to permit parents to purchase elite private schooling for their children. And, even if a successful argument could be given for why that was so important, the principle of educational equality might still require governments to take other measures, such as improving state schools so that they were effectively competing with elite private schools, or limiting inequality of wealth, or reducing child poverty, concentrations of which are a major barrier to providing good educational opportunities for less advantaged children.

The third objection appeals to efficiency. It is, or at least can be, socially inefficient to do what would be required to produce meritocratic educational equality, because it would result in a leveling down of educational provision and, consequently, reduced investment in the total stock of human capital and, ultimately, social wealth. At least in some circumstances this seems likely, and egalitarians are unlikely to dispute it. But social wealth is only one value; fairness in the competitions to access it, and how it is distributed, also matter. Educational egalitarianism describes a principle of fairness concerning access to the stock of social wealth, and egalitarians accept that justice will sometimes conflict with growth. Depending how much weight is placed on the principle, different judgments will be made concerning the likely tradeoffs.

A corollary of the motivation for the meritocratic principle is the idea that as inequalities of outcome narrow, educational equality becomes less important, because edu-

cation has a less important role in allocating people to advantages in the labor force. But education is not only valuable because it helps its recipients in social competitions; it is also intrinsically valuable, contributing as it does to personal growth and flourishing. Most egalitarians have a residual concern about the unfairness of some getting more of the benefits intrinsic to education than others through no effort or merit of their own. This concern has force even if non-educational outcomes are equalized.

The objection that rewarding natural talent but not social class is arbitrary suggests a much more radical principle of educational equality, one which attempts to compensate for inequality of talent, as well as for inequality of social class background:

> An individual's prospects for educational achievement should be a function only of that individual's effort, not of his or her social class background or natural talent.

This principle reflects the correct observation that natural talent is just as arbitrary from the moral point of view as social class. But it has two obvious problems. One is that, taken alone, it seems to justify concentrating massive resources on children with cognitive disabilities; the other is that it seems to justify leveling down educational achievement to the highest level that the lowest achiever reaches. Both consequences are unappealing. (See further exploration of these problems in Brighouse (2000: ch. 7).)

3. Adequacy

The main alternative regularly proposed to educational equality is the principle of educational adequacy. As with educational equality, there are several versions, all of which have the following form, but which specify X differently.

> Everyone should receive an education adequate for them to X.

At the most austere end of adequacy theorists is James Tooley, who demands education adequate to functioning in the economy; at the most demanding end are theorists like Debra Satz, Elizabeth Anderson and Amy Gutmann, who tie adequacy to the developed capacity to participate as an equal in political (Gutmann) and social (both Satz and Anderson) life. (For variants of this view see Gutmann (1999: 128–39); Curren (1995: 239–48); White (1994); Tooley (1996); Anderson (2007: 595–622); Satz (2007: 623–48).) Adoption of a principle of educational adequacy seems in most cases to be motivated by a more fundamental connection to the idea that adequacy, rather than equality or maximizing the position of the least advantaged, is what justice demands regarding the distribution of resources all-told. (See Frankfurt (1987) and Raz (1987: ch. 5) for statements of this view.) Tooley (2000) explicitly grounds his support for educational adequacy in the sufficientarian critique of a principle of equality of resources, rather than directly criticizing the case for educational equality. Anderson and Satz are both more direct in the criticism of various versions of the principle of educational equality, but Anderson (1999) at least has also endorsed a general principle of sufficiency as the core commitment of a theory of social justice, which is in turn grounded in a very extensive critique of equality as a general principle of justice. But, just as neither of the versions of educational equality surveyed above is necessarily tied to a more general principle of equality, one could endorse a principle of educational adequacy without

being a sufficientarian more generally; it is possible to endorse a principle of equality of condition but be unconcerned with specifically educational inequalities above a threshold of adequacy; and it is possible to believe that, once one is an adult, the chips should fall as they may as long as all competitors have had an adequate education.

What are the problems with adequacy principles? Imagine that everyone is adequately well educated (understanding adequacy however you plausibly might). Now imagine that new resources enter the educational system, and that whomever they are spent on, it will remain the case that everyone is adequately well educated. If adequacy were the sole distributive principle then there would no reason of justice at all to spend those resources on the least advantaged students. But this seems implausible; there is such a reason, which is that they, through no fault of their own, will have a worse education than others, and by spending the money on them society can alleviate that condition. That reason might not outweigh reasons to spend the money elsewhere; for example, if spending the money on more advantaged students would predictably secure better overall prospects for the less advantaged, which might well constitute another reason of justice (see below). But the principle of educational adequacy, offered as the sole and exclusive principle of educational justice, cannot recognize any reason of justice to spend it one way or another.

4. Benefiting the Least Advantaged

The efficiency objection to educational equality is sometimes posed specifically in terms of benefit to the least advantaged. In this version, the objection suggests an alternative principle to either version of educational equality which places a principled limit on the resources devoted to students with disabilities, and a reason not to level down achievement, namely:

> Education should be distributed in the way that maximizes the prospects for overall wellbeing of those whose prospects are poorest.

This principle has not been well explored in the literature. (I am not aware of anybody who proposes it directly except for Regina Schouten (forthcoming). See Rawls (1971), Clayton (2001), and Brighouse and Swift (2006) for arguments in favor of prioritizing the interests of the less advantaged in social policy generally.) Adopting this principle effectively abandons the idea, tacit in the previous principles, that there is a special principle of justice for education. Instead, it directly subordinates educational policy to an overarching principle of justice that demands maximizing the prospects of the least advantaged. The reasoning for the principle is the reasoning behind that principle. Given that the benefits to which education provides access are socially produced, and given that nobody has a claim that institutions be set up to ensure that they get to own some particular share, equality is the baseline principle. But an equal distribution which makes everyone worse-off than anyone would be in some available unequal distribution would be undesirable for everyone; so institutions should be designed to trigger productive behavior that benefits all, even at the cost of equality. Educational institutions are among the institutions that could play that role, indeed, that naturally do so in a highly educated and reasonably free society and economy.

Benefitting the least advantaged, though not an education-specific principle, is nevertheless a genuine and viable alternative to the principles discussed above. Unlike the

meritocratic version of educational equality it does not draw an arbitrary distinction between social and natural advantage. But unlike educational adequacy, it can explain why we should not be indifferent to inequalities above the adequacy level, or to how a sudden bounty of educational resources is distributed. And it has the nice feature, compared with the radical version of educational equality, that it does not demand leveling down of educational outcomes. But it also has the nice feature, compared with the meritocratic version, that it does not arbitrarily favor the talented; the talented get better educated, if they do, because that will benefit other less fortunate people. The radical version of educational equality would, if pursued monomaniacally, harm everyone, because severely cognitively disabled people are much more likely to have reasonable lives in a society that is productive in the way that a society can be only with a median and average level of educational achievement far in excess of what they could achieve in the best scenario; and the technological innovations made possible by high levels of education can not only benefit them materially, but educationally as well.

As with the other proferred principles, benefitting the least advantaged is problematic when offered as the sole and exclusive principle to guide the distribution of educational resources. Consider the most obvious problem; as with both versions of educational equality, realizing this principle fully might require extensive interference with family life. Even if removing children from the family home were ruled out, some other interventions, such as requiring talented students to be taught exclusively in a language that their parents do not understand, in order to ensure that they will be maximally productive, could cause substantial problems for family life.

5. Excellence

A final alternative to educational equality focuses on the value of educational achievement itself. John Wilson (1991: 29) makes the following proposal:

> Educational resources should be distributed to those who can make the most use of them.

(See also David Cooper (1980) for a trenchant statement of the principle that educational excellence is more important than concerns about the distribution of educational resources or achievement.)

The principle gets its appeal from two different sources. One is simply that educational achievement is intrinsically valuable, and should therefore be promoted. The other is some version of the principle of efficiency with regard to the production of educational excellence; to maximize excellence we would have to invest optimally, so it would be most urgent to invest in those whose capacity for achievement is high, and can be developed inexpensively. The two sources are in tension, though. The principle as stated would benefit some of the most able children enormously; but only *some* of the most able. For example, if a highly able child is sufficiently socially difficult, the cost of developing her talent might be so great that it would be more efficient to invest in a more docile, but less able, child; children who speak an unusual foreign language might be expensive to invest in, even though highly able. So whereas investing *enough* in the difficult or foreign language-speaking child might produce the highest levels of individual achievement by a single child, the efficiency thought counts against that. As stated, and taken as the sole criterion of distribution of educational opportunities,

though, the biggest problem with the criterion is that it is insensitive to the connection between educational achievement and the distribution of other rewards, both intrinsic and extrinsic. Benefitting the least advantaged may be a superior interpretation of the efficiency thought that underlies the principle. However, the thought that excellence is itself intrinsically valuable remains; and while it seems implausible that it should have so much weight that it should completely command the distribution of educational resources (which could, for example, have the consequence that children should be removed from their parents and placed in academic "hothouses"), it is plausible that it should play some role in determining distribution.

6. Hybrid Views

The theme so far has been that each of the proffered principles is problematic, when taken to be the sole guide to the distribution of educational resources. But several of the principles have some plausibility. One might, then, take the view that each of the starkly stated principles above captures a rational kernel, which has a place and needs to be weighed against the rational kernel in the other views. There is a reason to reduce the effect of social class, and of natural talent, on outcomes, but there are also reasons to seek higher levels of excellence, among them the reason provided by the desirability of producing higher levels of human capital that can be harnessed to the overall benefit of the least advantaged; and there is a reason to seek an adequate education for all. Or one might take the view that there is a rational kernel to *some*, but not all, of the principles. The task then, of constructing a theory of distributive justice for education is the task of identifying the reasons for adopting each principle which has some plausibility and showing how much weight they should have relative to one another.

For example, the excellence criterion described in the previous section looks implausible. But it also looks implausible that a society should be entirely unconcerned with promoting, as opposed to merely distributing, achievement. Adequacy principles concern both the promotion of achievement and its distribution. But one might think that they miss something out: any reference to individuals having a chance to realize their unique potentials, which is a thought that animates the excellence principle. Perhaps children have a right to an educational experience that enables them to realize their potential to some extent independently of whatever social use their skills can be put to. They may, then, become skilled at something (such as philosophy) which leads to less productivity than some other capability that, although realized to a lesser degree, is more socially valuable (such as diagnosing diseases). A principle demanding that resources be distributed so that the gap between academic potential and academic achievement be equal for each child, subject to the requirement that everyone has an education adequate to be an equal participant in social, political, and economic life (understood, for example, the way that Anderson (2008) understands that) would, then, be an example of a hybrid.

7. Balancing Values and Tradeoffs

Returning to the thought that just one of the above principles is the right guide to the answer "what is educational justice?," suppose something like the meritocratic principle is the right principle of justice for distribution in education. As I have observed, implementing it fully would require quite demanding public policy measures, because as things

stand a child's family background significantly determines her opportunities in life. Wealthier parents can improve their children's competitive prospects by paying for more expensive schooling, by creating a different peer group and ensuring they have relationships with other wealthier adults and, later, by subsidizing unpaid internships that enable the child to learn more and have a more impressive CV than those they are competing with in the labor market. Educated parents can help children by preparing them better to take up the opportunities presented by schools, and by supporting them better in taking up those opportunities. A teen who becomes addicted to alcohol or heroin has an advantage if her parents can and will pay for high-quality treatment. Inequalities due to such accidents of birth seem like paradigm cases of unfairness, hard to reconcile with any plausible conception of social justice. John Rawls famously expresses (1971: 265) one of the central conflicts with a plausible conception of equality of opportunity:

> [T]he internal life and culture of the family influence, perhaps as much as anything else, a child's motivation and his capacity to gain from education [so that] … even when fair opportunity (as it has been defined) is satisfied, the family will lead to unequal chances between individuals. Is the family to be abolished then? Taken by itself and given a certain primacy, the idea of equal opportunity inclines in this direction.

To say this is not to show that the meritocratic principle is false. It might be true—that is, it might express a genuinely valuable value—but so might other principles that, in certain circumstances, conflict with it. Consider one measure that might mitigate the inequalities caused through family life: removing all children from their parents and placing them in boarding schools for 12 hours a day, 6 days a week, 50 weeks of the year, from the age of 24 months. Would it be morally acceptable to do so? Most readers will be inclined to demur; parents have a strong interest in being able to have ample time to spend with their children, in being the people primarily in charge of their wellbeing, and the time left over by the described policy would be insufficient for them to realize that interest, even if it did no harm to the children. Allowing parents to establish and maintain intimate relationships with their children is an important value, almost certainly more important than educational equality (see Brighouse and Swift 2006). But the fact that it is more important does not mean that educational equality is unimportant, nor does it even mean that nothing may be done against the wishes of parents. The practice of making schooling compulsory and free at the point of delivery for all who seek it, which is common in contemporary wealthy countries, might prevent some parents from spending as much time as they would like with their children, and certainly affects the personal and social growth of their children, and in some cases this might make family life a tiny bit worse; but it (we presume) has big effects on the degree of educational inequality in the society, and that might justify it. Or consider a program within public schools which have a heterogeneous mix of children from different social classes requiring that they all come for a one-hour breakfast period at the start of the school day in order to ensure that low-income children start the day with a breakfast but are not stigmatized; if there were evidence that such a measure had substantial effects on the achievement of low achievers, it might be justified even though it deprives wealthier parents of some access to their children.

Many measures promoting educational equality would not be deleterious to family life. The family life of upper middle-class Britons is not worsened relative to that of

upper middle-class Americans by the practice of funding schools proportionally to disadvantage rather than inversely proportionally to disadvantage. Creating a targeted voucher system in which the value of the voucher is sensitive to educational need might reduce the achievement gap, but it would do so without cost to the family life of those who are not eligible for the voucher. And further measures might benefit the family life of both the disadvantaged and their more advantaged peers. Suppose that measures were found that substantially improved school productivity in such a way that not only drove up the achievement of low achievers but made the fraction of them who are very disruptive much more engaged in school. This might improve their family lives by reducing their parents' levels of stress and conflict in the home, but might also improve the family lives of higher achievers by reducing *their* parents' stress levels.

Children are not just adults-in-formation, and education should not be solely guided by their future prospects; childhood is itself a significant part of a person's life, and the quality of a childhood is intrinsically important, independently of its consequences for the quality of an adulthood. Recent work in the sociology of childhood has emphasized this, and it is implicit, too, in practitioners' concerns about matters such as bullying and testing. It is worth remembering, too, that schooling is compulsory. Children have no choice but to spend a significant number of their waking hours in the classrooms where their parents and teachers have placed them, and among other children whom they have not chosen, and many of whom they would not choose, as their companions. This places a special burden of justification on us when it comes to the quality of the time the children spend there, even though there are good enough reasons to require them to be in school. Philosophers, while none of them denies the independent importance of childhood as a stage of our lives, have not done much work figuring out how it constrains the delivery of education, or even what constitutes a good childhood other than the goods in childhood that prepare us well for adulthood. It is easy to imagine, though, that some of the means that would most improve some child's prospects for academic achievement might diminish the quality of her school days. It is plausible, for example, that in some circumstances frequent and rigorous testing would be involved in the most effective method for improving low-end achievement, but would make some low-end achievers excessively anxious at the time. Even if we were confident that a rigorous testing regime was crucial for the best strategy for improving a child's performance and thus reducing educational inequality, we might feel justified in sacrificing some of that achievement for the sake of not making their school days miserable.

Whereas children from higher socioeconomic classes might expect that their parents will celebrate and understand their educational success, and that it will lead them to have lives recognizably like those of their parents, educational success for many working-class children would exact the high price of alienating them culturally from their parents, siblings, and communities. If commonalities with one's family members are needed for relationships to remain close, and close familial relationships play an important role in underpinning a flourishing life, the social mobility generated by meritocratic educational equality might conflict with the flourishing of the socially mobile child.

8. Concluding Comment

Two tacit assumptions tend to underlie discussions of distributive justice in education; that there is a single guiding principle, and that principle is the same in both ideal and non-ideal contexts. But every proposed principle, once specified, seems problematic if

it is treated as the only guiding principle; an approach that assigns weight to various of the principles, and also balances them against other important values, might be better. And some non-ideal contexts are sufficiently far from the ideal that the various principles might have different weight in non-ideal than in ideal circumstances. Furthermore, none of the proffered principles is specified with sufficient detail to provide much guidance to decision-makers in some important non-ideal decisions; future work should rectify this.

Related Topics

Conservatism, Perfectionism, Pluralism, Luck Egalitarianism, The Difference Principle, Desert, Needs and Distributive Justice, The Capability Approach (and Social Justice), Intergenerational Distributive Justice, Equality, Freedom, Autonomy, Rights, Health, Marriage, Sex, and the Family

References

Anderson, E. (1999) "What is the Point of Equality?" *Ethics* 109: 287–337.

—— (2007) "Fair Opportunity in Education: A Democratic Equality Perspective," *Ethics* 117: 595–622.

Brighouse, H. (2000) *School Choice and Social Justice*, Oxford: Oxford University Press.

—— (2005) *On Education*, London and New York: Routledge.

—— and Swift, A. (2006a) "Equality, Priority and Positional Goods," *Ethics* 116: 471–497.

—— (2006b) "Parents Rights and the Value of the Family," *Ethics* 117: 80–108.

Clayton, M. (2001) "Rawls and Natural Aristocracy," *Croatian Journal of Philosophy* 1: 239–259.

Cooper, D. (1980) *Illusions of Equality*, London: Routledge and Kegan Paul.

Curren, R. (1995) "Justice and the Threshold of Educational Equality," *Philosophy of Education* 50: 239–248.

Frankfurt, H. (1987) "Equality as a Moral Ideal," *Ethics* 98: 21–43.

Goldin, C. and L. Katz (2010) *The Race Between Education and Technology*, Cambridge, MA: Harvard University Press.

Gutmann, A. (1999) *Democratic Education*, Princeton, NJ: Princeton University Press.

Jencks, C. (1988) "Whom Must We Treat Equally for Educational Opportunity to be Equal?," *Ethics* 98: 518–533.

Lareau, A. (2011) *Unequal Childhoods*, Berkeley: University of California Press, 2nd edn.

Lee, V. and D. Burkham (2004) *Inequality at the Starting Gate*, Washington, DC: Economic Policy Institute.

McLanahan, S. and G. Sandefur (1994) *Growing Up With a Single Parent*, Cambridge, MA: Harvard University Press.

Rawls, J. (1971) *A Theory of Justice*, Cambridge, MA: Harvard University Press.

Raz, J. (1987) *The Morality of Freedom*, Oxford: Oxford University Press.

Reardon, S. F. (2011) "The Widening Socioeconomic Status Achievement Gap: New Evidence and Possible Explanations," in Richard Murnane and Greg Duncan (eds.), *Social Inequality and Economic Disadvantage*, New York: Russell Sage Foundation.

Rothstein, R. (2004) *Class and Schools*, Washington, DC and New York: Economic Policy Institute and Teacher's College Press.

Satz, D. (2007) "Equality, Adequacy, and Education for Citizenship," *Ethics* 117: 623–648.

Schouten, R. (forthcoming) "Fair Educational Opportunity and the Distribution of Natural Ability: Toward a Prioritarian Principle of Educational Justice," *Journal of Philosophy of Education*.

Swift, A. (2004) *How Not to be a Hypocrite*, London: Routledge.

Tooley, J. (1996) *Disestablishing the School*, London: Institute of Economic Affairs.

—— (2000) *Reclaiming Education*, London: Continuum.

Waldfogel, J. (2005) *What Children Need*, Cambridge, MA: Harvard University Press.

White, J. (1994) "The Dishwasher's Child: Education and the End of Egalitarianism," *Journal of Philosophy of Education* 28: 173–182.

Wilson, J. (1991) "Does Equality (of Opportunity) Make Sense in Education?," *Journal of Philosophy of Education* 25: 27–31.

63
HEALTH

Norman Daniels

1. Overview

Health and its distribution are the focus of important issues for political philosophy. We begin with the basic question, "What is health?" (Section 2). Although some definitions tend to conflate health with all of well-being, one narrower definition that we use in what follows, in part because it fits well with actual work in epidemiology, public health, and health policy, remains controversial. One point of controversy is that some believe the narrow conception excludes a broad view of the answer to our second question, "What are the sources of population health and its distribution?" (Section 3). This belief is mistaken, for the narrower definition of health can be combined with a broad view of the causes of health, an important point because it turns out that social justice is actually good for our health. This fact about justice and health requires us to consider more carefully our third question, "What do we owe each other by way of protecting and promoting population health and distributing it fairly?" (Section 4). Though different accounts of justice answer this question in different ways, we concentrate on the implications of one liberal egalitarian view that extends Rawls's account of justice as fairness. Though this theory gives us some guidance in thinking about what we owe each other, it leaves considerable room for reasonable disagreement about how to incorporate in health policy two key objectives: improving population health and distributing that health fairly. This reasonable disagreement raises our fourth question, "How can we meet health needs fairly when we cannot meet them all?" (Section 5). After canvassing some options, we consider how one proposal that draws on democratic deliberative theory addresses, and we conclude by providing an answer to, a fifth question, "How should we understand a right to health or health care?" (Section 6).

2. What is Health?

According to the World Health Organization (WHO), health "is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity" (WHO 1946). Although many conceptions of health agree that health contributes to well-being and may be a constitutive component of well-being, health is not all there is to well-being. Someone could be completely healthy but lack complete social well-being for various reasons that do not bear on health. Further evidence against the WHO definition is that important areas of health measurement and health policy do not try to measure overall well-being in this way: epidemiologists, for example, consider the sources of departures from full health due to disease, injury, environmental

factors and genetic sources, but they have no way of measuring, and never are concerned with, shortfalls from the complete state of well-being that is encompassed in the WHO definition. Summary measures of population health also focus on departures from full health, but not complete well-being. Indeed, judging by these areas of science and policy, health is primarily viewed as the absence of disease and various kinds of injury. Since it is unlikely that they are all mistaken about what they are studying, the WHO definition seems out of line with the topic. Obviously, a concept of health should include both mental and physical health and the kinds of normal functioning that support the social interaction that typically contributes to mental and physical health, but the WHO definition is even broader than that.

Some people are attracted to the WHO definition because it seems to invite us to look at broad causes of health, not simply at disease agents. Nevertheless, we can adopt a narrower concept of health that more closely fits activities of the biomedical sciences and health policy and that nevertheless allows for a broad view of the causes of health and ill health (Field and Gold 1998; WHO 2002; IOM 2010). We combine a narrow definition of health with a broad view of the determinants of health in Section 3.

Others might believe that the WHO definition invites us to think about the social construction of health and therefore is incompatible with more "naturalistic" definitions of health. For them, complete social well-being seems to escape characterization as something we can specify without appeal to social attitudes and values. Although the debate about naturalism vs. social construction with regard to disease and health is ongoing, this is primarily a controversy about the nature of the concept of health and does not imply a specific definition of it (although some views about the nature of the concept might count against some specific definitions). In what follows, we avoid this debate in the philosophy of biology (where it focuses on the nature of biological functions; see Boorse (1997) and Kingma (2010)). We adopt a view of health as normal functioning that is neutral between stochastic and evolutionary (etiological) accounts of biological function and that admits a weak form of normativity (shortfalls from normal function must be harmful to the individual). Specifically, we follow Boorse's (1997) suggestion to consider Wakefield's (1992) account as a plausible alternative to his own. This eclectic view remains neutral between etiological and stochastic accounts of biological function, yet gives us a robust notion of health that still excludes extremely normative views (e.g. that view a disease as an unwanted condition).

The main point behind this eclecticism is that we should be able to adopt a view useful in political philosophy without resolving specific problems in the philosophy of biology. The reason for wanting to exclude strongly normative views that may admit of highly relativistic disagreements is that we would prefer questions about whether someone is in ill health to be more objectively decidable, even if the importance of their deficit in health for purposes of priority setting is a value judgment reasonable people might well disagree about. Political philosophy should focus on the latter question and should be able to take the former as a matter for the biomedical sciences, broadly construed. This is not to say that social attitudes do not sometimes skew conclusions about whether a problem is a health problem. Viewing "drapetomania" as the "running away disease of slaves," or masturbation and homosexuality as requiring a cure, or political opposition as a mental health problem (as in former Soviet Union) clearly is a mistaken view of disease and health, but it no more makes disease in general a social construction than calling whales "fishes" means biological classification as mammals or fish is just "social construction" (Daniels 2008).

Nevertheless, some advocates for people with disabilities find the notion of "normal functioning" demeaning, even if it is not intended to be (Silvers 2003). Some disabilities are the result of illness (severe depression), some the result of injury (quadriplegia from severing the spinal cord), some the result of environmental exposures (severe effects of lead exposure), and some the result of genetic or developmental causes, and people with some of these conditions might view themselves as "healthy" despite their deficit in functioning. For some, one mode of functioning (walking to achieve mobility) may be (partially) replaceable by another mode (a wheelchair). Yet, all disabilities involve the loss of some form of normal functioning, and that is the basis for saying there is a deficit in health, even if the person with the condition is otherwise healthy (Daniels *et al.* 2009). Some conditions can be compensated for far more than others by various forms of treatment or other inclusive social practices (modifying access to public pathways or buildings or adjusting work conditions), so the degree to which the deficit in biological or mental functioning interferes with diverse other activities can vary. The account adopted here is compatible with the most recent WHO (2001) view of disability.

3. What are the Sources of Population Health?

The main objective of this essay is to answer the questions of political philosophy raised in Sections 4–6 about what we owe each other by way of protecting and promoting health, how we can meet health needs fairly when we cannot meet them all, and what these answers mean for the claim that we have a right to health or health care. The answers to these questions, especially the first, would be much easier if the primary determinant of population health and its distribution was health care, broadly understood to include personal medical services and traditional public health measures. If health care were the primary determinant of population health and its distribution, then, to address the first of these questions, equitable access to health care would be what we owe each other by way of protecting and promoting health (Daniels 2008). But the situation is far more complex, despite the widely held myth about the effects of medicine's "magic bullets."

The myth, or more accurately semi-myth, since modern medicine has an important impact on population health, is that modern medicine has been primarily responsible for the longer, healthier lives we lead today compared to earlier centuries. After all, we read or hear every day about new medications, new diagnostic procedures, new surgical techniques—all marking great advances in medical practice visible within our lifetime, even as we take traditional public health measures for granted. Partly as a result, few people who become ill, or who have loved ones who are ill, ask for anything less than the best care possible, regardless of cost. Their clinicians also generally want to do the best they can for their patients, spurred perhaps by career and monetary incentives as well as ethical obligations.

In response to this demand, all developed societies spend considerable resources on meeting health care needs, with European countries expending 8–11 percent of gross domestic product (GDP) on health care, mostly on personal medical services. The U.S. is an extravagant outlier, spending over 17.6 percent of GDP on health care in 2009, 50 percent more than the next highest spending country, even while it is the only wealthy country falling well short of providing universal coverage (OECD 2009). Nevertheless, despite its very high expenditures, the U.S. has higher infant mortality rates and lower life expectancy than other wealthy countries. The poorer performance in the U.S.

is probably partly a result of considerable inefficiency in the fragmented U.S. health system, but, more importantly, these general indicators of population health are influenced by other factors than expenditure on personal medical services.

The belief that modern medicine and public health are the main forces underlying our longer, healthier lives, as compared to earlier centuries, was the subject of sharp criticism by Thomas McKeown (1962, 1976) in the middle of the twentieth century. He argued forcefully that major changes in transportation and nutrition accounted for a significant drop in mortality rates in Europe, before there was any impact of modern medicine or modern public health measures. Even if he underestimated the impact of public health measures (Szreter 2002) and, in the late twentieth century, the impact of modern medicine, his deflationary view has considerable force.

The most important line of evidence suggesting that we should look beyond equitable access to medical services, especially, for the sources of the distribution of population health derives from work in the last few decades of the twentieth century on social epidemiology. Most important are the Whitehall studies of the British civil servants (named after the headquarters of the British civil service) (Marmot *et al.* 1978; Marmot 1991). All subjects of these studies had access to a universal coverage system of good quality (the NHS); all had high basic levels of education; and none of them was poor. For these people, the higher their occupational status, the longer and healthier their lives. These robust findings about a socio-economic status (SES) gradient of health have been replicated in many countries and across many measures of health outcome. Higher SES correlates with better health, whether we measure status by income or wealth, education, or occupational status, and whether we measure better health by decreases in general mortality rates, some disease-specific mortality rates, and maternal and infant mortality rates, or by positive measures of population health, such as life expectancy and healthy life years lived.

The SES gradient of health indicates important health inequalities in a population that are a matter of policy concern. Indeed, the more one believes that the distribution of income, wealth, education, opportunity, and political power are unjust, then the more the resulting health inequalities that result from that distribution qualify as unjust (Kawachi and Kennedy 2002). These inequalities suggest that obligations to distribute health equitably, which many believe we have, are not fully met.

SES interacts with other social factors, such as race. In the U.S., for example, health disparities between blacks (and some other minority groups) and whites are large. Even if we control for access to insurance and thus to coverage for health care services, we find significant residual inequalities in various measures of health outcomes by race. Though SES accounts for five-sixths of the total health inequality between blacks and whites, this difference is itself the likely result of a legacy of exclusionary practices, since racism pushes disproportionately many minorities into lower SES groups. The residual black/white difference across all SES groups suggests that other factors focused on race make the gradient of health even worse for some minorities (Smedley *et al.* 2003).

Aside from the distribution of income and wealth, other important social determinants of health include the distribution of education, the support for effective means of political participation, provision of a robust social safety net, and the importance of social and cultural practices such as discrimination by race, gender, and ethnicity. In short, broader issues of social justice seem to promote population health and its more equitable distribution. These factors are increasingly being recognized globally, as evidenced by the WHO Commission on the Social Determinants of Health (2008).

To summarize so far: if population health and its distribution were simply a matter of access to appropriate personal medical services and traditional public health measures, then, arguably, equitable access to these would be the core of what we owe each other as a matter of justice. Population health and its distribution, however, are the result of a broader set of factors, even if health care is also important. Broader views about social justice, then, have relevance to what we owe each other with regard to health. For example, even if it turns out people in wealthy societies owe each other universal coverage to a comprehensive set of medical services as well as the provision of public health measures that protect all against various health risks, they also owe each other a just distribution of the full range of socially controllable factors that determine population health and its distribution.

4. What Do We Owe Each Other by Way of Protecting and Promoting Health?

Nearly everyone, political theorists included, believes that society should provide some ways of protecting and promoting health. Despite their objections to state efforts to improve welfare, even strongly libertarian views can support state protection of people against unconsented-to risks to health imposed by third parties. On such grounds, there can be support for "police protections" of health against some forms of infectious disease, environmental pollution, and behaviors dangerous to others, such as drunk driving. Most views, however, argue for much more robust forms of protection and promotion of health, even if they disagree about the grounds for them and their scope.

One key line of justification for a more robust set of obligations is based on the widely accepted view that improving or protecting people's health contributes to their welfare or well-being either directly or indirectly (Broome 2002; Broome argues further that health is not a separable component of well-being). That impacts on health affect well-being is central to utilitarian justifications for the belief that we have obligations to protect and promote health. The recent "macroeconomic" argument for investment in health, which has influenced development policies for many poorer countries, rests on the fact that investment in health leads to better economic growth and thus, more broadly, to improved well-being (Sachs 2001). Similar views were prominent in promoting the emergence of universal coverage systems in many countries. Utilitarian claims about what policies we should adopt go beyond simply protecting health against unconsented-to risks, and even beyond the provision of public goods that markets are unable to provide (clean water or air or basic biomedical research). We should promote health to the extent it promotes welfare.

Utilitarians and many other consequentialists, however, are interested in all the welfare consequences produced by improving population health, not just in the impact on health. Taking that interest seriously means that improving the health of working people, who will then produce more, has greater impact on aggregate welfare than improving the health of retirees; similarly, favoring the health of better paid workers over lower paid workers also contributes more to aggregate welfare. In the U.S. this would mean giving more priority to men than to women, who earn under 80 percent of what men earn, and more to whites than blacks, who earn even less than women on average. Contrary to widely held beliefs about fairness, utilitarianism thus understood favors those who are better off more than those who are worse off, where improving their health more produces greater total welfare.

More nuanced utilitarian views can counter this implication in various ways: such a policy would induce insecurity in retirees (a welfare cost); many of us have concern about our retired parents, and not favoring them has a cost in welfare. Nevertheless, since utilitarianism emphasizes the efficient conversion of resources, such as health care, into welfare, the theory presumably gives priority to those who will benefit more (or who will benefit others more) from our interventions. But favoring "best outcomes" in this way ignores the belief, which many share, that all people should have a fair chance at some significant benefit. In these and other ways, the theory seems to emphasize the efficient promotion of aggregate population health while ignoring common views about equity in the distribution of health.

These shortcomings are not decisive objections to utilitarian ways of justifying obligations to promote and protect health and to distribute that health equitably. They do suggest, however, that the emphasis on maximizing population health in a utilitarian framework might come at the expense of giving plausible support for other, widely held concerns for equity in health. Are there approaches in political philosophy that give support to both goals?

One approach that supports both goals is based on the following observation: Keeping people functioning normally, say by using health care and other resources to reduce the risks of ill health or to treat illness or injury when it occurs, makes a significant albeit limited contribution to protecting an individual's share of the opportunities available to her in her society, given her talents and skills (Daniels 2008). This relationship between keeping people functioning normally—keeping them healthy—and protecting their opportunities is important to several theories that emphasize the importance of opportunity as a matter of social justice. For example, if a theory supports a principle of justice that assures the protection of opportunity, such as Rawls's principle assuring fair equality of opportunity, then it provides a basis in justice for keeping people functioning normally.

To focus the discussion, consider how Rawls's principle of fair equality of opportunity can support both improving population heath and distributing health equitably (Rawls 1971). What the principle claims is that we must not only judge people by their talents and skills for access to jobs and offices, but also we must establish institutions, such as free public education, that give people a fair chance to develop their talents and skills despite other, allowable inequalities. Not only does justice require that jobs and offices be open to all, he argues, on the basis of the talents and skills they have, but also justice requires that we establish social institutions, such as free public schools (and arguably other early childhood interventions), that can help to reduce the inequalities in the development of talents and skills that result from allowable social inequalities and the impact of families on childrearing. The principle (suitably broadened to include functioning that goes beyond the Rawlsian focus on access to jobs and offices) aims us at keeping all people functioning normally—and thus implies improvement in aggregate health. But it also affirms that we are concerned with preserving equity in the distribution of opportunity—by aiming for (among other things) equity in health.

Other accounts of justice that are alternatives to Rawls (and critical of his view) also support the protection of opportunity. One such view holds that people should be compensated for any disadvantage or loss of welfare that results from bad "brute luck" as opposed to bad choices ("option luck") (Dworkin 1981; Arneson 1988; Cohen 1989), a view that has been dubbed "luck egalitarianism" (Anderson 1999). It shares with Rawls's view an emphasis on protecting opportunity, even if it gives more prominence to the

role of individual responsibility for health and other outcomes than Rawls's theory. Such a theory gives prominence to the protection of population health and its equitable distribution, though it may offer less protection for ill health that individuals are (partly) responsible for producing. Sen's (1980, 1992) (non-welfarist) view that the target of concerns about equality is the space occupied by capabilities arguably shares with Rawls a focus on opportunity. Capabilities are opportunities that people can choose to pursue. Neither the luck egalitarian nor the capabilities approach is a full account of justice, and so neither is a theory of justice for health. Neither reconciles its support for opportunity with concerns about liberty or efficiency, which arguably a comprehensive theory of justice must do. Still, both of these views support improving population health because of its impact on opportunity, and both support a more equitable distribution of health.

Despite the shared focus on opportunity, some points of disagreement among these views will lead to differences in the way they support the protection and promotion of normal functioning. Luck egalitarian (or prioritarian) views arguably place an account of what individuals have chosen to do and the effects of that on their health epistemically prior to an account of what we actually owe each other with regard to health. In short, we cannot say what we owe each other without knowing what the consequences of our choices about lifestyle are for health, whereas a Rawlsian account gives epistemic priority to claims about what we owe each other (Daniels 2011). Further, the luck egalitarian view might not justify providing incentives for our behaving more prudently with regard to lifestyle choices, because we simply owe less if people make imprudent (but informed and voluntary) choices. A Rawlsian account can provide a basis for justifying such incentives and so is more supportive of health promotion. Our main point, however, focuses on the shared view about protecting opportunity, not these differences.

What are the implications of support for both goals? First, we must reduce the risks of ill health and distribute those risks more equitably. This means arranging all the socially controllable factors—personal medical services, traditional public health measures, and the social determinants of health—in ways that carry out both tasks of risk reduction. Second, since we cannot prevent all losses of normal functioning, we must provide to everyone access to a reasonable array of services that address disease, injury, and disability. Determining what is in that reasonable array of services is another issue facing political philosophy, for reasonable disagreement about such choices is pervasive, and that raises further issues of how to meet health needs fairly.

5. How Can We Meet Health Needs Fairly When We Cannot Meet Them All?

Accounts of justice in political philosophy that give prominence to assuring opportunity can provide foundations for social obligations to protect and promote population health and its equitable distribution. Still, we cannot meet all health needs since the resources available to meeting health needs are always limited by other considerations. However important health or health care is, meeting health care needs competes with other important goods—the provision of education, employment, and adequate living conditions for all people. Although some health care resources are naturally scarce, such as transplantable organs, all health care resources are scarce because they must compete with other important goods.

Many decisions about what health care services to provide out of shared resources (for example, through taxes or through private or public insurance) are the subject of

reasonable disagreement. When a public or private insurer denies a not-fully-proven therapy that a patient and her clinician believe is a last chance at a life-saving intervention, they often trigger a loud public outcry. Value conflicts between stewardship, or the conservation of scarce resources, and compassion for those in need underlie the decision and the outcry. Reasonable people will disagree about the weights to be assigned each.

Similar reasonable disagreement is pervasive in other decisions about health policy, including coverage decisions within insurance schemes. Several "unsolved" rationing problems that involve such disagreement have been the subject of some discussion (Daniels 1993). The "priorities problem" asks how much priority we should give to those who are worse off with regard to current health, for example, those who have the worst prognosis if not treated. Two principled approaches would be to give them no priority or to give them maximal priority. On the former, we view it as equally important to deliver a unit of health benefit to a person regardless of where in a person's life it goes or how bad a person's health state. Thus cost-effectiveness analysis (CEA) uses a measure of health benefit, a quality adjusted life year (QALY), and claims it is a virtue of its approach that a QALY is a QALY wherever it goes. On the latter view, however, giving a unit of health to someone who is worse off is morally more important than giving it to someone better off, and the claim is that we should give absolute priority to those who are worst off. This approach risks creating a bottomless pit in which we devote all our resources to those worst off without regard to what other benefit we might produce by putting the resources elsewhere. Neither "principled" view matches the considered views of most people about equity: people want to give some priority to those who are worse off, but not absolute priority. Yet all intermediary points are ones where people reasonably disagree about the trade-offs they will allow.

Reasonable disagreement arises in the aggregation of health benefits: when are modest benefits to many allowed to outweigh significant benefits for a few? CEA says all aggregations should be allowed; some philosophers reject any aggregations. Both views fail to match more widely held views that some aggregations should be allowed, but others not. Similarly, the "best outcomes vs. fair chances" problem is also a focus of reasonable disagreement. CEA always favors best outcomes. Some philosophers argue for always giving people equal chances. Most people prefer intermediary solutions, but they disagree about trade-offs. Another problem on which there is reasonable disagreement is the problem of identified vs. statistical victims: should we give any moral weight to the disposition most people have to give some priority to identified victims? Again, there is disagreement about how much priority to give, if any.

CEA, as noted, embodies a controversial approach to each of these distributive problems, but it is one of the key economic tools for priority setting in health policy. Its failure to match public views, however, is not a decisive objection to it, for those views might be wrong and CEA correct. Nevertheless, this point is connected to a further problem: how much weight, if any, should we give to public attitudes toward these and other issues? If we know what public preferences—or even values—are, should we treat those as proxies for public decisions, or should they be but an input into a deliberative process addressing those problems? And what is the nature of such a process?

A problem in democratic theory underlies this cross-cutting issue: the fact that a value is widely held in some group does not mean that value should be one that is widely held and should control public policy. Democratic might—that is, decision by a majority—does not guarantee the right outcomes. Moral anthropology is descriptive; ethics is normative for all people, unless we slip into some form of moral relativism. If a

public value involves significant discrimination against some group, by race or gender, for example, we do not and should not accept that value as telling us what is the right thing to do. Racist societies do not make racism just even if their values—or their votes for it—support racist practices.

How can we assure both legitimacy and fairness for decisions, given this and the other unsolved rationing problems? One approach is to refine some of our priority setting tools, such as CEA, so that it incorporates "equity weights" and no longer is just a tool for maximizing health gain per dollar spent. Though some efforts have been made to construct such weights for single dimensions of the problem (such as age or severity of a condition), we so far lack a methodology for capturing the many interactions among dimensions of the problem (Haninger 2006). In addition, this approach assumes that if we get our view of social values right, we can treat equity weights as a proxy for democratic process, finessing the democracy problem.

An alternative approach might be to employ many philosophers to work on the unsolved rationing problems. When they solve them to our satisfaction, we can rely on those answers to set priorities. Unfortunately, we need to make decisions in real time.

We might then rely on various kinds of process to assure at least legitimacy. For example, we might hope that a market solution would emerge: people will buy different forms of insurance, embodying different resource allocations, and thus authorize those allocations by their purchase, construed as a form of consent. But this approach presupposes a range of options in a truly ideal market in which people get what they choose. We do not have such a market and cannot construct it.

A more plausible alternative is to appeal to a form of procedural justice (Daniels and Sabin 2002, 2008). Perhaps we can agree on what counts as a fair process, and we can agree to accept the outcome of that process as fair, even if we cannot (beforehand) agree what outcomes are fair. One proposal for such a process is that institutions be established at different decision-making levels in health systems and that the procedures for decision-making meet these conditions:

- Publicity: rationales for decisions are made public;
- Relevance: decision-makers, ideally including a broad range of relevant stakeholders, make decisions on the basis of reasons all consider relevant;
- Revisability: decisions are revised in the light of new evidence and arguments;
- Enforcement: the above conditions are met.

Meeting these conditions makes decisions accountable for their reasonableness. Why think adhering to this process yields greater fairness, even if we concede there is enhanced legitimacy?

Following Rawls's discussion, we can distinguish two main forms of procedural justice. In pure procedural justice, we lack prior agreement on a relevant principle for determining just outcomes, and we accept the outcome of a fair process as fair. Rawls offers gambling as an example: we accept the outcome of a fair spin of the roulette wheel as fair. In contrast, criminal trials constitute an example of impure procedural justice, since we have prior agreement on a relevant principle: convict all and only the guilty. We determine who they are through trials that pit adversaries against each other but are judged by neutral parties. If we later find conclusive evidence that someone we found guilty in a trial is innocent (say, through DNA evidence), then we should overturn the trial result.

Since we lack prior agreement on distributive principles specific enough to yield outcomes to decisions about allocating health care resources, the proposed process has some resemblance to pure procedural justice. On this view we have no basis for denying fairness to the outcome of a fair process. But the situation differs in two important ways from Rawls's example of gambling. Unlike gambling we should reject outcomes that violate requirements of justice, say, about non-discrimination. Further, again unlike the case of gambling, we can imagine arriving at a philosophically persuasive view about how to solve the priorities problem or any of the other unsolved rationing problems. Such a view might "defeat" decisions about fairness arrived at through the process. The "defeasible" fairness that results is the most we can claim for the outcome of our fair process.

6. What Does a Right to Health or Health Care Imply?

Whatever approach we take to addressing the problem of reasonable disagreement about priorities—equity weights, better philosophical solutions, procedural justice—the need to make decisions about priorities has implications for how we should understand a right to health care (or health). Specifically, a right to health care should be understood as a claim that individuals with certain health conditions can make to be provided with a reasonable array of services that address those needs. An array is reasonable for a given society if it includes services that emerge from the approach we take to determining a reasonable benefit package, given a society's technological capabilities and resource limits. A right to health care entitles individuals to contingent claims on services that are part of that array (Daniels 2008). In addition, the array must be fairly financed to assure those entitlements.

This restriction to a reasonable array of resources, as judged by what a fair deliberative process would choose, has important implications for how we understand a right to health or health care. Obviously, it makes no sense to claim that someone's right to health is violated if they lack health despite the efforts of a society to distribute all the socially controllable factors affecting health in a just way. If someone dies of a fatal disease that was not preventable or curable by a fair allocation of factors affecting health, then her right is not violated. So a right to health is not a right to health, period, but a right to a fair distribution of the socially controllable factors affecting health. Similarly, if we are talking more narrowly about medical services, a right to health care is a right to equitable (usually equal) access to a service that is part of a reasonable array of such services for that society.

One strong implication of this view is that from the right claim to health care, we cannot infer an entitlement to a specific treatment for a condition simply from the presence of that condition and the assertion—even constitutionally assured—of a right to health care. Only if that treatment is part of a reasonable array of services for that society (including, perhaps, what results from other societies providing assistance to that society as a matter of global justice), is it an entitlement. That is because all individuals are entitled only to access to that reasonable array of services, not to any treatment that might be available to some people anywhere in the cosmos.

The point is similar if we are talking about a human right to health care, as affirmed by international agreements on human rights. Such rights are progressively realized in each society. What counts as the entitlements such a right brings must reflect decisions made in a given society about the progressive realization of that right. Globally, there must be

a decision about what other societies owe a given society by way of assistance, in light of that society's decisions about how to progressively realize a right to health care. Arguably, a form of fair, deliberative process is needed to specify the content of a progressive realization of a right to health care in a given society (Gruskin and Daniels 2008).

Related Topics

Contractualism and Political Liberalism, Utilitarianism and Consequentialism, Luck Egalitarianism, The Difference Principle, Needs and Distributive Justice, The Capability Approach (and Social Justice), Equality, Democracy

References

Anderson, E. (1999) "What is the Point of Equality?" *Ethics* 109: 287–337.

Arneson, R. (1988) "Equality and Equal Opportunity for Welfare," *Philosophical Studies* 54: 79–95.

Boorse, C. (1997) "A Rebuttal on Health," in J.M. Humber and R.F. Almeder (eds), *What is Disease?*, Totowa, NJ: Humana Press.

Broome, J. (2002) "Measuring the Burden of Disease by Aggregating Well-Being," in C.J.L. Murray, J.A. Salomon, C.D. Mathers and A.D. Lopez, *Summary Measures of Population Health: Concepts, Ethics, Measurement and Applications*, Geneva: World Health Organization, <http://whqlibdoc.who.int/publications/2002/9241545518_Chap3.1.pdf>, viewed January 27, 2011.

Cohen, G.A. (1989) "On the Currency of Egalitarian Justice," *Ethics* 99: 906–44.

Daniels, N. (1993) "Rationing Fairly: Programmatic Considerations," *Bioethics* 7: 224–33.

—— (2008) *Just Health*, New York: Cambridge University Press.

—— (2011) "Social and Individual Responsibility for Health," in C. Knight and Z. Stemplowska (eds), *Distributive Justice and Responsibility*, Oxford and New York: Oxford University Press.

—— and Sabin, J.E. (2002) *Setting Limits Fairly: Can We Learn to Share Medical Resources?*, 1st edn, New York: Oxford University Press.

—— and Sabin, J.E. (2008) *Setting Limits Fairly: Can We Learn to Share Medical Resources?*, 2nd edn, New York: Oxford University Press.

——, Rose, S. and Zide, E.D. (2009) "Disability, Adaptation, and Inclusion," in K. Brownlee and A. Cureton (eds), *Disability and Disadvantage: Re-examining Topics in Moral and Political Philosophy*, New York: Oxford University Press.

Dworkin, R. (1981) "What is Equality? Part 1: Equality of Welfare," *Philosophy and Public Affairs* 10: 185–246.

Field, M.J. and Gold, M.R. (eds) (1998) *Summarizing Population Health: Directions for the Development and Application of Population Metrics*, Institute of Medicine, Division of Health Care Services Committee on Summary Measures of Population Health, Washington: National Academy Press, <http://www.nap.edu/openbook.php?record_id=6124&page=3>, viewed January 27, 2011.

Gruskin, S. and Daniels, N. (2008) "Justice and Human Rights: Priority Setting and Fair Deliberative Process," *American Journal of Public Health* 98: 1573–77.

Haninger, K. (2006) *The Role of Equity Weights in Allocating Health and Medical Resources*, Cambridge, MA: Center for Risk Analysis, Harvard University.

Institute of Medicine of the National Academies (IOM) (2010) *For the Public's Health: The Role of Measurement in Action and Accountability*, Committee on Public Health Strategies to Improve Public Health, <http://www.iom.edu/~/media/Files/Report%20Files/2010/For-the-Publics-Health-The-Role-of-Measurement-in-Action-and-Accountability/Report%20Release%20Slides.pdf>, viewed January 27, 2011.

Kawachi, I. and Kennedy, B. (2002) *The Health of Nations: Why Inequality is Harmful to Your Health*, New York: New Press.

Kingma, E. (2010) "Paracetamol, Poison, and Polio: Why Boorse's Account of Function Fails to Distinguish Health and Disease," *British Journal for the Philosophy of Science* 61: 241–64.

Marmot, M.G., Rose, G., Shipley, M. and Hamilton, P.J. (1978) "Employment Grade and Coronary Heart Disease in British Civil Servants," *Journal of Epidemiology and Community Health* 32: 244–9.

Marmot, P. (1991) "Inequalities among British Civil Servants: The Whitehall II Study," *The Lancet* 337: 1387–93.

McKeown, T. (1976) *The Role of Medicine: Dream, Mirage, or Nemesis?*, London: Nuffield Provincial Hospitals Trust.

—— and Record, R.G. (1962) "Reasons for the Decline of Mortality in England and Wales During the Nineteenth Century," *Population Studies* 16: 94–122.

The Organisation for Economic and Co-operation and Development (OECD) (2009) "Comparing U.S. Healthcare Spending with Other OECD Countries," *Seeking Alpha*, <http://seekingalpha.com/article/146992-comparing-u-s-healthcare spending-with-other-oecd-countries>, viewed January 24, 2011.

Rawls, J. (1971) *A Theory of Justice*, Cambridge, MA: Harvard University Press.

Sachs, J. (chair) (2001) *Macroeconomics and Health: Investing in Health for Economic Development*, World Health Organization, Commission on Macroeconomics and Health, Geneva: World Health Organization, <http://www.cid.harvard.edu/archive/cmh/cmhreport.pdf>, viewed January 27, 2011.

Sen, A.K. (1980) "Equality of What?," in McMurrin, S. (ed.), *The Tanner Lectures on Human Values*, Salt Lake City: University of Utah Press.

—— (1992) *Inequality Reexamined*, Cambridge, MA: Harvard University Press.

Silvers, A. (2003) "On the Possibility and Desirability of Constructing a Neutral Conception of Disability," *Theoretical Medicine and Bioethics* 24: 471–87.

Smedley, B.D., Stith, A.Y. and Nelson, A.R. (2003) *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care*, Institute of Medicine of the National Academies, Board on Health Sciences Policy, Washington, DC: National Academies Press, <http://www.nap.edu/openbook.php?isbn=030908265X>, viewed January 27, 2011.

Szreter, S. (2002) "Rethinking McKeown: The Relationship between Public Health and Social Change," *American Journal of Public Health* 92: 722–5.

Wakefield, J.C. (1992) "The Concept of Mental Disorder: On the Boundary between Biological Facts and Social Values," *American Psychologist* 47: 373–88.

World Health Organization (WHO) (1946) *Preamble to the Constitution of the World Health Organization as Adopted by the International Health Conference, New York, 19–22 June, 1946*, Official Records of the World Health Organization, no. 2, p. 100, signed on July 22, 1946 by the representatives of 61 States and entered into force on April 7, 1948, Geneva.

—— (2001) *International Classification of Functioning, Disability, and Health*, Geneva: World Health Organization.

—— (2002) *The World Health Report*, World Health Organization, <http://www.who.int/whr/2002/overview/en/index.html>, viewed June 23, 2012.

—— (2008) *Closing the Gap in a Generation: Health Equity through Action on the Social Determinants of Health*, Commission on Social Determinants of Health—Final Report, Copenhagen: World Health Organization Regional Office for Europe, <http://www.who.int/social_determinants/thecommission/finalreport/en/index.html>, viewed January 24, 2011.

64
MARRIAGE, SEX, AND THE FAMILY

David Archard

What can be said by political philosophy about the topics of marriage, sex and the family? Political philosophical scrutiny stands contrasted with aesthetic, moral or prudential appraisal. Yet, at first glance marriage, sex and family—the sphere of the intimate—seems to lie beyond the reach of politics, and thus of the discourse of political philosophy. The intimate is "private" and the private is essentially contrasted with the public or political. However, this characterization view rests on confusions about what is meant by the "private," a topic to which Section 1 is devoted.

Such an attitude also ignores the important claim that the "personal is political," a slogan of the second-wave feminists of the 1960s and 1970s. They drew attention to the fact that the patriarchal oppression of women by men is not restricted to the public practices and institutions of the work-place, law courts, and political assemblies. Indeed it may be displayed most acutely within the private domain of the intimate including, centrally, sexual relationships, marriage and the family. Phenomena such as domestic violence or the unequal division of labor within the home can be both more symptomatic of patriarchy, and at the same time more devastating in their effects upon women than the visible public gender-based injustices of discrimination in employment or the inequitable distribution of political and social rights.

1. Public and Private

One important mark of the liberal society, it is often remarked, is that it observes and protects the distinction between the private and the public. Surely the conduct of sexual relations, married life, and the enjoyment of the family belong, if nothing else does, to the protected sphere of the private.

However, this distinction is not a simple one. The bipolarity may not capture what is most important about social life, and it is debatable which of the two concepts within the distinction is primary (Benn and Gaus 1983). There are at least three relevant but different senses of "private" in play. First there is a putative right to privacy mooted as a fundamental civil right at the end of the nineteenth century (Warren and Brandeis 1890). Some philosophers deny that this suggested right is anything other than a cluster of rights each of which can already be found as elements of other basic rights (Thomson 1975). Others are more disposed to see such a right as protecting aspects of a person's life that are essentially private. These require protection against the public

gaze because, for example, their privacy is a condition of their having value or because human beings cannot healthily live if everything they are and do is open to the scrutiny of others (Nagel 1995, 1998). If there is a right to privacy in this first sense of the term it is because there are features of human beings and their lives that are properly marked out as private. Defecation and sex seem clear and evident examples. However, it is far less clear why family life should be protected by the right to privacy for the same reasons or to the same extent.

A second sense of private derives from the application of the familiar liberal principles of individual liberty constrained only by a harm condition (Mill 1859). On this understanding of "private," private matters are those that should not be subject to official regulation or social control inasmuch as they do not occasion harm to anyone other than their willing participants. However, it is not the case that the private in this second sense is delimited by clear and precise boundaries that can be fixed in advance of a determination of what should and what should not be a matter of public control. Within any liberal society a "frontier must be drawn between the area of private life and that of public authority. Where it is to be drawn is a matter of argument, indeed of haggling" (Berlin 1969: 124).

This sense of "private" is defined in the opposite fashion to the first sense. Thus it is not the case that a private matter is first picked out and then deemed, in virtue of its privacy, to be exempt from public control. Rather the proper limits to public control are fixed (as a result of argument and "haggling") such that the private is, in consequence, what can claim exemption from such control. For example, we would not define the family and its life as "private" and then justify the right to familial privacy in virtue of that definition. Instead we might judge that what happens within families for the most part falls outside the limits that have been set for the justified exercise of public power over the lives of individuals. These limits have been set by appeal not to the nature of the "private" but to the proper limits of personal liberty.

The third sense of "private" (or, more properly, non-public) characterizes the nature of reasons that might be advanced in justification of the exercise of coercive political power over individuals. To John Rawls, following Kant, is owed the idea of public reason (Rawls 1993). For Rawls modern liberal democratic societies are characterized by the "fact of reasonable pluralism." As a result of their exercise of the deliberative liberties—of thought, conscience, and speech—and in free association with other citizens, individuals form, and cleave to, very different understandings of the good or best way to lead their lives. Such differences are significant, deep-lying and enduring. A liberal state must recognize and respect these differences, as must its members, if reasonable terms of social cooperation with their distributed burdens and benefits are to be adopted and maintained.

Rawls, and political liberals following him, argue that the agreed rules and fundamental laws that govern our shared life must be the outcome of public deliberations in which citizens advance their case using only public reason. This is reason that is used in public, that justifies the use of the state's coercive power to sustain cooperation, and can command agreement in principle amongst reasonable persons. It thus comprises generally endorsed facts and simple principles of argument. By contrast, non-public reasons are those that citizens might employ to understand and justify their own particular conception of the good. These will derive from moral and religious beliefs that are not shared by their fellow citizens. Consider then how policy and law might be justified in respect of marriage, sexual behavior and the family. Individuals would be required to advance in defense of any proposal only what could be justified by the standards of public reason.

They are not permitted, for instance, to argue that gays or lesbians may not be married or be the legally recognized parents of dependent children because their lifestyle is sinful or morally depraved. However, they might argue that the best interests of children are not served by allowing same-sex couples to act as their guardians.

The three senses of private are clearly distinct; they provide different ways in which the sphere of the intimate may be both conceptualized and evaluated within political philosophy.

2. Sex

Sex matters to us. We are sexed beings who live in a highly sexualized world, surrounded by publicly displayed sexual images that advertise products and the desirability of certain lifestyles. The conduct of sexual relationships with others is a highly important part of our lives, such that the denial to us of the freedom to choose who, when, where and in what manner to have any such relationships constitutes a grave wrong. Punishing the achievement of sexual satisfaction creates "misery of a quite special degree" (Hart 1963: 22). There is, in consequence, an onus on the state to justify interferences with our sexual liberty.

A moral judgment that some kind of sexual behavior is wrong need not be a sufficient ground for the criminalization of that behavior. We also need to be satisfied that the moral costs of using the law are outweighed by the benefits of proscription. Nevertheless, an initial review of sexual morality is appropriate.

Here the broad contrast is between traditional, conservative defenses of an orthodox sexual morality which might be associated with, but need not rest for its ultimate defense upon, religious doctrine, and a liberal defense of the maximum degree of sexual freedom consistent with principles of individual consent and third-party harms (Belliotti 1993; Primoratz 1999). One basis for the conservative view understands sex as having a natural end or function, which is reproduction. Thus reproductive-type actions are the only ones permissible. The qualification "reproductive-type" allows that sex between a married husband and post-menopausal wife, for instance, is permitted. On this account all forms of non-heterosexual sexual activity, and all non-procreative forms of heterosexual activity—for instance, mutual masturbation, oral sex, anal sex—are illicit.

The stronger claims that only married heterosexual activity is morally permissible, and that the use of contraceptives is forbidden, require other arguments. These will appeal, respectively, to a defense of marriage discussed below, and a critique of the artificial by contrast with the natural. Conservatives need not believe that immoral sexual practices should be against the law; they will think that their immorality provides the state with reasons not to encourage, condone or facilitate these practices.

A different conservative argument is as follows. Homosexuality is abnormal because it is a functional misuse of genitalia for purposes other than reproduction. This entails not that it is immoral, only that it is seriously imprudent and productive of great unhappiness. Nevertheless, it would follow that the state has an obligation to enact such legislation as makes it less likely that individuals will acquire and act on homosexual dispositions, thereby having a very unhappy life (Levin 1984).

Conservative moral accounts that appeal to the natural end or purpose of sexual activity are vulnerable to the charge that sex does not have a single or even a dominant function. Sex can be procreative but need not be, and non-procreative sex can be an invaluable expression of mutual love or simply a source of enormous pleasure.

A defense of the view that conservative sexual morality should be legally enforced need only appeal to the purported facts that this morality is generally endorsed and that its enforcement is crucial to the continued integrity of society (Devlin 1965). Criticism of such a claim will distinguish between the basic morality that may indeed be socially essential, and the prevailing conventional morality that need not be and might, anyway, change (Hart 1963).

By contrast the liberal attitude to what sexual behavior should be legally allowed is based upon a strong presumption that individuals ought to be accorded sexual liberty constrained only by the harm principle. Such sex is private in the second sense distinguished above. Understanding what this constraint entails requires an appreciation both of consent and of third-party harms. Consistent with the legal precept, *volenti non fit injuria*, consensual sexual behavior cannot be harmful to its competent participants so long as such participation is willing and informed. Consent is a normatively basic power that transforms the moral relations between individuals, such that what would otherwise be impermissible is allowable. Conversely, what would be allowed is, in the absence of consent, disallowed. Thus, dissent makes sexual intercourse rape.

There are two principles of sexual consent: whatever is consensual is allowable, and whatever is not consensual is impermissible. The much criticized English legal judgment, *R. v. Brown* (1993), held, against the first principle, that extreme sadomasochistic sexual activity, even if conducted by consenting adults in private, should not be permitted. In support the English judges argued variously from the need to show that even consensual sexual activity serves a public good to the impropriety of legally condoning cruelty.

The principal difficulties facing the liberal account of sexual morality lie in specifying both exactly how consent is to be understood as given and when circumstances vitiate consent (Archard 1993). The law on rape is the site for much debate in respect of both problems. Feminists, for instance, will insist that the only acceptable standard for the giving and withholding of consent is the making of explicit verbal statements (Pineau 1989), whereas others will allow that men may reasonably rely on behavior that falls under accepted social conventions to infer the giving of consent (Husak and Thomas 1992). Again, there has been much discussion of what kinds of threats, and even offers, might suffice to vitiate a woman's consent.

The law on prostitution is another area of sexual activity in which an appeal to the basic freedoms of prostitutes to sell their sexual services, and of their clients to buy them, is contested. Criticisms of prostitution range from the feminist charge that it is degrading to women as a whole, to the claim, consistent with liberal principles, that prostitutes, or at least a substantial majority of them, do not pursue their professional activities voluntarily. Interestingly the putative severe harms done to prostitutes by their work, especially if it is pursued for many years, may give some liberals a reason to embrace limited paternalism (Marneffe 2009).

The liberal must also take account of third-party harms. These may include the general social harms of permitting some forms of sexual activity. For instance, prostitution can be criticized for subverting the valuable institution of marriage. A putative source of significant third-party harm is the public offense occasioned by sexual behavior—either at its public performance or in the bare knowledge of its private practice. Those who believe that the liberal state should not discount offensiveness need carefully to evaluate the emotions that inform reasonable, and properly weighted, offense (Nussbaum 2006).

3. Marriage

Marriage can be defined so that no choices are open to society as to who may enter into the state. If marriage is the union of one man and one woman, then polygamy and gay marriages are ruled out by conceptual fiat. However, if the definition of marriage is negotiable then important issues need to be resolved. These are basically twofold: who can be married and what does it involve?

Marriage is constituted by a set of rights, duties, permissions, privileges and exemptions that are legally recognized, protected and enforced. For example, married partners do not have to bear legal witness against one another; they can claim official financial benefits or tax breaks; individuals can claim citizenship in the country of their married partners. Marriage also confers a publicly acknowledged status on those who are married. Additionally it has a generally accepted social meaning. To be married is to stand in a certain privileged relation not enjoyed by the unmarried. This social meaning can be complex, contested, the product of various historical developments and of a diverse network of cultural practices and institutions. For example, marriage once was solemnized only through religious ritual and canon law in "the eyes of God." But this does not now need to be the case; indeed marriage has come to be recognized in some contexts as having an exclusively civil status.

Marriage may have value for those who are permitted to enter into it because it is an officially recognized state (Wedgwood 1999). It is the fact of being married that is valuable for its parties over and above the particular listed benefits that the married state brings with it. Then the choice to be married is of value only to those individuals who live within a society that has the institution of free marriage (Raz 1986).

Evaluation of marriage can thus turn on both what it involves—the package of rights and duties conferred, together with its social meaning—and on the scope of persons allowed to marry. If marriage is generally of significant value to those who choose it then it is prima facie unjust to deny some individuals the right to marry without demonstration of good reason for such exclusion. In addition to serving the interest of individuals in being married, there might be social or public interests served by the institution of marriage. Moreover those interests may be only or best promoted if marriage is defined in certain ways. For instance, if one believes that society has a legitimate interest in how children are brought up as its future members, and if one believes that children benefit from stable parenting, then one might have reason to promote marriage as the best way to ensure such an upbringing. Indeed one might have reason to think that society should make it difficult for married parents to dissolve a marriage (Galston 1996).

Of course a liberal state could choose not to recognize marriage publicly and to leave individuals the freedom to make their own arrangements. These might, for instance, be the subject of legally binding contracts voluntarily entered into, or they might involve endorsement of their relation by means of the ceremonial and ritual practices of religious institutions. Although the state should decide whether to enforce any such contracts, or to permit religious institutions to conduct certain ceremonies, this option would effectively amount to the privatization or disestablishment of marriage. Moreover, it could be seen as an obvious way to avoid the official favoring of some particular and contested conception of marriage (Metz 2007).

However, if the value of marriage lies precisely in its being the public recognition of a certain kind of relationship then there is a reason for the state not to disestablish the institution. Yet there then threatens a problem arising from a liberal state's commitment to

neutrality on the question of the good. A key principle of contemporary liberalism is that a state should not in its laws and policies presume the superiority of any conception of the good (Dworkin 1988). Neutrality is threatened by the public recognition of marriage in two respects. First the promotion of marriage as such—by means of rights, privileges and exemptions denied to the unmarried—might be thought to promote a particular conception of the good, namely that being married is morally better than being unmarried. Second, in its definition of what marriage involves and who can be married, the state might be thought to favor some lifestyles over others. For example, denying marriage to same-sex couples could be argued to discriminate against homosexuals and lesbians on the grounds that their mode of life is morally inferior to that of heterosexuals.

One reply to the first criticism is that the state need not directly promote marriage as instantiating a particular and controversial conception of the good. Indeed the state simply facilitates those who view marriage as embodying certain values in realizing those values. The state does not so much say that marriage is a virtuous condition, as allow those who do so regard it to enter into marriage (Wedgwood 1999). Nevertheless it might still be responded that the state is, in effect, endorsing some desires that are inauthentic or only the product of social pressures.

A second kind of reply is to insist that an appropriately defined marriage would satisfy the requirements of public reason, and, further, that its official support would be required if individuals are to attain goods understood as basic by liberalism. Thus Elizabeth Brake argues for legal recognition of only "minimal marriage." This is defined as one whose parties can be of any number and sex, and who are free to stipulate the nature of the relationship and its rights and duties. Moreover, minimal marriage, legally recognized, provides the social bases for caring relationships. Such bases are recognized within Rawlsian liberalism as primary goods (Brake 2010). Of course, the idea of minimal marriage consistent with public reason may be in tension with the conventional social meanings of marriage, those norms or expectations as to what marriage is, that give marriage its value for those who seek it.

The second threat to liberal neutrality lies in any definition of what marriage involves and who can be married that appears to favor some conception of the good. Conservatives will insist that marriage can only be of a certain kind. John Finnis, for example, defines marriage as a "lifelong and exclusive sexual commitment to a single spouse, in an institution, oriented towards, and socially supported for the sake of, the children whom this sexual union may well generate" (Finnis 2008: 396). However, an appeal to conventional understandings in support of this definition excludes the possibility of social change. Furthermore, a defense of the definition on explicitly normative grounds is controversial even by its own terms (Garrett 2008).

A state that denies an opportunity to marry to all those of its citizens who wish to do so violates both neutrality and equality. Yet consider the exclusion of same-sex couples from the possibility of legally recognized marriage. Here we must acknowledge a genuine difference of sincere opinion on the morality of same-sex relations. Either to publicly recognize such relations or not to do so would be to take the side of one set of opinions. Jeff Jordan has argued that in the case of a dilemma such as this, a refusal to recognize same-sex marriages is accommodating to the losing side in a way that recognition is not. This is because the liberal state can consistently both refuse legally to recognize same-sex marriages and legally tolerate same-sex relations (Jordan 1995). In response it has been suggested that by analogy Jordan's argumentative schema would tell against mixed race marriages (Boonin 1999).

It does not, of course, follow that permitting everyone who wishes to marry to do so is a welcome extension of marriage as traditionally understood. Some feminists resist such an inference, arguing that marriage is a fundamentally unjust institution or one that is so unfavorable to women that a freedom to marry should be thought of as no more welcome than a freedom to be a slave-owner (Card 1996, 2007). An argument of this kind might tell against marriage understood as constituted by a particular set of entitlements and duties or as having a particular social meaning. It need not tell against marriage as such.

4. Family

Whatever other disputes are entered into concerning the family it seems indubitable that the institution has been both enduring and widespread. Indeed it is hard to think of any society that has not had some form of family serving as the principal means by which its young have been reared. Nevertheless, we owe to Plato the idea that—at least with respect to the ruling class of his imagined *Republic*—there are good reasons to abolish the family. Plato's abolitionism is justified by eugenic perfectionism and the importance of securing the undivided collective loyalty of his rulers (Plato 1961). Amy Gutmann helpfully distinguishes between the ideal types of a "family state," of which Plato's *Republic* is a good example, in which the state or its agents take on sole responsibility for the upbringing of children, and a "state of families" in which the state devolves responsibility for the care of its young to families (Gutmann 1987: ch. 1). The latter is generally favored within liberal democratic societies, although the degree and extent of official involvement in the lives of families can obviously vary greatly.

The family has not been without its critics and the criticisms have varied from the psychoanalytic to the political. Defenses of the family can either be positive, extolling its unique and significant virtues, or "Churchillian" in the sense in which Winston Churchill gave faint praise to democracy as the worst form of government excepting all the others.

In discussions of the family it is important to heed the caution of social scientists who write on the subject. They note the enormous variety of observable family forms, and also note the extent to which a particular form—in which two married parents bring up their own biologically related offspring—is extolled or favored as *the* family par excellence. In order to avoid endorsement of any one family form, and in recognition of familial pluralism, many contemporary social scientists prefer not to use the definite article and, instead, to adopt the nomenclature of "families." However, it is perfectly possible to employ some minimal definition of the family which illuminates the common features of its various instances without displaying any preference for some particular kind.

At its most basic a family is a group of adults taking exclusive custodial care of dependent children. Any satisfactory normative account of the family must thus, in the first instance, say something about how some adults come to be in the position of acting as parents, and what rights and duties come with the assumption of that role. Here most of the interesting philosophical work has been done within "applied ethics" (Bayne and Kolers 2010). However, its relevance to political philosophy is clear. If someone does have a right to rear a child, then that is a right that the state should enforce—against other individuals who are not permitted to assume a custodial role in respect of the child—and it is a right that should constrain the actions of the state,

limiting the extent, for instance, to which it is permitted to intervene into familial life. Similarly, if there are parental duties of a certain kind then the state should enforce them.

To political philosophy is owed one influential if discredited view of parental rights. This is John Locke's proprietarianism (Locke 1698). Locke's theory of the justified acquisition of private property—through the admixture of labor with raw nature—extends naturally to procreation, although Locke himself resisted the extension for reasons widely thought inadequate to the task (Nozick 1974). Biological parents create their children and, in consequence, they might be thought to own them. Parental proprietarianism offers an account both of how parents come to have rights over children and of what those rights are. If they are property rights then it would seem that parents may alienate, sell and destroy the children they own. Parental proprietarianism is widely repudiated but it continues to exercise considerable influence in political and legal thinking about parental entitlements.

The liberal state is *parens patriae* and, as such, owes a duty of care and protection to the weak and vulnerable within its jurisdiction, including, most obviously, children. On the liberal view, broadly conceived, parents should be entrusted by the state to care for the children within their care—children they might have a right or presumptive liberty to look after but which they do not own—the state only intervening when the standard of care provided falls below a critical threshold. This is normally when the children are suffering or are at risk of suffering significant harm.

Appeal to the interests of children will also provide an answer to the question of what kinds of family the liberal state should permit. It would violate the precept of neutrality to deny any prospective parents the opportunity to rear children on the grounds that their lifestyle was sinful or morally base. Note, crucially, that a combination of cultural changes, biotechnological developments, and legal reforms means that an extraordinary range of familial possibilities present themselves, possibilities defined in terms of the number, sexual orientation and marital relationship of the adults, and their relation to the dependent children. Conservative defenders of the traditional family can avoid violating the principle of liberal neutrality by displaying the putative harms that are done to children, and, beyond, to society at large by the toleration or encouragement of non-traditional family forms (Almond 2006).

Within political philosophy criticism of the family has mainly focused on the family as a site and as a source of injustice. In respect of the first, feminists have called attention to the inequitable nature and consequences of a gendered division of labor within the traditional family (Okin 1989). They have also noted the extent to which political philosophy, historically, has failed to acknowledge this injustice and has done so precisely inasmuch as the domestic sphere, to which women are consigned, has been regarded as private and beyond its proper scope (Okin 1979). Women are normally assigned to the role of domestic carer whilst the man is the principal earner. The consequences are unjust in two ways. First, women are denied the same employment prospects as men; second the family serves as a "school of injustice" by inculcating in the children an acceptance of, and passage into, those same gendered roles.

Such criticism tells only against the family as traditionally constituted, and is in danger of presuming that the family can only have a certain form (Kymlicka 1991). Moreover, it is not clear that a liberal can do anything other than acknowledge the lexical priority of personal liberty, which would allow individuals to form familial associations with an unequal or at least differentiated distribution of roles if they so chose.

Inasmuch as equal liberty is demanded by justice, it would not then be appropriate somehow to balance justice against liberty (J. Cohen 1992).

The family is a source of injustice inasmuch as disadvantage (and advantage) is inherited from one's parents. It seems impossible to combine liberal commitments to equality of life prospects, formal equality of opportunity, and the freedom of parents to rear children as they choose (Fishkin 1983). A liberal state can obviously regulate the transmission of advantage in respect of the inheritance of money and property; it might even be open to it to regulate in respect of genetic inheritance. However, evidence suggests that the most potent source of the differences in a child's eventual life prospects can be attributed to qualities of parenting which are correlated with socio-economic position and lie beyond obvious means of control.

A liberal egalitarian who views inequalities in life prospects as unfair if attributable to brute circumstantial luck thus has a serious problem with the family. For no one chooses their family, and yet the circumstances of our birth make a huge difference to how well (or badly) our life subsequently goes. The problem prompted John Rawls to ask, if only rhetorically, whether the family should be abolished (Rawls 1999: 448), and has stimulated Rawlsian liberals to explore whether liberal justice can be consonant with the existence of the family. One important reply argues that parents derive something of great value from the discharge of their duty to care for children, that this cannot be derived from any other kind of relationship or activity, that the pursuit of this valuable activity of parenting is a choice permitted within the terms of social justice, but must be pursued only to the extent that there is no significant subversion of social justice (Brighouse and Swift 2006, 2009).

5. Conclusion

It is tempting to view the sphere of the intimate—sex, marriage and family—as beyond the scope of the political; and to do so because this sphere seems private, of enormous personal importance, and not properly subject to the brutish supervision and control of public institutions. Yet the personal is political not only inasmuch as these intimate affairs can, in themselves, be politically evaluated, but also because their effects on our lives are significant and enduring (G. A. Cohen 1997).

Related Topics

Plato's Political Philosophy, Locke, Liberalism, Feminism and the History of Political Philosophy, Luck Egalitarianism, Intergenerational Distributive Justice

References

Almond, B. (2006) *The Fragmenting Family*, Oxford: Oxford University Press.

Archard, D. (1993) *Sexual Consent*, Boulder, CO: Westview Press.

Bayne, T. and Kolers, A. (2010) "Parenthood and Procreation," *Stanford Encyclopaedia of Philosophy*, at <http://plato.stanford.edu/entries/parenthood/>, accessed March 1, 2011.

Belliotti, R. (1993) *Good Sex: Perspectives on Sexual Ethics*, Lawrence, KS: University Press of Kansas.

Benn, S.I. and Gaus, G.F. (eds.) (1983) *Public and Private in Social Life*, New York: St. Martin's Press.

Berlin, I. (1969) "Two Concepts of Liberty," in *Four Essays on Liberty*, Oxford: Oxford University Press.

Boonin, D. (1999) "Same-Sex Marriages and the Argument from Public Disagreement," *Journal of Social Philosophy* 30(2): 251–9.

Brake, E. (2010) "Minimal Marriage: What Political Liberalism Implies for Marriage Law," *Ethics* 102: 302–37.
Brighouse, H. and Swift, A. (2006) "Parents' Rights and the Value of the Family," *Ethics* 117: 80–108.
—— (2009) "Legitimate Parental Partiality," *Philosophy and Public Affairs* 37(1): 43–80.
Card, C. (1996) "Against Marriage and Motherhood," *Hypatia* 11(3): 1–23.
—— (2007) "Gay Divorce: Thoughts on the Legal Regulation of Marriage," *Hypatia* 22(1): 24–38.
Cohen, G.A. (1997) "Where the Action Is: On the Site of Distributive Justice," *Philosophy and Public Affairs* 26: 3–30.
Cohen, J. (1992) "Okin on Justice, Gender, and Family," *Canadian Journal of Philosophy* 22(2): 263–86.
Devlin, P. (1965) *The Enforcement of Morals*, Oxford: Oxford University Press.
Dworkin, R. (1988) "Liberalism," in *A Matter of Principle*, Cambridge, MA: Harvard University Press.
Finnis, J. (2008) "Marriage: A Basic and Exigent Good," *The Monist*, 91(3/4): 388–406.
Fishkin, James (1983) *Justice, Equal Opportunity, and the Family*, New Haven, CT: Yale University Press.
Galston, W. (1996) "The Reinstitutionalization of Marriage: Political Theory and Public Policy," in D. Popenoe, J. B. Elshtain and D. Blankenhom, eds., *Promises to Keep: Decline and Renewal of Marriage in America*, Lanham, MD: Rowman and Littlefield.
Garrett, J. (2008) "Why the Old Sexual Morality of the New Natural Law Undermines Traditional Marriage," *Social Theory and Practice* 34(4): 591–622.
Gutmann, A. (1987) *Democratic Education*, Princeton, NJ: Princeton University Press.
Hart, H.L.A. (1963) *Law, Liberty and Morality*, Oxford: Oxford University Press.
Husak, D. and Thomas, G. (1992) "Date Rape: Social Conventions and Reasonable Mistakes," *Law and Philosophy* 11: 95–120.
Jordan, J. (1995) "Is it Wrong to Discriminate on the Basis of Homosexuality?" *Journal of Social Philosophy* 25(1): 39–52.
Kymlicka, W. (1991) "Rethinking the Family," *Philosophy and Public Affairs* 20(1): 77–97.
Levin, M. (1984) "Why Homosexuality is Abnormal?" *The Monist* 67: 251–83.
Locke, John (1698) *Two Treatises of Government*, at <http://www.lonang.com/exlibris/locke/>, accessed March 1, 2011.
Marneffe, P. de (2009) *Liberalism and Prostitution*, Oxford: Oxford University Press.
Metz, T. (2007) "The Liberal Case for Disestablishing Marriage," *Contemporary Political Theory* 6(2): 196–217.
Mill, J.S. (1859) *On Liberty*, at <http://www.constitution.org/jsm/liberty.htm>, accessed March 1, 2011.
Nagel, T. (1995) "Personal Rights and Public Space," *Philosophy and Public Affairs* 24(2): 83–107.
—— (1998) "Concealment and Exposure," *Philosophy and Public Affairs* 27(1): 3–30.
Nozick, R. (1974) *Anarchy, State, and Utopia*, Oxford: Basil Blackwell.
Nussbaum, M. (2006) *Hiding from Humanity: Disgust, Shame and the Law*, Princeton, NJ: Princeton University Press.
Okin, S.M. (1979) *Women in Western Political Thought*, Princeton, NJ: Princeton University Press.
—— (1989) *Justice, Gender and the Family*, New York: Basic Books.
Pineau, L. (1989) "Date Rape: A Feminist Analysis," *Law and Philosophy* 8(2): 217–43.
Plato (1961) *The Republic*, trans. Paul Shorey, in Edith Hamilton and Huntington Cairns, eds., *The Collected Dialogues of Plato*, Princeton, NJ: Princeton University Press.
Primoratz, I. (1999) *Ethics and Sex*, London: Routledge.
Rawls, J. (1993) *Political Liberalism*, New York: Columbia University Press.
—— (1997) "The Idea of Public Reason Revisited," *The University of Chicago Law Review* 64(3): 765–807.
—— (1999) *A Theory of Justice*, Revised Edition, Oxford: Oxford University Press
Raz, J. (1986) *The Morality of Freedom*, Oxford: Clarendon Press.
R. v. Brown (1993) All ER 75.
Thomson, J.J. (1975) "The Right to Privacy," *Philosophy and Public Affairs* 4: 295–314.
Warren, S. and Brandeis, L. (1890) "The Right to Privacy," *Harvard Law Review* 4: 193–220.
Wedgwood, R. (1999) "The Fundamental Argument for Same-Sex Marriage," *The Journal of Political Philosophy* 7(3): 225–42.

Further Reading

Soble, A., ed. (2006) *Sex from Plato to Paglia: A Philosophical Encylcopedia*, 2 volumes, Westport, CT: Greenwood Press; Archard, D. (2003) *Children, Family and State*, Aldershot: Ashgate; Archard, D. (2010)

The Family: A Liberal Defence, London: Palgrave Macmillan; Archard, D. and Benatar, D. (2010) *Procreation and Parenthood: The Ethics of Bearing and Rearing Children*, Oxford: Oxford University Press; Cave, E.M. (2003) "Marital Pluralism: Making Marriage Safer for Love," *Journal of Social Philosophy* 34(3): 331–47; Kleingeld, J. (1998) "Just Love? Marriage and the Question of Justice," *Social Theory and Practice* 24(2): 261–81; Brake, E. (2009) "Marriage and Domestic Partnership," *Stanford Encyclopedia of Philosophy* <http://plato.stanford.edu/entries/marriage/>, accessed March 1, 2011; Brake, E. (2012). *Minimizing Marriage: Marriage, Morality and the Law*, New York: Oxford University Press.

65
WORK

Nien-hê Hsieh

In this chapter, I survey arguments that have been advanced for policies and institutions that are intended to regulate the manner in which work is carried out in the formal economy.

Following Philippe van Parijs, I define work as activity "that is geared to the production (whether pleasurable or not) of a benefit that is external to the performance of the activity itself—and is, therefore, also capable of being enjoyed by others" (1995: 137–8). This is not to say that work must be motivated by the fact of benefiting another person. An individual might be motivated to work solely by what she will receive in return. This definition also does not deny that work can be enjoyable or have intrinsic value. Gregory Pence (2001), for example, distinguishes three categories of productive activity: (1) "laboring," which is generally repetitious, lacking in intrinsic satisfaction and leaves little choice about how and when the work is done; (2) "workmanship," which involves greater discretion, pride of the worker in the products of work, and the use of higher human faculties; and (3) "callings," which is activity "directed towards accomplishing goals which tie in with larger goals in the community and world" (2001: 93–4). Although work is often most closely associated with laboring, as understood in this chapter, work can be just as much a matter of workmanship or a calling.

In focusing on work in the formal economy, I do not mean to deny the value of unpaid work, such as care for children by parents or volunteer services. Nor do I mean to downplay the significance of debates regarding the performance of work outside the formal economy, including debates about the appropriate division of labor between men and women in the household (Hochschild 1989; MacKinnon 1987; Okin 1989). I do not mean to overlook the important question of what kinds of work are permissible and excluded from the formal economy (Anderson 1993; Radin 1996; and Satz 2010). Rather, in focusing on the manner in which work is carried out in the formal economy, the chapter has two aims.

First, in the light of the continued dominance of market capitalism as a way to organize economic activity, one aim of the chapter is to examine ways in which contemporary scholarship has taken up traditional concerns about the impact of market capitalism on the manner in which work is carried out. By the manner in which work is carried out, I have in mind the activities performed by workers and the organization of that activity as opposed to other dimensions of work, such as pay or the distribution of income. According to one prominent tradition of thought—namely, Marxism—the manner in which work is organized under market capitalism is a matter of concern. The manner of organization is the *wage-labor relationship*, in which an individual is hired to perform some agreed-upon set of tasks over a certain period of time under the direction and supervision of another individual in exchange for an hourly wage or salary.

The second aim of this chapter is to highlight ways to respond to the criticism that contemporary philosophical scholarship has framed issues of justice largely as matters of distribution and has paid insufficient attention to questions about the nature and organization of economic production (Gould 1988: 133; Young 2006: 91). The aim is to point to ways to take up questions not only about justice in distribution, but also about justice in production.

The chapter is organized as follows. In the Section 1, I discuss ways in which the traditional Marxist concern with exploitation has been taken up in contemporary scholarship. In the Section 2, I turn to consider ways in which the concern with alienation has been addressed. In neither of these cases does it appear that the source of concern is the manner in which work is organized under the wage-labor relationship. Instead, in the case of exploitation, the injustice lies in the distribution of the means of production, and in the case of alienation, the aim is to ensure that workers have a meaningful range of options in deciding where to work. In the third section, I discuss a series of arguments that criticize the wage-labor relationship on liberal egalitarian grounds that relate directly to the manner in which work is organized and carried out. These arguments, however, do not specify what is distinctive about work as an area of concern from the perspective of justice. I conclude by outlining what this discussion suggests for developing a full account of justice in production.

1. Exploitation: A Matter of Justice in Production or Justice in Distribution?

On traditional interpretations of Marxism, workers are *exploited* under the system of wage labor in that they are not paid for the full value of what they produce by their labor. This difference results from the fact that capitalists own the means of production and that capitalists will hire workers only when they are able to extract this "surplus value." Exploitation is eliminated from the production process once private ownership of the means of production is replaced by collective ownership and workers no longer operate under the wage-labor relationship. If understood as a matter of justice, the wage-labor relationship, it seems, is inherently unjust. On this interpretation, the injustice lies in the manner in which work is organized under market capitalism (Kymlicka 2001: 177–8; Wood 2004: 242–4).

Under contemporary accounts of exploitation, however, the wage-labor relationship itself no longer appears to be the object of concern. Instead, injustice arises mainly because the distribution in access to the means of production deviates from a just distribution, which is determined independently of the manner in which work is organized. That is, it is not that the wage-labor relationship is inherently unjust but, rather, that the wage-labor relationship gives rise to injustice when the background distribution of access to the means of production is unjust. To see this, consider two contemporary analyses of exploitation.

In analyzing the justice of market transactions, G.A. Cohen writes:

> Marxists say that capitalists steal labour time from working people. But you can only steal from someone that which properly belongs to him. The Marxist critique of capitalist injustice therefore implies that the worker is proper owner of his labour time; he, no one else, has the right to decide what will be done with it.
>
> (1990: 366)

Underlying this first interpretation of what is wrong with the wage-labor relationship is a commitment to the *self-ownership thesis*, which is the libertarian claim that a person has rights over herself as a master has over a chattel slave, and according to which she owes no product or service to anyone else (Cohen 1995). The self-ownership thesis alone, however, cannot identify what is unjust about the wage-labor relationship. For example, self-ownership is held by many commentators to be consistent with entering into a wage-labor relationship (Nozick 1974); what matters instead is the background distribution of resources against which the terms of the relationship are settled (Kymlicka 2001: 181).

A second approach is to define exploitation explicitly in terms of an unjust distribution in access to the means of production. John Roemer's account is an example of such an approach (1982a, b, c, 1988). On Roemer's account, an individual is exploited under an economic regime if she would be better off in a hypothetical situation of resource equality in which she withdrew her labor along with her per capita share of external resources. According to Roemer, workers would be better off in a hypothetical situation of resource equality, and as such, are exploited under market capitalism. Richard Arneson (1981) also offers an account along these lines. On his account, the relevant comparison is how workers would fare under an egalitarian distribution not only of external resources, but also of talents. Workers are said to be exploited if they are made worse off by undeserved inequalities in either external resources or talents that others are able to use to their advantage. On both of these accounts, capitalists are able to extract surplus value from workers, so there is the possibility of exploitation. However, the transfer itself is not the source of injustice. Rather, what determines whether a given transfer is unjust is whether it is the result of an unjust distribution in access to the means of production.

In contrast to the traditional interpretation of exploitation, contemporary accounts, it seems, do not take the wage-labor relationship itself to be the relevant locus of injustice. If exploitation is understood as a matter of justice, on contemporary accounts, exploitation is not so much a matter of justice in production as it is a matter of justice in distribution. The question arises whether there are other characterizations of the wage-labor relationship that can be understood on contemporary terms as matters of justice in production. To address this question, the chapter turns to the Marxist critique that workers under the wage-labor relationship are alienated.

2. Alienation: A Case for Meaningful Work

In *The Manifesto of the Communist Party*, Frederick Engels and Karl Marx describe the impact of the wage-labor system on the content of work. "The work of the proletarians," they write, "has lost all individual character, and consequently, all charm for the workman. He becomes an appendage of the machine, and it is only the most simple, most monotonous, and most easily acquired knack, that is required of him" (Tucker 1978 [1972]: 479). Engaging in monotonous work represents one way in which workers under the system of wage-labor are *alienated*—that is, they are unable to develop and exercise the full range of powers that are associated with engaging in free productive activity, which Marx sees as essential to human nature. Workers are alienated because work is organized not toward the exercise of their full range of productive powers but, rather, toward increasing profits. Workers also are alienated because they do not control the means of production and thereby lack the opportunity to conceptualize the broader

ends of which their work is a part. On traditional interpretations of Marxism, alienation is a result of the manner in which work is organized (Wood 2004: chs. 2 and 3).

In contemporary scholarship, work that is structured to avoid the condition of alienation is sometimes referred to as "meaningful work." As defined by Richard Arneson, it is work that is

> interesting, that calls for intelligence and initiative, and that is attached to a job that gives the worker considerable freedom to decide how the work is to be done and a democratic say over the character of the work process and the policies pursued by the employing enterprise.
>
> (1987: 522)

In defining meaningful work as such, note that the criticism of alienation may be raised in a contemporary context not just about manufacturing work, but also about mental and psychological work, for example, that takes on the commoditized nature of manual work in the manufacturing sector (Hochschild 1983).

In the first instance, it might appear that contemporary accounts of alienation, as in the case of exploitation, do not take the wage-labor relationship itself to be the relevant object of concern. Consider, for example, Arneson's discussion of whether the state ought to ensure that all work qualifies as meaningful. Arneson argues against such a proposal on grounds that it fails to respect the preferences of people who are willing to trade off meaningful work in favor of other valuable goods, such as leisure time. "The core socialist objection to a capitalist market," he writes, "is that people who have fewer resources than others through no fault of their own do not have a fair chance to satisfy their preferences." The appropriate state response, according to Arneson, "is to tinker with the distribution of resources that individuals bring to market trading" (1987: 537).

The concern that the state in contemporary society ought to respect divergent preferences, however, does open the way for policies and institutions aimed at regulating the way in which work is carried out, and not just the underlying distribution of resources. David Miller (1989) offers one such argument. One way to understand the respect for divergent preferences is with reference to the liberal ideal that the state in contemporary society ought to remain neutral with respect to promoting different conceptions of the good life, which might include considerations about the value of work and the manner in which work is carried out. John Rawls, for example, acknowledges the value of meaningful work from the perspective of human flourishing and autonomy, but the value of meaningful work is not part of Rawls's argument for the institutions that comprise the basic structure (1999: 463–4). Miller (1989) argues that under market capitalism, worker cooperatives are at a disadvantage relative to capitalist firms. These disadvantages arise not because worker cooperatives are inherently less productive, but in part because individual economic actors, such as banks, will find it in their interest to transact with capitalist firms over worker cooperatives even if everyone would be just as well off with worker cooperatives. Under such conditions, if no one conception of the good life is to be favored with regard to the manner in which work is carried out, then worker cooperatives will require some form of subsidy if both they and capitalist firms are to be available as meaningful options for workers. Note that generalizing this argument relies on the view one adopts as to why capital hires labor, a question that is a matter of some debate (Boatright 2004; Dow 2003; Dow and Putterman 2000; Elster 1989; Hansmann 1996; Putterman 1988).

Neutrality-based arguments along these lines are not simply arguments about the distribution of opportunities and resources (Hsieh 2008: 76–7). They also are arguments for regulating and shaping the manner in which work is carried out in the formal economy. In this respect, they concern themselves directly with the organization of economic production. They are arguments about justice in production. At the same time, in contrast to traditional Marxist accounts of exploitation and alienation, neutrality-based arguments are not arguments about the inherent injustice of the wage-labor relationship. The point of neutrality-based arguments is that citizens have a meaningful range of options from which to choose and not that one manner of organizing work is inherently unjust or morally preferable to another. In turn, it could be asked whether there are arguments that criticize the wage-labor relationship on grounds that relate directly to the manner in which work is carried out, but do not invoke specific conceptions of the good life. In what follows, I discuss a set of arguments that claim to do just that.

3. Liberal Egalitarianism at Work

In this section, I discuss a series of liberal egalitarian arguments for granting workers a right to participate in the management and governance of the economic enterprises in which they work. By liberal egalitarianism, I have in mind a view committed to equal liberty for citizens along with some degree of social and economic equality, taking Rawls's theory of justice as a paradigmatic account (1999). These arguments are considered liberal egalitarian in the sense that they offer good reasons, on liberal egalitarian grounds, for rejecting the wage-labor relationship as the appropriate manner for organizing work. There are three main lines of argument, which I discuss elsewhere in greater detail (Hsieh 2008).

One line of argument is what Joshua Cohen calls the "parallel case argument" (1989: 27). According to this line of argument, the conditions that help ground a citizen's right to participate in the democratic governance of the state are also found in economic enterprises as far as workers are concerned. Unless one denies that citizens have a right to participate in political democracy, workers have a right to participate in the management and governance economic enterprises in which they work. Among the most well-known examples of this line of argument can be found in the work of Robert Dahl (1985) and Michael Walzer (1983). "To say that [democracy] is *not* justified in governing economic enterprises," writes Dahl, "is to imply that it is not justified in governing the state" (1985: 111). He writes, "like a state, then a firm can also be viewed as a political system in which relations of power exist between governments and the governed" (1985: 11). Walzer advances a similar argument, writing that "an economic enterprise seems very much like a town" (1983: 300). Versions of the parallel case argument also can be found in the work of Iris Marion Young (1979) and Christopher McMahon (1994).

The parallel case argument is an intuitively appealing line of argument. The claim to democratic governance in the political sphere is one that admits of a number of justifications and holds wide intuitive appeal, and many parallels can be drawn between the position of citizens in relation to government officials and the position of workers in relation to their managers and employers. For example, citizens are subject to the power of government officials, and if there are costs to exiting work, then workers also can be subject to the exercise of power. Economic theory points to a variety of costs. One is the result of a worker's investment in developing firm-specific human capital (Blair and

Stout 1999). Second, because it is costly to monitor workers, employers will find it in their interest to pay workers more than the market-clearing wage so that workers face a cost to exiting the firm (Akerlof and Yellen 1986). Third, there may be costs associated with finding and transitioning to a new job.

At the same time, objections have been raised that the relevant parallels cannot be drawn between states and economic enterprises. Consider, for example, discussions around the specific accounts by Dahl (Mayer 2001a, b; Dahl 2001) and McMahon (Hsieh 2007; McMahon 2007; Moriarty 2007; van Oosterhout 2007; and Smith 2007). Generalizing from these discussions, one objection concerns the difference in the range of matters over which government officials and managers have discretion. This difference is not only a matter of degree, but also a matter of kind, such that the considerations that help ground a right to political democracy do not apply to economic enterprises. A second objection concerns the power of enforcement. "If a firm doesn't like the way you do your job," asks Jan Narveson, "can it send men with guns who will put you in prison if you don't do it the way the boss says?" (1992: 53). A third objection that has been raised relates to the ends for which the mechanisms of governance are intended in the state and in economic enterprises. Economic enterprises are purposive. They have a limited set of ends that largely concern economically productive activity. In contrast, in the case of the state, it could be said there are no such comparable ends, or if there are, the ends are extremely underspecified. Although there is debate about this objection (Phillips and Margolis 1999; Hartman 2001; Moriarty 2005) along with the others, these objections should give us pause in moving too quickly from the case for political democracy to the case for worker participation in the management and governance of economic enterprises.

A second approach to grounding a claim on the part of workers to participate in the management and governance of economic enterprises is to ask what, if anything, is wrong about the wage-labor relationship from the perspective of the worker as an autonomous agent. In the contemporary literature, two answers that have been given are as follows. The first answer is that absent an institutional guarantee to participate in the management and governance of economic enterprises, workers lack adequate means to protect important interests from being adversely affected by managerial decisions (Brenkert 1992; Hsieh 2005; McCall 2001). The second is that because workers' actions are determined by managerial decisions, respect for workers' autonomy requires that they have a right to participate in the making of those decisions (Archer 1995; Brenkert 1992; Ellerman 1992; Gould 1988; McCall 2001; Werhane 1985). From the perspective of liberal egalitarianism, both the adequate protection of individual interests and respect for individual autonomy represent plausible bases on which to ground a worker's right to participate in the decision-making processes that affect those interests and her exercise of autonomy.

One question that has been raised about these arguments is whether they ground a similar right on the part of other parties affected by managerial decisions, including customers, suppliers, or local community members. Jeffrey Moriarty (2010) argues that they do. "The case for enfranchising some other stakeholders," he concludes, "is at least as strong, according to the reasons given in [these arguments], as the case for enfranchising some employees" (2010: 381). Moriarty's argument raises the question whether there is anything about the wage-labor relationship, at least as identified by these interest-based and autonomy-based arguments, that is unique or distinctive as grounds for concern. This is not to deny there might be something morally problematic

about the wage-labor relationship from the perspective of the worker as an autonomous agent. Rather, the point is there is nothing about her status as a worker that gives rise to a special concern.

A similar point can be made about a third line of argument for a right to participate in the management and governance of economic enterprises or a right to meaningful work. Arguments of this sort make the case that in the absence of such alternatives to the wage-labor relationship, workers will come to lack the capabilities or dispositions that are of concern to liberal egalitarians. Adina Schwartz, for example, argues that workers who lack meaningful work will be "made less capable of and less interested in rationally framing, pursuing, and adjusting their own plans during the rest of their time" (1982: 637). Along similar lines, building on the work of Carole Pateman (1970), Joshua Cohen argues for workplace democracy on grounds that the wage-labor relationship tends to foster "passivity and a narrower basis of political judgment" on the part of workers (1989: 28–9). Cohen (1989) generalizes this argument into what he calls the "psychological support argument." According to Cohen's version of the argument, two psychological conditions are of "special importance in a well-functioning democracy." The first is the "sense that social arrangements are malleable and subject to improvement, and that one's own efforts can contribute to their improvement." The second is "the capacity to judge in terms of common good, and an effective desire to act on such judgments." Because capitalist work relations "vest final authority in the owners of capital, they limit the extent of intra-firm democracy, thereby fostering passivity and a narrower basis of political judgment" (1989: 28–2).

Elsewhere, I have termed arguments of this sort, *formative arguments* (Hsieh 2008: 77–9). The question that arises for these arguments is whether working under the wage-labor relationship results in a diminished capacity for autonomy and political judgment or whether it is the lack of adequate opportunities, both in and out of work, to sustain a capacity for autonomy and broader political judgment that is the problem (Hsieh 2008: 78–9). If it is lack of opportunities both in and out of work that diminishes the capacity for autonomy, then the reasons for regulating or replacing the wage-labor relationship are largely pragmatic. That is, it might be because there are constraints to finding opportunities outside of work to sustain a capacity for autonomy and broader political judgment that work becomes a focal point for formative arguments. It is not because the wage-labor relationship is inherently a source of concern.

4. Why Work?

The discussion in the preceding section should give us pause before concluding we need to invoke a conception of the good life to criticize the wage-labor relationship as being inherently unjust or morally problematic. There is reason to hold that liberal egalitarianism contains within it the resources to criticize the wage-labor relationship not only as a matter of justice in distribution, but also as a matter of justice in production. At the same time, these arguments do not give us reason to limit this criticism to the domain of work. The case for granting non-workers a right to participate in the management and governance of economic enterprises is at least as strong, it seems, for enfranchising workers. And, in the case of the formative argument, more needs to be said as to what makes work the relevant domain of concern.

This brings us to what I take to be a key challenge for contemporary scholarship, at least of the liberal egalitarian sort, in moving toward a fuller account of justice in

production. On traditional Marxists' accounts, work is special. Work is the source of value and engaging in productive activity is essential to human nature. For these reasons, it is not difficult to understand why the manner in which work is conducted and organized is itself an object of concern. In moving to contemporary critiques of the wage-labor relationship, there has been a shift. While the focus on work and the wage-labor relationship has remained the same, the reasons for focusing on work and the wage-labor relationship are no longer unique to them. The distinctiveness of work as a domain of concern has fallen aside along the way.

In one sense, perhaps this is the way it should be. Injustice can occur in almost any domain of life and in this way, we avoid the risk of fetishizing work as the principal site of injustice. At the same time, if we are to develop a fuller account of justice in production, it seems we would do well to take up the challenge to understand if there is anything unique or distinctive about the domain of work as a matter of justice, and if so, what it is. Without that understanding our picture of justice in production remains incomplete.

Acknowledgments

I thank Gerald Gaus and Fred D'Agostino for their support and encouragement in writing this chapter. I also thank Harvard Business School for providing a hospitable environment in which to complete it.

Related Topics

Marx, Liberalism, Marxism and Contemporary Political Thought, Left Libertarianism, Freedom, Autonomy, Power, Authority and Legitimacy, Democracy, Paternalism, Moralism and Markets

References

Akerlof, G. and J. Yellen, eds. (1986) *Efficiency Wage Models of the Labor Market*, Cambridge: Cambridge University Press.

Anderson, E. (1993) *Value in Ethics and Economics*, Cambridge, MA: Harvard University Press.

Archer, R. (1995) *Economic Democracy*, Oxford: Clarendon Press.

Arneson, R. (1981) "What's Wrong with Exploitation?," *Ethics* 91: 202–27.

—— (1987) "Meaningful Work and Market Socialism," *Ethics* 97: 517–45.

Blair, M. and L. Stout (1999) "A Team Production Theory of Corporate Law," *Virginia Law Review* 84: 247–328.

Boatright, J. (2004) "Employee Governance and Ownership of the Firm," *Business Ethics Quarterly* 14: 1–21.

Brenkert, G. (1992) "Freedom, Participation, and Corporations: The Issue of Corporate (Economic) Democracy," *Business Ethics Quarterly* 2: 251–69.

Cohen, G.A. (1990) "Marxism and Contemporary Political Philosophy, or Why Nozick Exercises Some Marxists More Than He Does Any Egalitarian Liberal," *Canadian Journal of Philosophy* Supplementary Volume 16: 363–87.

—— (1995) *Self-Ownership, Freedom and Equality*, Cambridge: Cambridge University Press.

Cohen, J. (1989) "The Economic Basis of Deliberative Democracy," *Social Philosophy and Policy* 6: 25–50.

Dahl, R. (1985) *A Preface to Economic Democracy*, Berkeley: University of California Press.

—— (2001) "A Right to Workplace Democracy? Response to Robert Mayer," *The Review of Politics* 63: 249–53.

Dow, G. (2003) *Governing the Firm: Workers' Control in Theory and Practice*, Cambridge: Cambridge University Press.

—— and L. Putterman (2000) "Why Capital Suppliers Usually Hire Workers: What We Know and What We Need to Know," *Journal of Economic Behavior and Organization* 43: 319–36.
Ellerman, D. (1992) *Property and Contract in Economics: The Case for Economic Democracy*, Oxford: Blackwell.
Elster, J. (1986) "Self-realization in Work and Politics: The Marxist Conception of the Good Life," *Social Philosophy and Policy* 3: 97–126.
—— (1989) "From Here to There: If Co-operative Ownership Is So Desirable, Why Are There So Few Co-operatives?" *Social Philosophy and Policy* 6: 93–111.
Gould, C. (1988) *Rethinking Democracy*, Cambridge: Cambridge University Press.
Hansmann, H. (1996) *The Ownership of Enterprise*, Cambridge, MA: Belknap Press.
Hartman, E. (2001) "Moral Philosophy, Political Philosophy, and Organizational Ethics: A Response to Phillips and Margolis," *Business Ethics Quarterly* 11: 643–87.
Hochschild, A. (1983) *The Managed Heart: Commercialization of Human Feeling*, Berkeley: University of California Press.
—— (1989) *The Second Shift: Working Parents and the Revolution at Home*, New York: Viking Press.
Hsieh, N. (2005) "Rawlsian Justice and Workplace Republicanism," *Social Theory and Practice* 31: 115–42.
—— (2007) "Managers, Workers, and Authority," *Journal of Business Ethics* 71: 347–57.
—— (2008) "Survey Article: Justice in Production," *The Journal of Political Philosophy* 16: 72–100.
Kymlicka, W. (2001) *Contemporary Political Philosophy*, 2nd edn., Oxford: Oxford University Press.
MacKinnon, C. (1987) *Feminism Unmodified: Discourses on Life and Law*, Cambridge, MA: Harvard University Press.
McCall, J. (2001) "Employee Voice in Corporate Governance: A Defense of Strong Participation Rights," *Business Ethics Quarterly* 11: 195–213.
McMahon, C. (1994) *Authority and Democracy: A General Theory of Government and Management*, Princeton, NJ: Princeton University Press.
—— (2007) "Comments on Hsieh, Moriarty and Oosterhout," *Journal of Business Ethics* 71: 371–9.
Mayer, R. (2001a) "Robert Dahl and the Right to Workplace Democracy," *The Review of Politics* 63: 221–47.
—— (2001b) "A Rejoinder to Robert Dahl," *The Review of Politics* 63: 255–7.
Miller, D. (1989) *Market, State and Community: Theoretical Foundations of Market Socialism*, Oxford: Clarendon Press.
Moriarty, J. (2005) "On the Relevance of Political Philosophy to Business Ethics," *Business Ethics Quarterly* 15: 455–75.
—— (2007) "McMahon on Workplace Democracy," *Journal of Business Ethics* 71: 339–45.
—— (2010) "Participation in the Workplace: Are Employees Special?," *Journal of Business Ethics* 92: 373–84.
Narveson, J. (1992) "Democracy and Economic Rights," *Social Philosophy and Policy* 9: 29–61.
Nozick, R. (1974) *Anarchy, State, and Utopia*, New York: Basic Books.
Okin, S. (1989) *Justice, Gender, and the Family*, New York: Basic Books.
Pateman, C. (1970) *Participation and Democracy Theory*, Cambridge: Cambridge University Press.
—— (2002) "Self-Ownership and Property in the Person: Democratization and a Tale of Two Concepts," *Journal of Political Philosophy* 10: 20–53.
Pence, G. (2001) "Towards a Theory of Work," Philosophical Forum 10(1978–79) reprinted in *Philosophy and the Problems of Work: A Reader*, ed. Kory Schaff, Lanham, MD: Rowman and Littlefield: 93–106.
Phillips, R. and J. Margolis (1999) "Toward an Ethics of Organizations," *Business Ethics Quarterly* 9: 619–38.
Putterman, L. (1988) "The Firm as Association versus the Firm as Commodity," *Economics and Philosophy* 4: 243–66.
Radin, M. (1996) *Contested Commodities*, Cambridge, MA: Harvard University Press.
Rawls, J. (1999) *A Theory of Justice*, revised edn., Cambridge, MA: Belknap Press.
Roemer, J. (1982a) *A General Theory of Exploitation and Class*, Cambridge, MA: Harvard University Press.
—— (1982b) "Property Relations vs. Surplus Value in Marxian Exploitation," *Philosophy and Public Affairs* 11: 281–313.
—— (1982c) "New Directions in the Marxian Theory of Exploitation and Class," *Politics and Society* 11: 253–87.
—— (1988) *Free To Lose: An Introduction to Marxist Economic Philosophy*, Cambridge, MA: Harvard University Press.

Satz, D. (2010) *Why Some Things Should Not Be For Sale: The Moral Limits of Markets*, Oxford: Oxford University Press.
Schwartz, A. (1982) "Meaningful Work," *Ethics* 92: 634–46.
Smith, J. (2007) "Managerial Authority as Political Authority: A Retrospective Examination of Christopher McMahon's Authority and Democracy," *Journal of Business Ethics* 71: 335–8.
Tucker, R. (ed). (1978 [1972]) *The Marx-Engels Reader*, 2nd edn., New York: W.W. Norton.
van Oosterhout, J. (2007) "Authority and Democracy in Corporate Governance?" *Journal of Business Ethics* 71: 359–70.
Van Parijs, P. (1995) *Real Freedom for All: What (if Anything) Can Justify Capitalism?*, Oxford: Oxford University Press.
Walzer, M. (1983) *Spheres of Justice*, New York: Basic Books.
Werhane, P. (1985) *Persons, Rights, and Corporations*, Englewood Cliffs, NJ: Prentice-Hall.
Wood, A. (2004) *Karl Marx*, 2nd edn., New York: Routledge.
Young, I.M. (1979) "Self-Determination as Principle of Justice," *The Philosophical Forum* 11: 30–46.
—— (2006) "Taking the Basic Structure Seriously," *Perspectives on Politics* 4: 91–7.

66
PUNISHMENT

Mark R. Reiff

A surprisingly large number of questions arise under the heading "the philosophy of punishment." Here, in what roughly conforms to the intellectual order in which these questions might be addressed, although not necessarily in their order of importance, are some of them:

- *What is punishment?* Is it punishment to put someone in prison who prefers this to life outside? Is it punishment to put someone to death who wants to die? From whose point of view should these questions be decided? (Reiff 2005: 77–98, 116–141)
- *Whom may we properly punish?* Can we punish the mentally incompetent, or children? How do we tell whether someone has the capacity to control their actions or the actions of other people to a sufficient extent to be subject to punishment for them? (Hart 2008)
- *What is the relationship between punishment and responsibility, in both the causal and moral sense?* Can we impose punishment without fault? Is there such a thing as collective responsibility? As collective punishment? (Duff 1990, 2005, 2008; Feinberg 1970; Hart 2008; Hart and Honoré 1985; Reiff 2008)
- *What may we punish people for?* Only for criminal wrongs, or can we punish people for civil wrongs too? What considerations should be taken into account here? The offender's intent? The harm caused? The degree to which the harm caused can be compensated, if at all? (Duff 1996; Feinberg 1984; Hart 1963; Husak 2008; Reiff 2005)
- *How can punishment be morally justified?* Does the justification for the institution of punishment differ from the justification for a particular act of punishment? Does the justification for the threat of punishment differ from the justification for its actual imposition? Is punishment not only morally permitted, but (at least in some cases) morally required? (Acton 1969; Cottingham 1979; Duff 1986, 2001; Feinberg 1970; Rawls 1955; Simmons *et al.* 1995; von Hirsch 1993; Walker 1991, 1999)
- *How do we determine what kind and amount of punishment is due?* Are attempts to be punished as severely as completed crimes? Are there moral limits on the form and extent of punishment we may impose? Is capital punishment permissible? Torture? What is the role of forgiveness and mercy? (Davis 1986; Kramer 2011; Murphy and Hampton 1998; Reiff 2005; Shue 1978; Sussman 2005)
- *What defenses should be available to someone charged with a punishable offense?* What is the difference between a justification and an excuse? (Duff 2004)
- *Who has the right to impose punishment?* Is this the exclusive province of the state, or can punishment be imposed by private citizens as well? (Husak 1992; Locke 1988; Murphy 1985; Shafer-Landau 1996; Simmons 1992)

- *Who decides whether the accused is guilty and if so, how they should be punished?* By what process are these decisions to be made? What restrictions might apply to whatever decision-making process we employ? (Duff *et al.* 2004, 2006, 2007)
- *What is the role of punishment in international law?* Are the issues different here, or only the extent of the crimes? Are nations to be punished just like individuals, or is it wrong to impose hardships on people for the actions of their leaders? (May 2005, 2007, 2008)
- *What is the role of punishment in transitional justice?* Are actual prosecutions necessary, or are truth and reconciliation commissions sufficient? (Elster 2004)

Of course, many of these questions are intertwined in various ways, some could be derived from and therefore subsumed by others, and some have necessarily been left out. This list is not exhaustive, nor has it been reduced to only its most fundamental elements. Finally, the answer to some of these questions, especially to the question of justification, may heavily influence if not actually determine the answer to many of the others. Nevertheless, it is helpful to see the range of issues that can arise here before we begin to examine some of them in greater depth.

1. What is Punishment?

Surprisingly few theorists spend much time discussing this particular question. Most simply state that punishment is "hard treatment" that is threatened or imposed as a deliberate response to some threat or act of wrongdoing. But while this does capture something broadly true about many instances of what we would call "punishment," it is also misleadingly simplistic, for it disguises and possibly even prejudges what are, in fact, some important issues that arise when deciding what should be regarded as punishment and what should not. For example, is hard treatment of the wrongly convicted "punishment," or is hard treatment punishment only when it is inflicted on the truly guilty? Must hard treatment be deliberately imposed to constitute "punishment," or can punishment come about in other ways as well? And how do we decide what constitutes hard treatment anyway?

To avoid these pitfalls, the definition I propose is both broader and more neutral than this: punishment, in its most general sense, refers to an undesirable change in a person's well-being that is seen to result from, or is deliberately imposed as a response to, some act or omission committed by that person or by others for whom he is deemed responsible. Under this definition, punishment accordingly has two elements: an evaluative element, which determines whether the change in well-being is desirable or undesirable, and a relational element, which determines whether the connection between the change in well-being and the actual or potential act of wrongdoing is sufficient or insufficient. The former tells us whether a particular change in well-being is a burden or a benefit, and the latter tells us which burdens count as burdens of punishment and which do not. The question of whether it is possible to "punish" the innocent is left open, as is the question of whether all harsh consequences can constitute punishment or only those deliberately imposed by human agents. Most importantly, this definition makes clear that we cannot determine what punishment is until we have decided from whose point of view the question is to be determined.

The latter issue is especially important, for the relevant point of view seems to change depending upon what we see as the purpose of the putative instance of punishment we

have in mind. If we are thinking of punishment as a *deterrent*, for example, the evaluative and relational elements must be satisfied from the point of view of the potential violator. We cannot deter a potential violator by threatening to change his well-being in a way he views as desirable, and we cannot deter a potential violator by threatening to change his well-being in a way he views as undesirable if he does not also understand what behavior will activate the threatened change. If we are thinking of a punishment as *retribution*, however, as some sort of treatment a wrongdoer *deserves* to experience rather than as a threat designed to deter future wrongdoing, the point of view from which the evaluative and relational elements must be satisfied is less clear. Here it seems that the wrongdoer's personal conception of the good and peculiar notions of causation are not controlling, and are perhaps even irrelevant. Most of us think there is a difference between receiving a bouquet of flowers and a slap in the face or a slug from a .45, but most of us would not consider the former to be an act of retribution even if the person to whom the flowers were given had unusual reactive attitudes and experienced this as a burden rather than a benefit. Indeed, if we were to think otherwise we might be inclined to tailor our acts of retributive punishments to each individual wrongdoer's personal conception of the good, an enterprise that seems uncivilized, to say the least, and possibly barbaric. This was, after all, the entire purpose of Room 101, the infamous punishment room used by the totalitarian government of Big Brother to enforce submission to its will in George Orwell's novel *1984*. If we remain convinced that this kind of personalized punishment is not something a civilized society should support, then when it comes to retributive punishment, what seems to matter is whether the evaluative and relational elements are satisfied from the point of view of the victims of the wrongdoing, or perhaps from the point of the view of society as a whole, the semi-objective point of view of the proverbial "reasonable man," rather than the potentially idiosyncratic point of view of the wrongdoer himself.

Our determination of whether the relational element is satisfied is also likely to shift depending on the purpose of the purported instance of punishment we have in mind. If the purpose is deterrence, then the relational element is satisfied whenever the potential violator understands that the violation he is contemplating will provide a *justification* for the imposition of a negative change in well-being, as would be the case when the change comes about through imposition of a criminal legal remedy. But why should we treat only negative changes in well-being that are deliberately imposed by human agents in response to a violation as punishment? Surely the scope of changes in well-being that a potential violator will consider in deciding whether to commit a violation and which will therefore actually enter into his practical reasoning is far broader than this. Not only will a potential violator consider changes in well-being that are brought about as a deliberate response to wrongdoing, a potential violator will also consider certain changes in well-being that would be merely a *consequence* of the contemplated wrongdoing when deciding whether to commit this particular wrongful act. The change in well-being need not even be a *direct* consequence of the wrongful act—certain *indirect* consequences may be sufficiently connected to the wrong under consideration to figure into the potential wrongdoer's decision-making process as well. Typically, these indirect or secondary consequences will be the predictable follow-on effects of a particular change in well-being. Injuries to my person might cause me to suffer financial difficulties, which in turn could create difficulties in my personal life, and so on. But not every undesirable change in well-being that is part of the chain of events that begins with an act of wrongdoing will seem sufficiently connected to that act to be seen as

punishment for it. A potential wrongdoer may consider the possibility that he will crash into a tree in deciding whether to drive at an unsafe rate of speed, but he will probably not consider the possibility that he will be exposed to a serious disease if he is injured in the crash and sent to the hospital for treatment, or that his home could be more easily burgled if he is required to stay in the hospital overnight. The trick is distinguishing between the consequences a potential wrongdoer is likely to consider in determining whether to commit a wrong, even though they might be indirect, and those that are too exotic or remote to be taken into account.

If we are thinking of punishment as a form of retribution, however, we are likely to interpret the relational requirement more expansively. There are two reasons for this. First, we tend to discount the effect of future changes in well-being in deciding how to behave now, and thus the same negative change in well-being might not be sufficiently connected to the violation to be a deterrent but may be sufficiently connected to be retributive when it becomes an undiscounted present reality rather than a discounted future possibility. Second, we are likely to be willing to consider a much wider range of changes in well-being as being "caused" by a wrongful act when we are determining whether such a change is retributive than we would if we were merely looking at whether that change would function as a deterrent. And not only that, the more serious the violation, the more indirect the consequence may be and still be sufficiently related to the wrongful act to constitute punishment for it. Perhaps this explains our intuitive sense of what we sometimes call "divine" retribution. We might be unwilling to look very far along the chain of subsequent events for retributive punishment if we are dealing with a pickpocket; we might be willing to look very far indeed if we are dealing with a murderer.

2. What Forms Can Punishment Take?

There are many ways to categorize the forms that punishment can take. We can distinguish between punishments on the basis of their source, that is, whether they are imposed by private persons or the state, or on the basis of the particular aspect of well-being on which they have an effect, that is, whether they have an effect on our economic, physical, moral, social, spiritual, or emotional well-being, or we could distinguish between punishments on the basis of some distinctive characteristic that each exhibits. If we do the latter, however, we have to be careful, for if we try to have categories that both exhaust all the possibilities and do not overlap with one another we could end up with too many categories or too few. It is therefore important not to be too rigid about this, and to define our categories so that they capture something essential about each particular form of punishment without necessarily being too concerned if our categories overlap.

With that in mind, the forms that punishment can take can perhaps best be divided into six partially overlapping categories. These include physical sanctions, strategic sanctions, moral condemnation and regret, social criticism and the withdrawal of social cooperation, automatic sanctions, and of course traditional legal remedies such as fines and imprisonment. Most of these forms of sanctions will be familiar, but I will mention some examples. Spraying Mexican marijuana fields with the herbicide paraquat in the 1970s was not only a way of eradicating the crop, it was also way of discouraging users from smoking it, for they risked physical sanctions in the form of additional negative health effects if they did. While individual consumers often have little leverage over

businesses who deal with them, large customers can exercise substantial influence over their business partners by threatening to take their business elsewhere, a use of their strategic power that can ensure that those amenable to it will consistently deal with their business partners in good faith. The mere threat of moral condemnation and regret keeps a good portion of the ordinary population from violating the rights of others, and the threat of the withdrawal of social cooperation, such as would likely occur following a negative entry on a credit report, keeps most of us paying our credit card bills even though our balances are small enough in most cases that it is extremely unlikely that if we did not pay we would be sued.

With all these forms of punishment, the particular sanction imposed can be especially powerful it its imposition is automatic. Automatic sanctions are negative changes in well-being that arise out of wrongful conduct but do not require an intentional act by anyone in order to initiate. The most famous of these is perhaps the Doomsday Machine from the movie Dr Strangelove, but automatic sanctions would also include things like the injuries one risks when one drives on the wrong side of the road or otherwise drives recklessly, the dog bite one risks by attempting to burgle or otherwise trespass on a dog-owner's property, the "severe tire damage" that results from trying to exit a parking lot without paying, and so on. Finally, there are the traditional legal remedies of fines and imprisonment, although the actual category of legal remedies is much broader than this, for it would also include things such as self-help where permitted by law, various kinds of legal forfeitures, and the various penalties that can be associated with violating licensing laws. In any case, what is important to note is that legal remedies are only one form of sanction, and that punishment almost always involves a combination of sanctions, and need not include a remedy that requires some sort of judicial process at all.

3. How Can Punishment be Morally Justified?

The question of justification has been and remains one of the central questions in discussions of the philosophy of punishment, for even after thousands of years, punishment is still thought to be morally problematic. "It is morally problematic because it involves doing things to people that (when not described as 'punishment') seem morally wrong" (Duff and Garland 1994: 2) and, one might add, because even after all this time, there is still no consensus about what this justification might be, or even whether a viable justification exists. On the contrary, even after all these years, the number of competing justifications on offer is still increasing (see, e.g., Cottingham 1979; Duff 1986, 2001; Honderich 2006; Walker 1991, 1999). Rather than describe any particular proffered justification in depth, I will accordingly simply describe the general categories into which such justifications fall.

Traditionally, there have been three: consequentialist justifications, retributivist justifications, and justifications that somehow draw on both consequentialist and retributivist elements and are therefore considered "mixed." Consequentialists argue that punishment is justified if and only if its good effects, all-things-considered, outweigh its bad effects, either in this particular case (if one is an act consequentialist) or in a certain class of cases (if one is a rule consequentialist). These "good effects" include the incapacitation of the offender so he cannot offend again, at least not straight away, and perhaps even his rehabilitation, although continuously high recidivism rates suggest that rehabilitation is an unrealistically utopian goal. But the most important effect is usually claimed to be the general deterrence of crime, that is, crime by those other than

this particular offender. Of course, as Bentham pointed out, actual punishment is not necessary to produce this general deterrent effect; all that is required is that potential wrongdoers *believe* that punishment will be harsh, swift and certain, and this belief could be produced by simply publicizing punishments that had not in fact taken place, thereby achieving the morally desirable effect without having to pay the morally troubling cost (see Reiff 2005: 181–7). There is, accordingly, an indirectness to the main consequentialist justification for punishment that most consequentialists ignore, even if all their empirical claims are true. But the principal objection to the consequentialist justification is that even if punishment must actually be administered to have the requisite deterrent effect, it does not seem to be essential that those subject to such punishment actually be guilty of the offense for which that punishment is being imposed. Under the right circumstances, consequentialist reasoning would endorse punishing the innocent as well.

Non-consequentialists, of course, find this possibility very morally troubling (think, for example, of the version of the Kantian categorical imperative that admonishes us to never use someone as a mere means, but to always treat others as ends-in-themselves), and some consequentialists do as well. But non-consequentialists are also skeptical of the consequentialists' empirical claims, especially the claim that punishment has the alleged general deterrent effect, or at least the claim that it has this effect in all cases, and they are also skeptical that punishment would be unjustified if it does not. Accordingly, non-consequentialists reject the idea that an act of punishment can be justified by its effects. Instead, they try to articulate a justification for punishment that does not depend on empirical claims, although most non-consequentialists realize that somewhere lurking in the background of their seemingly non-consequentialist justifications is some sort of empirical claim, which makes it surprisingly difficult in many cases to decide into which category a proffered justification should be placed, lest every view simply be categorized as mixed. In light of this, some theorists have proposed rethinking the traditional categorization of punishment theories as either consequentialist, non-consequentialist, or mixed in favor of more informative divisions based on the extent to which the theory relies on empirical rather than conceptual claims (see Davis 2009). But the overriding goal of all non-consequentialist justifications is to show that for certain kinds of conduct, punishment is in some sense "deserved" by those who are subject to it. And, of course, the necessary corollary of this is that if someone is innocent, punishment is not deserved, and therefore it would be wrong to impose it, or rather treatment that is otherwise wrong would remain wrong for there would be no reason to disregard the wrongness of it.

Retributivists themselves can be divided into two different groups. Negative retributivists believe that punishment is morally permitted, while positive retributivists believe that punishment is morally required. The most famous example of this latter kind of retributivist is probably Kant, who said that

> even if civil society were to be dissolved by the consent of all its members … the last murderer remaining in prison would first have to be executed, so that each has done to him what his deeds deserve and blood guilt does not cling to the people for not having insisted on his punishment, for otherwise people can be regarded as collaborators in this public violation of justice.
>
> (Kant 1996: 106)

Although modern positive retributivists generally recognize that there is a limit to the amount of justice that any society can afford, and thus unlike Kant do not claim that we have an absolute, indefeasible duty to punish those who deserve it, even modern positive retributivists are less willing than their negative counterparts to recognize forgiveness and mercy as reasons for imposing less than the amount of punishment that is otherwise due.

Note, however, that there is another way that the negative–positive distinction can be cashed out. This is to say that positive retributivists believe that desert is both a necessary and a sufficient condition for just punishment, while negative retributivists believe that desert is a necessary but not a sufficient condition—some other (that is, consequentialist) ground must be shown before punishment will be just. Under this version of the distinction, negative retributivists can say that punishing the innocent is unjust, but they cannot say that punishing the guilty is just without referring to this other consequentialist justification. Obviously, these two ways of giving content to the negative–positive dichotomy are inconsistent with one another, for one could be a negative retributivist in the latter sense and yet be a positive retributivist in the former. So one has to be careful when one encounters or uses these terms to be sure one is clear about the sense in which the term is being employed.

Finally, there are some who argue that the practice of punishment cannot be justified at all, and the practice should accordingly be abolished. This "abolitionist" view can be anti-consequentialist, in that it might reject the idea that consequentialist moral reasoning can be used to justify anything. Or it could simply reject the idea that punishment has either instrumental or intrinsic value. It seems that in every generation there are a small number of (usually very vocal) theorists who press this view, and it has some serious contemporary advocates (see, e.g., Boonin 2008). It has never, however, managed to attract many followers.

One of the most popular retributivist views currently on offer is the communicative view. Under this view, the imposition of punishment is designed to open a moral dialogue between the offender and the community against which he has offended. While most retributivists tend to see punishment as something done *to* offenders, communicative theorists see punishment as a kind of process, something done *with* offenders, a process in which both the community and the offender are important participants. The purpose of punishment for this brand of retributivist is accordingly to communicate to the offender the censure that his or her crime deserves (Duff 1986, 2001). This is similar to expressive theories of punishment, which also justify punishment by reference to the expression it makes of society's disapproval of the offender's action (see, e.g., Feinberg 1965; von Hirsch 1993). It is also similar to educative theories of punishment, which aim not just at communicating society's censure, but at educating the offender so that he can understand why what he did was wrong and be in a position to rehabilitate himself and reform (see, e.g., Hampton 1984). But the communicative view goes beyond this because communication implies a two-way conversation, not a one-way lecture. "It aims not just to communicate censure but to thereby persuade offenders to repentance, self-reform, and reconciliation" (Duff 2001: xix). It is thus important for communicative theorists that the offender be involved in this process, that he actually *answer* for what he has done, that he be the *subject* of the process of punishment and not just its *object*.

In may be, however, that all these attempts at justifying punishment are somewhat misguided, for they look for a single, overriding purpose, when there are in fact two.

In my view, the justification for threatening or imposing punishment differs depending on whether one is thinking about the previolation or postviolation state of affairs (Reiff 2005). In the previolation state of affairs, the purpose of the threat of punishment is to facilitate social cooperation. To do this, the threat of punishment must be sufficient to make most people behave as if they believed their rights will not be violated, perhaps after taking some minor precautionary measures, such as locking one's doors. People will behave in such a manner, in turn, as long as it is rational to believe that the threat of punishment is sufficient to make most potential violators prefer to remain in the previolation state of affairs. The goal of issuing previolation threats of punishment is accordingly to produce such a rational belief throughout the general population.

In the postviolation states of affairs, in contrast, the purpose of punishment is to facilitate social conflict. But by this I do not mean make social conflict more common or violent—on the contrary, I mean that when social conflict occurs, postviolation punishment should be used to maximize the chances that the conflict will be resolved with the minimum amount of damage to the fabric of social cooperation. Toward this end, the measure of punishment due is the amount that will produce an amount of suffering in the wrongdoer that is roughly equivalent to the uncompensated suffering caused by the wrongdoing, subject to the applicable moral limits on the form and extent of suffering that may be imposed, for this is all that a reasonable people who have been subjected to wrongdoing can reasonably demand. Postviolation punishment under this view is accordingly deserved, but it is justified by the effect it has on those who are beneficiaries of the right that has been violated as well as the effect it has on the wrongdoer himself.

4. How Do We Determine the Form and Extent of Punishment that is Due?

As I said, postviolation punishment should be roughly equivalent to the uncompensated suffering caused by the wrongdoing. But I do not mean by this that the punishment should fit the crime in both nature and extent. Indeed, modern retributivists almost universally reject the biblical *lex talionis* view—the idea of an eye for an eye and a tooth for a tooth—and I am certainly not advocating a return to that. As Blackstone says, theft cannot be punished by theft, defamation by defamation, forgery by forgery, or adultery by adultery (Blackstone 1979: 13). And what do you do when a one-eyed man puts out the eye of a man with two eyes, or vice versa? In one case, an eye for an eye seems too severe a penalty; in the other it seems too slight. But almost no one rejects the idea that the punishment should fit the crime in some meaningful sense. Some base their view on an alternative theory after rejecting *lex talionis* outright; others argue that it is not necessary to take the biblical formulation so literally, and that a plausible interpretation of *lex talionis* is that it merely requires rough equivalence in any event (Waldron 1992). In either case, however, both groups agree that instead of trying to match the punishment and the crime so literally, punishment should be largely uniform, chosen from among a relatively small set of physical, financial, social, and moral penalties.

This avoids the problems that arise from aiming for strict identity, of course, but it raises problems of a different sort. If punishments must be chosen from among a relatively small set of changes in well-being, there will be many cases in which we will have to "translate" one type of injury into another, and this type of translation can be

difficult. The greater the difference between the injury caused by the wrongdoing and the injury to be inflicted as punishment, the more "incomparable" or "incommensurable" the two forms of injury might be, and the more difficult this translation process becomes. For example, translating certain injuries into monetary penalties might be relatively unproblematic, since monetary penalties can often be set by reference to market prices, but there is no market for imprisonment, and translating injuries into terms of imprisonment could be very difficult.

But most modern retributivists also deny that the suffering created by the wrongdoing is the proper object for comparison. Instead, the generally accepted view is that the relevant object of comparison is the "moral gravity" of the offense. This includes consideration of the harm caused by the wrong but is not necessarily dominated by it (see Hart 2008: 233–4). Intention is another factor—intentionally causing harm is morally more serious than causing the same harm negligently or accidentally. Motive also seems to be a factor. An intentional harm motivated by racial hatred is more serious than one motivated by jealousy. The identity of the victim is also relevant. The murder of a child is morally more serious than the murder of an old man even if the motives behind the crimes are identical. Exactly how all these elements interrelate, however, is not entirely clear (Nozick 1974: 59–63). The calculation of moral gravity accordingly has a somewhat mysterious, intuitive quality to it.

As a result, using moral gravity as one side of the comparative equation creates numerous practical difficulties that are difficult to overcome. Whatever incommensurability problems we would encounter by trying to translate many types of injuries into a small number of retributive punishments are multiplied many times over, for it is even more difficult to translate moral gravity (itself a compilation of difficult to compare considerations) into some amount of punishment than it is to translate one form of suffering into another. Indeed, while the consideration of moral gravity might work reasonably well for adjusting the *comparative* severity of punishments for a wide range of wrongs (robbery should be punished less severely than murder, attempts less severely than completed crimes), it is of little use when trying to decide where to start. No matter how morally serious the wrong, no absolute minimum amount of punishment can be derived from this fact alone; some additional criteria must be supplied. And while it might be possible to derive anchoring points for this ordinal scale by linking the most serious crime to the most serious punishment we are prepared to impose and working down from there (see von Hirsch 1993), the lack of any substantive connection between the form and extent of punishment and the particular crime makes this solution rather unsatisfying.

It is also not clear that putting "moral gravity" on the other side of the comparative equation ultimately leads to anything different than a comparison of suffering. All of the factors that contribute to moral gravity are factors precisely because they influence the overall suffering caused by the violation. Take two violations that are basically identical except for the level of intent involved. Even though the physical harm caused by two wrongs is similar, our normal responses to these wrongs might differ. The intentional violation is likely to be more emotionally disturbing than the negligent one, both to the immediate victim and to other members of society who learn of it (see Bentham 1996). For example, a murder is potentially socially volatile in a way a negligent killing is not, simply because of the greater range and depth of the suffering caused by the level of intent involved. If the motive behind the murder is especially pernicious (say it was motivated by racial hatred rather than greed or jealousy), it will be more disturb-

ing still, and its potential social volatility even greater. And if the victim is particularly defenseless, say a child, the ripples of emotional disturbance that travel through society will once again be even more profound and potentially more disruptive. In contrast, if the victim is particularly evil and unsympathetic, say a drug dealer or a pedophile, the emotional disturbance the murder causes in society at large might be rather limited indeed. In each case, whether we use some conception of moral gravity to determine the seriousness of a violation or use the extent of physical, psychological, and financial suffering the violation will inflict on society as a whole, the ordinal rankings we come up with are likely to be similar.

5. What Kind of Conduct Should Lead to Punishment?

The conventional view is that wrongs should be divided into two categories: the civil and the criminal. Civil wrongs are to be compensated; only criminal wrongs are to be punished. This, in turn, makes deciding whether punishment is due dependent on whether the wrong at issue is properly characterized as civil or criminal, and this in turn is usually seen as dependent on whether the wrong in question is sufficiently serious in terms of its broader societal effects to be the public's business and not merely a private matter. In response to this view, some theorists argue that the test should not be whether the wrong has significant societal effects, for many of the wrongs we currently accept as civil clearly have such effects and many wrongs we currently accept as criminal do not, or at least not always, have such effects. On the contrary, the real dividing line between the civil and the criminal—between whether compensation or punishment is due—should be determined by the intrinsic nature of the wrong itself. Of course, in deciding what intrinsic factors make a wrong more criminal than civil and what factors have the opposite effect, the same factors might come into play, so the real difference between these views could be merely semantic. In either case, we are still looking for some criterion that establishes the difference between the civil and the criminal (see generally Duff 2010; Husak 2008).

But there is another possibility here. Instead of focusing on whether a wrong is properly classified as civil or criminal, we might recognize that punishment as well as compensation can be due for any wrong, regardless of whether it is civil or criminal. What determines whether punishment is due is not the nature of the wrong, but whether the wrong causes incompensable suffering, for compensable suffering should be compensated, thereby minimizing the degree of morally questionable punishment that might be due. The only reason to categorize some wrongs as criminal rather than merely civil, then, is that in some cases the punishment that can be meted out for the incompensable suffering caused by a particular wrong—and there is always punishment that can be meted out even if we are talking about moral wrongs that are not even legal wrongs, much less criminal legal wrongs—is not sufficient to make the underlying right enforceable unless criminal legal sanctions are also made available (Reiff 2005). Of course, we still need to decide what kind of acts are wrongs and what are not, and this might indeed depend on the intrinsic nature of the act in question or on its societal effects or both, for not all acts that cause incompensable suffering are wrongs, but this has nothing to do with whether compensation or punishment is due. The latter question is *determined by* the need to enforce whatever rights we have; it is not a factor in *determining* what rights we have.

Related Topics

Utilitarianism and Consequentialism, Desert, Rights

References

Acton, H. B., ed. (1969) *The Philosophy of Punishment*, London: Macmillan.

Bentham, Jeremy (1996) *An Introduction to the Principles of Morals and Legislation*, ed. J. H. Burns and H. L. A. Hart, Oxford: Oxford University Press [1798].

Blackstone, William (1979) *Commentaries on the Laws of England: A Facsimile of the First Edition of 1765–1769*, Chicago, IL: University of Chicago Press.

Boonin, David (2008) *The Problem of Punishment*, Cambridge: Cambridge University Press.

Cottingham, John (1979) "Varieties of Retribution," *Philosophy Quarterly* 29: 238–246.

Davis, Michael (1986) "Harm and Retribution," *Philosophy and Public Affairs* 15: 236–266.

—— (2009) "Punishment Theory's Golden Half Century: A Survey of Developments from (about) 1957 to 2007," *The Journal of Ethics* 13: 73–100.

Duff, R. A. (1986) *Trials and Punishments*, Cambridge: Cambridge University Press.

—— (1990) *Intention, Agency and Criminal Liability: Philosophy of Action and the Criminal Law*, Oxford: Basil Blackwell.

—— (1996) *Criminal Attempts*, Oxford: Oxford University Press.

—— (2001) *Punishment, Communication, and Community*, Oxford: Oxford University Press.

—— (2004) "Rethinking Justifications," *Tulsa Law Review* 39: 829–850.

—— (2005) "Strict Liability, Legal Presumptions and the Presumption of Innocence," in A. P. Simester ed., *Appraising Strict Liability*, Oxford: Oxford University Press, pp. 125–149.

—— (2008) "Responsibility and Liability in Criminal Law," in Matthew H. Kramer, Claire Grant, and Ben Colburn, eds., *The Legacy of H. L. A. Hart: Legal, Political, and Moral Philosophy*, Oxford: Oxford University Press, pp. 101–120.

—— (2010) "Towards a Theory of the Criminal Law?" *Aristotelian Society Supplementary Volume* 84: 1–28.

—— and David Garland, eds. (1994) *A Reader on Punishment*, Oxford: Oxford University Press.

Duff, R.A., Lindsey Farmer, Sandra Marshall, and Victor Tadros, eds. (2004) *The Trial on Trial Volume I: Truth and Due Process*, London: Hart Publishing.

—— (2006) *The Trial on Trial Volume II: Judgment and Calling to Account*, London: Hart Publishing.

—— (2007) *The Trial on Trial Volume III: Towards a Normative Theory of the Criminal Trial*, Oxford: Hart Publishing.

Elster, Jon (2004) *Closing the Books: Transitional Justice in Historical Perspective*, Cambridge: Cambridge University Press.

Feinberg, Joel (1965) "The Expressive Function of Punishment," *The Monist* 49: 397–423.

—— (1970) *Doing and Deserving*, Princeton, NJ: Princeton University Press.

—— (1984) *Harm to Others*, Oxford: Oxford University Press.

Hampton, Jean (1984) "The Moral Education Theory of Punishment," *Philosophy and Public Affairs* 13: 208–238.

Hart, H. L. A. (1963) *Law, Liberty, and Morality*, Oxford: Oxford University Press.

—— (2008) *Punishment and Responsibility*, 2nd edn., Oxford: Oxford University Press.

—— and Tony Honoré (1985) *Causation in the Law*, 2nd edn., Oxford: Oxford University Press.

Honderich, Ted (2006) *Punishment: The Supposed Justifications Revisited*, London: Pluto Press.

Husak, Douglas (2008) *Overcriminalization: The Limits of the Criminal Law*, Oxford: Oxford University Press.

—— (1992) "Why Punish the Deserving?" *Nous* 26: 447–464.

Kant, Immanuel (1996) *The Metaphysics of Morals*, ed. Mary Gregor, Cambridge: Cambridge University Press.

Kramer, Matthew H. (2011) *The Death Penalty Redux: A Philosophical Investigation*, Oxford: Oxford University Press.

Locke, John (1988) "Second Treatise of Government," in *Two Treatises of Government*, ed. Peter Laslett, Cambridge: Cambridge University Press.

May, Larry (2005) *Crime Against Humanity: A Normative Account*, Cambridge: Cambridge University Press.

—— (2007) *War Crimes and Just War*, Cambridge: Cambridge University Press.

—— (2008) *Aggression and Crimes Against Peace*, Cambridge: Cambridge University Press.
Murphy, Jeffrey G. (1985) "Retributivism, Moral Education, and the Liberal State," *Criminal Justice Ethics* 4: 3–11.
—— (2004) *Getting Even: Forgiveness and its Limits*, Oxford: Oxford University Press.
—— and Jean Hampton (1998) *Forgiveness and Mercy*, Cambridge: Cambridge University Press.
Nozick, Robert (1974) *Anarchy, State, and Utopia*, Oxford: Blackwell.
Rawls, John (1955) "Two Concepts of Rules," *The Philosophical Review* 64: 3–32.
Reiff, Mark R. (2005) *Punishment, Compensation, and Law: A Theory of Enforceability*, Cambridge: Cambridge University Press.
—— (2008) "Terrorism, Retribution, and Collective Responsibility," *Social Theory and Practice* 34: 209–242.
Shafer-Landau, Russ (1996) "The Failure of Retributivism," *Philosophical Studies* 82: 289–316.
Shue, Henry (1978) "Torture," *Philosophy and Public Affairs* 7: 124–143.
Simmons, A. John (1992) "The Right to Punish," in *The Lockean Theory of Rights*, Princeton, NJ: Princeton University Press.
——, Marshall Cohen, Joshua Cohen, and Charles Beitz, eds. (1995) *Punishment: A Philosophy and Public Affairs Reader*, Princeton, NJ: Princeton University Press.
Sussman, David (2005) "What's Wrong with Torture?," *Philosophy and Public Affairs* 33: 1–33.
von Hirsch, Andrew (1993) *Censure and Sanctions*, Oxford: Oxford University Press.
Waldron, Jeremy (1992) "Lex Talionis," *Arizona Law Review* 34: 25–51.
Walker, Nigel (1991) *Why Punish?* Oxford: Oxford University Press.
—— (1999) "Even More Varieties of Retribution," *Philosophy* 74: 595–605.

Further Reading

R.A. Duff and David Garland, eds., *A Reader on Punishment* (Oxford: Oxford University Press, 1994) is the best collection of some of the most influential early works on punishment; R.A. Duff, *Punishment, Communication, and Community* (Oxford: Oxford University Press, 2001) is a key text by one of the most important contemporary writers on punishment; Joel Feinberg, *Harm to Others* (Oxford: Oxford University Press, 1984) is a key text on criminalization; H.L.A. Hart, *Punishment and Responsibility* (Oxford: Oxford University Press, 2nd edn. 2008) is a classic text by one of the world's foremost legal scholars; Mark R. Reiff, *Punishment, Compensation, and Law: A Theory of Enforceability* (Cambridge: Cambridge University Press, 2005) is a comprehensive study of the relationship between punishment, compensation, and enforceability.

67
TERRORISM

C. A. J. Coady

Contemporary world politics is weighed down by the fear of terrorism, as once (according to the boast of Marx and Engels) Europe was haunted by the specter of communism. That specter endured for most of the twentieth century and cast a shadow over much that was not strictly communist; similarly, anxiety about terrorism often stretches to quite diverse realities. It is therefore important to start a philosophical discussion of terrorism by seeking clarity about the best way to understand the term.

We should focus initially on the idea of "a terrorist act" rather than on "terrorism" or "terrorist" since it is likely that we apply the latter terms to projects, people or policies because of something distinctive about what is done. Among the scores of competing and overlapping definitions in the huge literature on terrorism, there is one feature that commonly recurs even if qualified and variously compounded by other features. This is the idea that terrorist acts are violent acts targeting innocent people: the outrage such acts create is often focused on their intentionally killing, or seriously injuring, those who shouldn't be attacked. This is one of the messages sent by calling such acts "indiscriminate"—they fail to discriminate properly their targets.

1. The Tactical Definition and its Competitors

So I will define a terrorist act as follows: "A political act, ordinarily committed by an organized group, in which severe violence is intentionally directed at non-combatants (or 'innocents') or their significant property." The term "terrorism" can then be defined as "the tactic or policy of engaging in terrorist acts." Let us call this a "tactical" definition. There is a debate about whether threats to do such acts should be included in the definition. There are plausible reasons for and against, so here it might be best not to include threats in the explicit definition, but regard them as impinging on the borders of the defined concept.

I call this a tactical definition because its focus is upon acts that are distinctive sorts of means towards political goals. Other definitions have quite different emphases. Some concentrate upon the type of agent that deploys political violence, others upon the characteristic effect aimed at by the actions. Call the former political "status definitions," and the latter "impact definitions." Some definitions combine both strands but they should be considered separately because they can stand in opposition.

Political status definitions typically claim that terrorism is the use of political violence by sub-state groups against the state, or sometimes, more rarely, merely against democratic states. Walter Laqueur, for instance, while disclaiming the need for definitions, insists that terrorism "is an attempt to destabilize democratic societies and to show that

their governments are impotent" (Laqueur 1986: 87). As a conceptual maneuver that would have the curious consequence that terrorism could never be directed against non-democratic regimes, an implication that would surely be absurd. But the usual political status definition, without the Laqueur restriction, does capture something about the common discourse on terrorism that concentrates heavily upon non-state groups and individuals using violent means to achieve objectives against state authorities. This is most plausibly seen, however, as a feature of the political climate in which the discourse frequently operates rather than as a definitional feature. If you are in a relatively stable society, violent threats to its order by groups inside or outside the polity will often have more salience than most uses of state violence against insiders or outsiders. This picture will look very different if the stability is maintained by brutal tyranny that the sub-groups are resisting. Outsiders seldom regarded the violence used by the "Arab Spring" revolutionaries in the 2011 upheavals in the Middle East as terrorism though the besieged dictators regularly denounced it as such. In such cases, sympathetic observers see the state's violence as suspect and at least open to the charge of terrorism. This illustrates the importance of determining what type of political violence counts as terrorist, not what agents are resorting to it.

The tactical definition has two major advantages over political status definitions. The first is that it gives a clear sense to the phrase "state terrorism" which will be an accurate description of the state's resort to violent attacks upon innocent citizens. If terrorism describes a tactic rather than an ideology or an anti-state orientation, then the way is clear to determine when a terrorist act has occurred without any presumption about what type of agent has performed it. It also makes it possible to criticize states for using terrorist methods not only against their own populations but against other states. So Neville Chamberlain plausibly characterized the state-authorized bombing of enemy civilian populations in World War II as "mere terrorism" (Fuller 1972: 280). A second advantage of the tactical approach is that it allows us to distinguish between the different sorts of violence used by revolutionary and insurgent groups, even if all are morally suspect. For unjust revolutions we can still distinguish between those that resort to terrorism and those that don't. For justified revolutions (if there are any) we can likewise distinguish between those that resort to terrorism and those that don't. If terrorism, tactically defined, is shown to be morally wrong, then we can give limited praise to revolutionaries who have no just cause but refrain from terrorism and, equally, we can criticize legitimate revolutionaries when they resort to terrorism. I have no enthusiasm for violent revolution, but if there can be just wars then there seems valid logical and moral space for a just revolution. Yet those definitions of terrorism, such as the political status versions, that make all sub-state violence against the state count as terrorism and, in addition, reject terrorism as immoral, must conclude that all revolutions are immoral. Not only does this seem implausible, but it is worth noting that such a condemnation would be a peculiar absurdity in the mouths of those whose currently legitimate state owes its origins to a bloody revolution that they naturally regard as justified.

Impact definitions, on the other hand, tend to expand rather than restrict the sorts of acts that can be terrorist. Just how far they extend them depends on how they incorporate aspects of the political status definition. If they insist that only creating some specific effect, usually fear, by sub-state agents can count, then they face the problems already noted, but some of them avoid this consequence by making no restriction on agents. The most expansive version of an impact definition is Robert Goodin's which counts as terrorist any act which is performed "with the intention of frightening people

for political advantage" (Goodin 2006: 156). This makes no reference to the act's being violent and so would group together as equally terrorist both the action of a concerned climate scientist delivering a factual speech which he rightly hopes will frighten his audience sufficiently to moderate their contributions to climate destruction, and the bombing of a busload of schoolchildren by a political activist who wants to bring about a change in military policy. Such a definition is clearly too remote from ordinary usage to be of practical or theoretical utility.

While insisting on this sort of utility, some tactical definitions make much of a specific type of impact. They include a strong emphasis on the intent to create fear and sometimes the intent to influence the behavior of a group other than that specifically attacked. So, Igor Primoratz, in a typical definition of this kind, defines terrorism as: "the deliberate use of violence, or threat of its use against innocent people, with the aim of intimidating some other people into a course of action they otherwise would not take" (Primoratz 2004: 24). (It should be noted that Primoratz does not restrict his definition to political acts. I use "political" in a broad sense to cover many religiously motivated acts that have an obvious political dimension, but to contrast with merely criminal attacks upon innocent people. Where criminal acts have a direct political dimension they would also qualify under the definition.) These are basically tactical definitions with an important impact dimension to them. They are not expansive, as Goodin's definition is, indeed in one respect they are restrictive since they would rule out as terrorist some acts that tactical definitions without the stress upon fear and indirect purposes would include. Such tactical definitions, for instance, would count as terrorist those acts that aim at creating anger rather than fear in order to provoke an overreaction, whereas the Primoratz-style definitions would have to say that such attacks upon non-combatants were not terrorist. This strikes me as distinctly counter-intuitive. The intended effect of seeking to influence a group other than those attacked is a different matter. On one interpretation, it must be the case since if the terrorist act is intended to kill all in group A, the political purposes it serves must be brought about by members of some other group B; this is so inherent in the circumstances of such an attack that it is hardly worth building into the meaning of "a terrorist act." On the other hand, if the act is not certain to kill all of group A, and the perpetrators know that, then it might well be that they are aiming to bring about the political results they seek through the agency of those in group A who survive. So the reference to seeking to influence a different group from the one attacked is either redundant or mistaken. Nonetheless, it will not, for the most part, matter much if it is included in the definition with the qualification "normally intended" and much the same is true of the fear impact. My own preference is for a definition that leaves as much as possible regarding terrorist intentions and motives open to empirical research.

2. The Search for "Distinctive Significance"

The issue of fear raises another question that has begun to receive closer attention recently. It seems related to definition but purports to go beyond it by asking what is the distinctive significance of terrorism. This takes us partly into the territory of the moral status of terrorism, which we will treat below, but it has sufficient affinity to definitional questions to deal with here. Jeremy Waldron (2004) has made a contribution to this enterprise as has, more recently, Samuel Scheffler (2006). Scheffler argues that it is possible to be relatively neutral about definitional matters and still pose his question. To do this, he chooses certain examples of politically oriented violence that people "would

not hesitate, prior to analysis, to classify as instances of terrorism" (Scheffler 2006: 2) and he seeks to analyze what is distinctively morally repellent about them. The answer is that such acts are committed "in order to induce fear or terror in others, with the aim of destabilizing or degrading (or threatening to destabilize or degrade) an existing social order" (Scheffler 2006: 3). These he calls "the standard cases" though curiously this nomenclature is not intended (he says) "to beg the very questions of definition that I said I would not be addressing." I call this curious because a selection of standard cases to the exclusion of what others would equally readily pre-theoretically classify "unhesitatingly" as terrorist acts is an obvious definitional move, even if it is definition by paradigm case rather than explication of necessary and sufficient conditions. Significantly, many of these other excluded cases are clear counter-examples to Scheffler's thesis, so he treats them as somehow peripheral instances of terrorism, or, for one set of cases, as not terrorism at all but "terror."

Scheffler's strategy excludes from consideration the numerous examples of (what would otherwise be) terrorist acts in which the aim of the perpetrators has nothing to do with "destabilizing or degrading an existing social order" such as terrorist acts (in the sense of the tactical definition) aimed at securing the release of political prisoners, or the removal of an occupying force, or the removal of an oppressive government. None of these need aim at the degrading or destabilizing of the existing social order to which the victims belong; the attackers need not care about that social order, other than for its influence on the specific grievances they hope to remedy. Of course the expression "destabilizing or degrading an existing social order" is open to unconstrained interpretations, and one can imagine political commentators who would view any removal of a government, or pressured release of political prisoners as "destabilizing or degrading an existing social order." I doubt, however, that Scheffler wants to count himself in that camp. Nor, as he admits, does a great deal of state terrorism have his favored aim, indeed much of it aims to bolster the existing social order, and, for this reason, he wants to call it "terror" rather than "terrorism" thereby exhibiting a previously disavowed interest in definition, though this time a purely stipulative one. In short, Scheffler's version of the moral significance of terrorism helps promote a distorted understanding of terrorism that could have serious consequences for counter-terrorist strategies.

3. Some Counter-examples and Clarifications

The tactical definition has been criticized for excluding clear cases of terrorist acts. Seumas Miller has argued, for instance, that targeted killings and violent torture by South African government agents of dissidents who have been using non-violent methods to overthrow that violently oppressive government would clearly be state terrorism though its targets are not innocent (Miller 2009: 38). Virginia Held cites the suicide-bombing attack upon the United States marine base in Beirut, Lebanon in 1983 which killed 241 American servicemen, and the attack on the *USS Cole* in Yemen in 2000 as cases of terrorism in which combatants were the targets (Held 2008: 55). As Stephen Nathanson has pointed out, Miller's South African example and a similar one concerning attacks upon those connected with the Union Carbide disaster in Bhopal in 1984 are significantly ambiguous on the question of innocence (Nathanson 2010: 51–5). If the South African dissidents are, from the government's perspective prosecuting an evil (however non-violently) then the government's violent response to them does not provide a counter-example to the tactical definition; it is clearly different from a gov-

ernment attack upon a crowd of black shoppers at a market, even though it is of course a great wrong since the government is mistaken in thinking that the protest is an evil of any sort. On the other hand, if we think that the non-violence of the demonstration makes it palpably obvious that the protesters are not in any way prosecuting an evil then we have to regard them as innocent, and again the tactical definition is intact.

Held's examples also seem to contain a tacit assumption of innocence. She thinks the fact that these attacks upon US military personnel are "routinely offered as examples of terrorism" shows the defects of the tactical definition, but, on the contrary, such routine reactions should be viewed with suspicion. It is understandable that any nation will regard "unprovoked" attacks upon their armed forces as outrageous and reach for the strongest terms of condemnation they can find. In the case of the United States, this is compounded by the "American exceptionalist" conviction that the nation offers the world a uniquely benign mission of security and civilization. In the Beirut case, the US "peace-keeping" troops were widely perceived by locals as actively favoring the Maronite Catholic faction in the Lebanon civil war and US naval gunfire had supported Lebanese army action. Those who see the US troops in their midst as foreign elements engaged in coercive violence or the standing threat of it, can, accepting the tactical definition, plausibly argue that attacks upon those troops are not terrorist. Those who see them, in spite of their uniforms, arms and deployment for foreign political purposes, as being in context non-combatants or innocents can rightly describe attacks upon them as terrorist and still accept the tactical definition. What is difficult is to admit their role in coercive deployment of violence but still insist that resistance to it is terrorist.

The discussion of these objections suggests the influence of an idea of terrorism as connected to *any* illegitimate deployments of political violence, or, in terms of just war theory, any violations of the *jus ad bellum* or the *jus in bello*. This would account for a tendency, possibly at work in the objections of Miller and Held, to classifying certain morally repellent acts of violence as terrorist that the tactical definition does not. Such a concept might be coherent, but it has the huge disadvantage that all wars and revolutions that failed the tests of the *jus ad bellum* would consist only of terrorist acts.

A final difficulty with the tactical definition concerns whether to read "intentional killing of the innocent" as "intending to kill those who are actually (objectively) innocent" or as "intending to kill those who are believed (subjectively) to be innocent." There are pressures both ways. In favor of the objective construal is the fact that we presumably don't want to say that the hijackers who killed the people in the Twin Towers in Manhattan and sacrificed the two planes' passengers were not terrorists because they mistakenly (and pretty weirdly) believed, as Osama bin Laden seems to have done, that all those people were combatants or somehow guilty. In favor of the subjective construal is the fact that if some people have an understandably mistaken belief that the people they kill are combatants when they are not, then we wouldn't want to call their actions terrorist, or them terrorists, for example, when the enemy force innocent villagers into military uniforms and push them within range of our fire. Perhaps it is best to adopt an objective construal for the most part, but allow that unavoidable mistakes about status count against attribution of "terrorist."

4. The Moral Assessment of Terrorism and the Issue of "Innocence"

What then of the morality of terrorism? Utilizing the tactical definition, we can question the morality of terrorist acts via the tradition of just war thinking with reference to

the *jus in bello* principle of discrimination. This has the advantage that the tradition is widely respected at least in theory in Western military academies and training, and has played a significant part in the formulation of the legal regulation of warfare in the UN conventions and other declarations at the basis of humanitarian law and international military law. This is not to claim that the tenets of just war theory are uncontentious. Recently there has been a great deal of philosophical work criticizing, revising and rejecting aspects of the theory, but most of this is itself a testimony to its intellectual vitality. In connection with the tactical definition, the major issue concerns the meaning of innocence and of non-combatants in the principle of discrimination prohibiting the intentional targeting of non-combatants.

There are three crucial questions about the status of non-combatants. Who are non-combatants, why do they have moral significance, and how far should their proposed immunity extend? These are questions that have recently been subjected to considerable philosophical scrutiny. First, who counts as a non-combatant? A rough criterion, and one that has been influential, is that a non-combatant is anyone who is not bearing arms. But this test is open to several objections. Soldiers seem to be prime candidates for combatant status but they are not always bearing arms, as when they are on leave or resting behind the front lines. Moreover, there are many people who play an important role in prosecuting a war who do not bear arms, but can be argued with some plausibility to be legitimate targets in war. These include those parts of the political leadership most involved in directing a war, scientists designing new weapon systems during a war, and workers in munitions factories. We might think of the soldiers on leave as part of the army that is engaged in fighting and as standing ready to resume hostilities. Equally, work in the munitions factory is surely integrated directly into the war effort in a way that can make the factory and its operators a legitimate target. The "civilian" politicians are hardly mere civilians when they scheme, plan and control violent hostilities from the safety of an office. Yet these modifications point towards abandoning the criterion of "bearing arms" in favor of something more explanatory that seems to underlie it. One move away from this is to cash the idea of non-combatant in terms of an underlying idea of innocence and combatant in terms of the lack of such innocence, but this plausible move gives rise to an important ambiguity. The move is plausible because it gives some moral significance to the classifications combatant/non-combatant (so answering our second question) since attacking those who are guilty of prosecuting some grave offence makes more moral sense than attacking those who are not. But some who invoke this move are insistent that they do not mean to rely upon an idea of moral innocence or guilt. For them "innocent" simply means "not harming" (Anscombe 1981: 67; Nagel 1979: 53–74, esp. 70). Here the idea is that the crucial thing about those we call combatants is that they are in the process of inflicting some grave harm so that whether they are fully morally responsible for doing so is beside the point. This, in turn, raises a question about why we can be entitled to direct lethal violence at those who are not fully morally responsible for the wrongs that they do. In the philosophical literature on war, these questions have been hotly debated recently and issues have been raised about the "innocence" of unjust combatants compared to just ones, about degrees of moral responsibility of coerced and/or ignorant soldiers, and about the non-innocence of large numbers of civilians. Much of this concerns what sort and degree of responsibility or liability, if any, would be required to make moral sense of the combatant/non-combatant distinction (Coady 2008b; Fabre 2009; McMahan 2009; Lazar 2009, 2010; Shue and other papers in Rodin and Shue 2008).

Part of the difficulty is caused by the fact that the idea of a combatant or "non-innocent" in the context of political violence is influenced by two different strands of thought: one is concerned with a type of individual agency and the other with a type of public role. If we concentrate on the former we will emphasize some more or less strict notion of moral or other responsibility, if we focus on the latter we will be concerned with classification under a public or political status. One theorist, Stephen Nathanson, has suggested resolving the difficulty by a conjunctive account whereby a person is only an "innocent" if they satisfy the conditions both of being morally innocent and of having non-occupancy of a military role. This has the advantage, in accordance with international law and what seems to be "common sense," that soldiers will still be legitimate targets who have been coerced into fighting or are non-culpably ignorant of the wrong they are doing and hence are arguably morally excused or even morally non-responsible for their actions. Ordinary citizens will not occupy the relevant role and will also not be *sufficiently* morally responsible for the killing in the war, and so will be "innocent" in the relevant sense.

Nathanson's suggestion rescues the ideas of "innocence" and non-combatancy from some of their intuitive problems, but it has its own costs. Most importantly, there seems to be insufficient philosophical rationale for the public role criterion if it is entirely independent of the agency criterion. Why should it matter morally which role a person occupies unless we know something about the moral significance of her acts in that role? The role of soldier, for instance, is not seriously specified by the sort of unusual clothing the soldier wears, but by what the soldier characteristically is engaged in doing qua soldier. This suggests that a concentration on some typical sort of intentional behavior or agency geared to the prosecution of harms must play some part in our understanding of combatant and, by contrast, non-combatant. (Just what sorts of agency conditions are relevant is a subject of great debate and some confusion. When considering liability to attack in war, Jeff McMahan argues (2009: 35) for the strict condition of moral responsibility for an objectively unjustified threat of harm. He considers various excuses as capable of removing blame but not full responsibility. Others differ about the necessity for, and nature of, moral responsibility.)

5. Justifications of Terrorism

The violation of the principle of discrimination counts against the moral justification of terrorism, but theorists who reject the principle can attempt to justify resort to terrorist acts on grounds of its promoting some significant good, and even those who accept it can argue that its violation can be legitimate in some circumstances. The former group includes many defenders of terrorism who insist that the goals pursued by some favored group are sufficient to justify attacks upon the innocent that will promote those goals. They take a purely instrumentalist approach to moral justification, but seldom make any serious effort to show how these tactics are likely to promote their ends, and the history of much terrorism seems to show a conspicuous lack of success in doing so. Most defenders of the Allied terrorist bombing of German cities in World War II, for instance, justify it as necessary to achieve victory efficiently over an evil enemy principally responsible for the horrors of the war. Yet many commentators argued then and later that it had no significant effect upon promoting that victory (Grayling 2006). There are so many terrorist campaigns that have been counter-productive: the Palestinian terrorist campaigns (not that all their attacks have been terrorist by the tactical

definition) seem to have achieved little for their plight, nor have Chechen or Basque terrorist attacks achieved their ends. It is sometimes hard to establish the results of terrorist attacks decisively since terrorists do not need to show that the terrorist acts alone have produced success: they can form part of a package of tactics and success or lack of it may involve a conjunction of tactics. Even so, a justification of terrorism should require some strong evidence that the resort to killing and maiming non-combatants is likely to make a significant contribution to success.

Some who argue instrumentally for terrorism seek to avoid morality and offer only "realist" arguments. The theoretical school of political realism is sometimes invoked as claiming that morality is irrelevant to warfare. These theorists are, however, more ambiguous on the role of morality in international politics than a merely amoralist interpretation would allow, and they often insist on "the national interest" as an overriding good that seems to have a moral flavor about it. Sub-state groups could adopt a similar goal relating to their primary interests. So understood, realism offers a particularly narrow theory of the morality of war, political violence and international relations more generally, and it is hard to motivate rationally the restrictive scope it offers. Why should the interests of some particular nation or group triumph over considerations of the human rights of innocent non-nationals, or a host of common moral decencies (see Beitz 1979; Coady 2008a: chs. 1 and 2).

The consequentialism inherent in this sort of realism is extended by utilitarian theory, which more plausibly does not restrict the goods pursued to the partiality of national or group interest. So utilitarianism has often been thought to permit terrorist acts since it seems possible that such acts might in some circumstances promote the greater overall good. Any appearance of easy permissiveness about the utilitarian approach, however, is misleading since terrorist acts face no simple task in the calculation of expected utility when it is combined with the altruistic component in utilitarianism. Even were terrorism shown to be effective in achieving certain valuable group ends, such as overthrow of an oppressive regime or victory in a just war, it would still have to be shown that this is not outweighed by other damaging effects of the terrorist campaign, such as the deaths of innocent people, the sufferings of their loved ones, the cast of mind created in the group resorting to terrorism and its consequences for their post-victory behavior, and the example set to countless others who face a temptation to terrorist acts though without the same prospects of success. A further point about utilitarianism is that it should not be automatically assumed that it must reject the principle of discrimination. There are many forms of utilitarianism and even simple act utilitarianism can adopt the principle as a "rule of thumb," but more profoundly a strong version of rule utilitarianism can endorse the principle as an absolute one and hence offer, as does Stephen Nathanson (2010: ch. 14), an unconditional rejection of terrorist acts.

6. Two Other Forms of Justification

Two other forms of justification need mentioning. One is that given by Virginia Held in terms of distributive justice and the other is that of "dirty hands." Held argues that in certain circumstances where there exist violations of fundamental human rights of one group and not of another it might be justifiable to redistribute rights violations so that the previously immune group members suffer them. Assuming that terrorism involves relevant human rights violations, this will license acts of terrorism on grounds of justice. The basic scenario is a society S_1 in which group A enjoys the satisfaction of

human rights both morally and legally but group B does not. This is contrasted with a society S_2 where both groups equally enjoy respect for such rights. Held's argument is that, given certain empirical assumptions about the effectiveness of terrorism in moving from S_1 to S_2 and the unavailability of non-terrorist means to that end, distributive justice allows members of group B to use terrorism against members of group A. Held is skeptical of definitions of terrorism but her conception of terrorist acts seems given by a disjunctive condition which she offers as "perhaps" sufficient to turn political violence into terrorism, namely, either the intentional harming of non-combatants or the intentional spreading of fear (Held 2008: 76). Assuming that her justification extends to acts covered by the first disjunct, then her distributive account implies that the intentional killing of a sufficient number of babies or young children in group A is required by, or at least in accord with, justice. Leaving aside the problem of how such killings could promote a transition from S_1 to S_2, the proposal seems to misconstrue the nature of distributive justice which is concerned with the fair distribution of goods not harms. Consider her example of a redistribution that she sees as parallel to her proposal, namely, the rectifying of unequal provision of police services in a community. Where a society dominated by group A provides very poor police arrangements to establish security and personal safety rights for members of group B, there is clearly a case in justice for redistribution. But the case requires redistribution of protective police services, not directly redistributing deaths and injuries. Held thinks of distributing police more fairly with the consequence that there will be more deaths/injuries/burglaries in group A and fewer in group B. This may well be a regrettable but required consequence of a fair distribution of protective services. But it is still protection that is being distributed, not killings and maimings per se. What the example obscures is that the terrorism proposal is aimed at licensing agents to intentionally commit murders by way of distribution.

A second form of justification is that of supreme emergency. In his book *Just and Unjust Wars*, Michael Walzer (2006 [1977]) invoked the category to justify certain acts of terrorism by states. As he makes clear in later writings (Walzer 2004: 45–6) supreme emergency is the name for situations that require agents to get "dirty hands" by violating some deep, otherwise absolute moral constraint. The acts of terrorism Walzer justifies are the Allied bombings of German cities in the early phases of World War II that intentionally killed many thousands of German civilians. Walzer is clear that these (and the later city bombings) were acts of terrorism and that terrorist acts are totally immoral, but he believes that sometimes in extreme circumstances, necessity, especially political necessity, licenses such acts to avoid disasters. Of course, Walzer is not unique in thinking that deep moral prohibitions can be overridden by weighty considerations about dire consequences of maintaining them or even by conflicting duties. It is rare to find contemporary moral philosophers of any stamp who accept even the strongest moral constraints as absolute, beyond exception. So, the peculiar situation exists that most ordinary folk think that terrorism can never be morally justified and most moral philosophers think it can. But Walzer's view is even more unusual because he thinks that terrorist acts remain deeply immoral in supreme emergency cases but their perpetrators are right to do them. Other deontologists, such as W.D. Ross and his intuitionist followers, would see the acts as morally right though the necessity to do them would be regrettable (Ross 1930: 28). Walzer's view is paradoxical because the acts are both right and wrong. His position is nearer to those who allow for strong moral dilemmas in ethics where morality cannot deliver a verdict because some act is both right and wrong, but the difference is that for Walzer there *is* a verdict because the necessary act must be done.

What to make of all this is the subject for another entry, but we must note that Walzer is reluctant to draw the dirty hands conclusion in the case of sub-state terrorism. His emergency justifications for state terrorism seem equally applicable to sub-state agents faced with supreme emergencies, but in a biting critique of excuses and justifications for sub-state terrorism Walzer, initially at least, leaves no room for their recourse to the dirty hands defense (Walzer 1988). When this essay was republished in his collection *Arguing About War* (2004), he added a parenthetical paragraph admitting the bare possibility that the defense could be available to sub-state terrorists but only if the oppression they were responding to was "genocidal in character" and he claimed that no such actual cases had occurred. He construes (2004: 54) this as involving "an imminent threat of political and physical extinction." Elsewhere he says (2004: 49) that supreme emergencies threaten "either the elimination of the people or the coercive transformation of their way of life." But if the second clause is what "genocidal" means then quite a few sub-state groups may plausibly appeal to it. If we are going to make exceptions for terrorist acts then we should be consistent about it.

Finally, those who argue that the moral prohibition on terrorist acts is defeasible may nonetheless hold that there should be an absolute legal prohibition of such acts. This is the position of McMahan (2009: 234–5). It is parallel to the position that euthanasia might be morally permissible in some circumstances but should always be illegal. Even if a supreme emergency story has plausibility in very specific circumstances, it would be unworkable and probably disastrous if given some general applicability in legislation. A problem with this is that something of the same generality with the same problems may be involved in its moral permissibility. This is what seems to lead Nathanson to support solely by rule utilitarian arguments an absolutist moral prohibition on terrorist acts. My own view, argued elsewhere (Coady 2008b), is that there are good non-utilitarian arguments for an absolute moral prohibition on terrorism.

Related Topics

Utilitarianism and Consequentialism, Justice and Borders, War

References

Anscombe, G.E.M. (1981) "Mr. Truman's Degree," in her *Collected Philosophical Papers of G.E.M. Anscombe, vol. iii: Ethics, Religion and Politics*, Oxford: Basil Blackwell.

Beitz, Charles R. (1979) *Political Theory and International Relations*, Princeton, NJ: Princeton University Press.

Coady, C.A.J. (2008a) *Messy Morality*, Oxford: Oxford University Press.

—— (2008b) *Morality and Political Violence*, Cambridge: Cambridge University Press.

Fabre, C. (2009) "Guns, Food, and Liability to Attack in War," *Ethics* 120, 36–63.

Fuller, J.F.C. (1972) *The Conduct of War, 1789–1961*, London: Eyre Methuen.

Goodin, R. (2006) *What's Wrong with Terrorism?*, Cambridge: Polity Press.

Grayling, A.C. (2006) *Among the Dead Cities: Was the Allied Bombing of Civilians in WWII a Necessity or a Crime?* London: Bloomsbury.

Held, V. (2008) *How Terrorism is Wrong: Morality and Political Violence*, New York: Oxford University Press.

Laqueur, W. (1986) "Reflections on Terrorism," *Foreign Affairs* 65, 86–100.

Lazar, S. (2009) "Responsibility, Risk, and Killing in Self-Defense," *Ethics* 119, 699–728.

—— (2010) "The Responsibility Dilemma for *Killing in War*: A Review Essay," *Philosophy and Public Affairs* 38, 180–213.

McMahan, J. (2009) *Killing in War*, Oxford: Oxford University Press.
Miller, S. (2009) *Terrorism and Counter Terrorism: Ethics and Liberal Democracy*, Oxford: Blackwell.
Nagel, T. (1979) "War and Massacre," in Nagel, T. *Mortal Questions*, Cambridge: Cambridge University Press.
Nathanson, S. (2010) *Terrorism and the Ethics of War*, Cambridge: Cambridge University Press.
Primoratz, I. (2004) "What is Terrorism?" in Primoratz, I. (ed.) *Terrorism: The Philosophical Issues*, Hampshire: Palgrave Macmillan.
Rodin, R. and Shue, H. (eds.) (2008) *Just and Unjust Warriors*, Oxford: Oxford University Press.
Ross, W.D. (1930) *The Right and the Good*, Oxford: Clarendon Press.
Scheffler, S. (2006) "Is Terrorism Morally Distinctive?" *Journal of Political Philosophy* 1–17.
Waldron, J. (2004) "Terrorism and the Uses of Terror," *Journal of Ethics* 8, 5–35.
Walzer, M. (1988) "Terrorism: A Critique of Excuses," in Luper-Foy, S. (ed.) *Problems of International Justice*, Boulder, CO: Westview.
—— (2004) *Arguing About War*, New Haven, CT: Yale University Press.
—— (2006 [1977]) *Just and Unjust Wars: A Moral Argument with Historical Illustrations*, New York: Basic Books.

Further Reading

Coghlan, T., "Al-Qaeda as strong as ever, says CIA man who trailed bin Laden," *The Times*, May 27, 2011. This article argues that Western misunderstanding of the nature of al-Qaeda and of its terrorist agenda is crippling the military counter-terrorism measures against that organization.
Meggle, G. (ed.), *Ethics of Terrorism & Counter Terrorism*, Ontos, Frankfurt am Main, 2005. A collection of essays from English-speaking philosophers and Continental theorists.
Primoratz, I. (ed.), *Terrorism: The Philosophical Issues*, Hampshire: Palgrave Macmillan, 2004. A comprehensive collection of essays by leading figures in the field.
Steinhoff, U., *On the Ethics of War and Terrorism*, Oxford: Oxford University Press, 2007. A provocative attack on some staples of just war theory and a defense of terrorist attacks on some who would normally be considered innocent.

68

PATERNALISM, MORALISM, AND MARKETS

Mark D. White

According to political philosopher Gerald Dworkin (2010), paternalism is "the interference of a state or an individual with another person, against their will, and defended or motivated by a claim that the person interfered with will be better off or protected from harm." While one individual can certainly act paternalistically toward another, such as when a person takes the car keys from a drunken friend to keep him from driving while inebriated, state paternalism receives much more attention because of the state's ability to establish and enforce coercive policies and regulations, and so it is on state paternalism that this chapter focuses.

Paternalistic laws and policies can take many forms, including prohibitions (such as on the use of illegal drugs) and mandates (such as regarding motorcycle helmets), both meant to directly control behavior; excise taxes (such as on tobacco) or subsidies (such as on tobacco cessation programs), both meant to influence behavior through changing incentives; and, most subtly, so-called "nudges," policies that rearrange the structure of options in a choice situation, taking advantage of common cognitive dysfunctions to guide the decision-maker toward certain choices. All of these types of paternalistic regulation involve "a usurpation of decision-making, either by preventing people from doing what they have decided or by interfering with the way in which they arrive at their decisions" (G. Dworkin 1988: 123).

As we shall see, standard critiques of paternalism focus on its effects on the autonomy and well-being of the persons whose choices are blocked or manipulated. While the ethical problems with this interference, and possible justifications for it, have been widely discussed, what remains relatively neglected is the motivation of paternalism, the state's concern for the well-being of the regulated, which I will argue is an inappropriate attitude for the state to take toward its citizens, and also necessarily involves the substitution of outsiders' values and interests for those of the regulated, which defeats the stated purpose of paternalism. Furthermore, regardless of the nature of the paternalistic regulation or prohibition—whether it aims at benefitting the citizen's "moral health"—this value substitution is inherently moralistic because of the presumption of superior knowledge and judgment on the part of regulators or legislators. I will conclude by arguing that the market, representing the maximal degree of free choice for each

person consistent with all others enjoying the same, provides the only guarantee of respect for dignity and protection against paternalistic value substitutions.

1. The Nature of Paternalism and Standard Critiques

As Joel Feinberg notes in *Harm to Self* (1986), the concept of paternalism is difficult to discuss analytically because of its widespread negative connotation; merely the accusation that a law is paternalistic clouds rational discussion of its merits. People may favor paternalistic policies, especially when they mostly affect *other* people, but they normally do not want them associated with paternalistic motives. But why? After all, many scholars claim that government has a duty to "care" for its citizens; Ronald Dworkin (2000: 1) takes "equal concern [to be] the sovereign virtue of political community." If this purpose for the state is accepted, then paternalism would be treated as a natural part of governance. But care for the well-being of the citizens can often conflict with respect for their autonomy and choices, which can be considered the core issue with paternalism: regulators might be well-intentioned, but they fail to respect citizens as responsible, autonomous agents, and they also dismiss their own ignorance with respect to people's true interests.

In his review of several treatments of paternalism, Dan Brock (1988: 551) describes this conflict between two values:

> One value is that of autonomy or self-determination, the interest persons have in making significant choices about their lives for themselves and in pursuing the courses chosen without interference from others. The other value is individual well-being or good which paternalistic action seeks to protect or promote. These values will be in conflict when the action chosen by the subject appears to be contrary to his well-being or good. The potential paternalist's alternatives then are either to respect the subject's autonomy and not protect his well-being or to infringe his autonomy by interfering to protect his well-being. So understood, the issue of paternalism requires a determination in any particular case of which value—autonomy or well—being is more important or weighty.

At risk of oversimplification, autonomy and self-determination represent a deontological point of view, deriving from the moral philosophy of Immanuel Kant or the John Stuart Mill of *On Liberty*, while the concern for well-being is based on the simple utilitarianism of Jeremy Bentham (rather than the more elaborate version of Mill's). (Below I will state this distinction in yet another way, representing the conflict between care and respect.)

Indeed, the seminal critique of paternalism comes from Mill (1859: 10–11) in the form of his famous harm principle, which mandates that government power must only be used to protect persons from each other, not from themselves:

> [T]he only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others. His own good, either physical or moral, is not a sufficient warrant. He cannot rightfully be compelled to do or forbear because it will be better for him to do

> so, because it will make him happier, because, in the opinions of others, to do so would be wise, or even right. … The only part of the conduct of any one, for which he is amenable to society, is that which concerns others. In the part which merely concerns himself, his independence is, of right, absolute. Over himself, over his own body and mind, the individual is sovereign.

To Mill, paternalistic action is only justified if the person being influenced is not acting voluntarily or in full knowledge of her actions. For instance, he argues (1859: 89) that

> if either a public officer or any one else saw a person attempting to cross a bridge which had been ascertained to be unsafe, and there were no time to warn him of his danger, they might seize him and turn him back, without any real infringement of his liberty.

In modern times this has been termed *soft paternalism*, as opposed to *hard paternalism*, which justifies interference with even voluntary, "open-eyed," rational action which is believed to be against the agent's interests (Feinberg 1971). Even among advocates of (only) soft paternalism, there are disagreements over the meaning of voluntary and how to judge it; we will return to this below. For the purposes of this chapter, by paternalism we will be referring to hard paternalism, since this is the more contentious meaning of the term.

As mentioned above, the concept of autonomous self-determination can be traced further back to Immanuel Kant (1785), who maintained that every rational being is capable of autonomous choice, that is, choice according to the moral law of his or her own determination, even in the face of opposing influence from external authority or internal desires and inclinations. This capacity of autonomous choice grants every person a dignity, an intrinsic value, incomparable and incalculable, which demands respect from all other persons by virtue of it. Extended to the political realm, this dignity (and the autonomy on which it is based) also requires respect from the state, and therefore official coercion is justified only to protect citizens from coercion from each other:

> Resistance that counteracts the hindering of an effect promotes this effect and is consistent with it. Now whatever is wrong is a hindrance to freedom in accordance with universal laws. But coercion is a hindrance or resistance to freedom. Therefore, if a certain use of freedom is itself a hindrance to freedom in accordance with universal laws (i.e., wrong), coercion that is opposed to this (as a *hindering of a hindrance to freedom*) is consistent with freedom in accordance with universal laws, that is, it is right.
>
> (Kant 1797: 231)

By extension, state coercion is never to be used to protect citizens from themselves:

> A government that was established on the principle of regarding the welfare of the people in the same way that a father regards his children's welfare, i.e., a *paternal government*—where the subjects, like *immature* children unable to distinguish between what is truly useful or harmful to them … such a govern-

> ment is the worst *despotism* we can think of (a constitution that subverts all the freedom of the subjects, who would have no freedom whatsoever).
>
> (Kant 1793: 290–1)

The deleterious effect of paternalism on personal autonomy is clear, but it is less clear how paternalism which is aimed at increasing the well-being of individuals might actually harm it. In addition to his autonomy-based argument, Mill (1859: 78) also makes a utilitarian argument against paternalism, based on the superiority of self-knowledge:

> [T]he strongest of all the arguments against the interference of the public with purely personal conduct is that, when it does interfere, the odds are that it interferes wrongly, and in the wrong place. On questions of social morality, of duty to others, the opinion of the public, that is, of an overruling majority, though often wrong, is likely to be still oftener right; because on such questions they are only required to judge of their own interests; of the manner in which some mode of conduct, if allowed to be practiced, would affect themselves. But the opinion of a similar majority, imposed as a law on the minority, on questions of self-regarding conduct, is quite as likely to be wrong as right; for in these cases public opinion means, at the best, some people's opinion of what is good or bad for other people.

This sentiment is echoed by critics of so-called libertarian paternalism (discussed below); for instance, Claire Hill (2007: 448) argues that paternalistic regulators

> sometimes speak as though they have access to the knowledge of what people really want apart from what they choose. This position is ultimately untenable. … As convenient and tempting as it may be to extrapolate from our own introspection that others want what we do, or should, want, we simply have no access to others' beliefs and desires.

And if we purport to measure individuals' well-being according to their own standards or preferences, paternalism will fail to maximize that measure of well-being if laws and regulations are determined by a different set of standards. But aside from the utilitarian aspect, presuming to increase a person's well-being according to standards or values other than her own has ethical implications for her dignity, which we discuss in the next section.

2. Interests, Dignity, and Value Substitutions

As we saw above, according to soft paternalists, the only justification for interference with a person's choice out of concern for his or her interest is when that choice is believed not to be sufficiently voluntary. Feinberg (1986: 115) defines voluntary choice as present when

> the chooser is "competent" … he does not choose under coercion or duress … he does not choose because of more subtle manipulation … he does not choose because of ignorance or mistaken belief … he does not choose in circumstances that are temporarily distorting.

Note that these standards are formal or procedural, assessing the decision-making process rather than the outcome of it. In other words, if a person satisfies Feinberg's conditions (or similar ones proposed by others), then his action is deemed to be voluntary; but if he fails any of these conditions, then the state can legitimately interfere because his choice is not his own.

Another problem with paternalism is that it sets aside the agent's own interests in favor of an external set of values judged to represent them. I refer to this aspect of paternalism as *value substitution*, with which there are two distinct but interrelated problems, one epistemic and the other ethical. The epistemic problem, discussed above, is that absent an expression of actual choice or consent, it is impossible for regulators to know the true interests of the persons whose behavior they are regulating. This implies that, even absent any concerns regarding respect or autonomy, regulators are not as well-positioned to advance the actual interests and well-being of persons as are the persons themselves, and any interests on which official policy is made are externally imposed on the persons being regulated. The ethical problem is introduced when regulators take the step from presuming to know the true interests of persons to actively influencing their decisions in promotion of those presumed interests. Both problems betray a lack of respect for persons' own deeply held values and preferences—which together I call their "interests"—as well as their ability (and their right) to act in pursuit of them (consistent with all else doing the same, which is to say within the boundaries of ethics and justice).

On what basis do regulators substitute their own presumptions regarding persons' interests for their actual ones? Implicit in any such regulation is the belief that persons do not always make decisions in their own true interests. Of course, persons often recognize this themselves, especially when self-control problems are involved: people sign up for automatic savings programs, weight control centers, addiction support groups, and so forth, to counter their own impulsive or weak-willed action. But these persons are seeking help voluntarily and in recognition of their true interests along with their personal problems, and it is their choice to seek help that reveals these true interests. In other words, a person can recognize (or not) that she is not making decisions in her own interests, but there is no way for third parties to observe her behavior and make that determination absent knowledge of her true interests (or formal evidence of nonvoluntary choice). In the interest of maintaining respect for dignity and autonomy, the presumption should always be that a person makes choices in his or her own interests, unless that person expresses otherwise (through words or action). Since it is each individual who knows if his or her decision is poor, it is up to each individual to seek help for this problem—and if the state wants to offer help for those who voluntarily seek it, that raises no problems with paternalism (though the availability of private alternatives raises different questions of state involvement). However, if the state imposes a "solution" on the populace as a whole to "help" select citizens overcome their bad judgment or weak wills, we risk a paternalistic "tyranny of the majority" in which some citizens deign to interfere wrongfully with the autonomous, personal choices of all.

Regulators might honestly believe that they have persons' true interests at heart, but since this information is unavailable to them absent explicit expression—which they implicitly reject by choosing paternalistic intervention anyway—they necessarily substitute their own interests for persons' own, in the belief that the interests presumed by the regulators are "obvious." Examples abound: clearly, according to the current state

of medical science, persons who choose to consume foods loaded with trans fats are not making choices solely in the interest of their long-term health. Drivers who eschew seat belts or helmets are not making choices solely in the interest of their safety, and those who drive above the optimal, fuel-efficient speed are not choosing their velocity purely according to their pocketbooks. And new employees who do not enroll in a 401(k) retirement plan from the first day of their employment are not making choices solely in the interest of their long-term financial security. But who is to say these are the interests in which people *must* make decisions? As Hill (2007: 450) argues,

> what is a better guide than people's choices? Even if people may really want something else, what might that be, and on what grounds can we claim we have access to it that gives us a better claim on what they are going to do than what they otherwise would choose?

Persons' actual interests are often much more complex and nuanced than simply "health" or "wealth." A person might consume a doughnut loaded with trans fats to bring back fond memories of eating doughnuts with her beloved grandfather who recently passed—or she simply likes the taste despite the adverse health consequences. A new employee might delay enrolling in a 401(k) plan because he is saving money for a down payment on a home, sending money to his parents, donating to a charity devoted to stopping hunger in Third World countries—or spending it on frivolous, hedonistic pursuits. And a motorcyclist might choose to ride without a helmet simply because she likes the feel of the wind through her hair. In the spirit of Mill, as long as such decisions do not cause harm to other persons, and they are not based on any procedurally flawed decision-making process, they should be protected and respected, not manipulated in a third party's idea of persons' true interests.

On what do regulators base their presumptions regarding persons' interests? Paternalistic regulations are not tailored to each individual, so regulators make assumptions regarding what the average person would reasonably hold as interests: wealth, health, safety, and so forth. It does not take a cynic to suspect that regulators choose the interests that they promote out of a sincere conviction that they are the "proper" interests in which persons "should" make decisions. They do not want persons eating unhealthy foods "just" because they taste good, or spending money that "should" be devoted to their retirement. But it is very difficult to justify this presumption on the part of regulators, much less actions taken to guide persons to make "better" decisions in pursuit of the "approved" interests. Of course, this is not to say that there is not usually some obvious prudence in holding wealth, health, and safety as interests, or that we cannot attempt to persuade persons whom we see making poor decisions (in our judgment) that there might be a better choice available. But if we respect persons as autonomous, rational agents, then we have no right to influence their decisions unduly—after our best attempts at persuasion, persons' choices are their own, and the consequences are theirs to bear. (We will have more to say on consequences, and responsibility for them, below.)

Some supporters of paternalism argue that regulation will guide people into making decisions that are in their best interests, or satisfy their preferences, if these concepts are taken to represent a person's "ideal" or "rational" interests. (See Griffin (1986: 11–16) on rational preferences.) For example, the mandatory use of seatbelts in automobiles might be justified by saying that if drivers or passengers thought about it, they would

agree that they should wear seatbelts, and the laws protect them from irrational carelessness or forgetfulness. Similarly, they assume that people would not eat fatty foods if they "realized" their true interests in health. But this is the same issue as before: any external judgment that a person's revealed preferences or interests are irrational reflects the interests of the regulator. Whether we call them rational, informed, or real, these "preferences" are artificially and arbitrarily constructed by someone other than the agent herself and cannot be held to represent her true interests. Furthermore, any policymaker's judgment about what should comprise an agent's well-being necessarily involves the preferences of the policymaker himself.

This is the approach often taken by supporters of behavioral law and economics and "libertarian paternalism," popularized by Richard Thaler and Cass Sunstein in their bestselling book *Nudge* (2008), who recommend "nudging" persons into making the choices they "would have made" were they not subject to cognitive biases, weakness of will, or sheer laziness. An example is automatic enrollment of new employees in 401(k) retirement plans with an option to withdraw, rather than nonenrollment with an option to enroll, based on the assumption that it is the new employee's true interest to enroll, but she might irrationally choose not to. And Sunstein and Thaler (2003: 1172–73) write,

> If employers think (correctly, we believe) that most employees would prefer to join the 401(k) plan if they took the time to think about it … then by choosing automatic enrollment, they are acting paternalistically by our definition of the term … steer[ing] employees' choices in directions that will, in the view of employers, promote employees' welfare.

Similarly, J. D. Trout (2005: 433) writes that paternalistic intervention designed to counter the effect of cognitive biases "promotes the agent's autonomy by intervening when the agent's decision is not one that, if fully informed and cognitively unbiased, the agent would have made." But Robert Sugden (2008: 232) asks,

> [h]ow, without making normative judgements, do we determine what counts as complete information, unlimited cognition, or complete willpower? Even if we can specify what it would mean to have these supernatural powers, how do we discover how some ordinary human being would act if he were somehow to acquire them?

Ultimately, libertarian paternalism is no different from any other type of paternalism which, as in Gerald Dworkin's definition, interferes with choices to promote a person's own good, and it suffers from the same ethical and epistemic problems (White 2011b: 180–94).

3. Moralisms Old and New

There are two ways in which paternalistic can be regarded as moralistic. The most obvious way, and the one most discussed in the literature on paternalism and liberalism, is regulation intended to protect or guard the moral character of citizens, rather than the more concrete interests in health, wealth, and so forth. Feinberg (1986: 27) refers to this as *moralistic legal paternalism*, characterized by laws intended to "prevent *moral harm*

(as opposed to physical, psychological, or economic harm) to the actor himself." Laws prohibiting victimless crimes such as drug use, prostitution, and gambling are often portrayed as moralistic, as are regulations of sexual behavior such as homosexual relations, sodomy, and the use of contraception. The latter are often regarded as the most pernicious cases of moralistic regulation because of the privileged status given sexual relations due to their essentially private nature. Indeed, the legal status of homosexuality inspired the famous Hart–Devlin debates over *legal moralism*, which regards regulation of behavior widely considered immoral without any necessary regard for the well-being of individuals engaging (or not) in such behavior (Devlin 1965; Hart 1963; see also R. Dworkin 1966; Sartorius 1972). It is but a small jump, by adjusting the motivation of moralistic regulations, to get to paternalism, in which the well-being, moral or amoral, of the individual is the focus (Ten 1972; G. Dworkin 2005).

But the other, less common way to interpret moralism in paternalism is through this chapter's focus on value substitution, in which a third party's judgments about an agent's proper interests are substituted for the agent's own judgments. Thus, we can see that paternalism in general contains a certain moralistic aspect, whether it deals with "traditionally" moral topics such as sex or drug use, or less emotionally charged issues such as food or money. The issue in either type of case is the autonomy of the individual to set her own interests and to pursue them within the bounds of justice that protect other people from her actions. Consider the automatic 401(k) enrollment plan, which implicitly assumes that retirement planning is a "better" or more prudent use for marginal income than any other use the employee might have for it. This type of paternalism need not be directly concerned with an agent's moral interests, but it does involve moral judgment over what her interests should be, and in that way is as moralistic as the more traditional paternalism aimed at an agent's moral well-being.

Relying upon this sense of moralism in value substitutions, we can see that paternalism can also be understood in terms of the debate over ethics of care versus ethics based on respect. Advocates of care ethics, such as Virginia Held (2006) and Michael Slote (2007), argue that relationships based on caring, as exemplified by those between mothers and their children, should form the basis for all relationships between persons, even anonymous strangers, as well as relationships between citizens and the state (displacing traditional conceptions of justice). This can be contrasted with schools of ethics, such as Immanuel Kant's, that focus on respect and then filter their ideas about beneficence through it. For instance, recognizing our duty to help others, Kant (1797: 448–49) wrote that

> since the favor we do implies that his well-being depends on our generosity, and this humbles him, it is our duty to behave as if our help is either merely what is due him or but a slight service of love, and to spare him humiliation and maintain his respect for himself.

More to the point of value substitution and paternalism, he wrote:

> I cannot do good to anyone in accordance with *my* concepts of happiness (except to young children and the insane), thinking to benefit him by forcing a gift upon him; rather, I can benefit him only in accordance with *his* concepts of happiness.
>
> (1797: 454)

In these terms, paternalism, particularly when value substitution is emphasized, seems motivated by care rather than respect, which is problematic given the nature of the relationship between the state and its citizens.

On one level, this is simply one way to restate the conflict between well-being and respect or between utilitarianism and deontology. But the care/respect dichotomy highlights the attitudes that argue for or against paternalism, as well as instances when paternalistic intervention is or is not appropriate. For instance, few would dispute that parents should interfere with the decisions of their children (though when and to what extent they should do this is debatable), or that a person might be justified in acting coercively toward a good friend she knows very well to save him from making the same mistake over and over again. These are relationships in which care is appropriate, specifically because the first person knows the second one well enough to presume to act in his or her true interests. But we would not expect to see mild acquaintances, much less total strangers, behave like this; respect seems more important than care in these cases, if only because one person does not know enough about the other to give "care." Beneficence toward persons one does not know well needs to be tempered by respect to a greater degree than care between family or good friends; as Kant said in the quoted passage above, beneficence must be practiced in a way that serves the other person's concept of happiness, which requires a certain degree of knowledge about the other person. He then invokes social distance:

> For in wishing I can be *equally* benevolent to everyone, whereas in acting I can, without violating the universality of the maxim, vary the degree greatly in accordance with the different objects of my love (one of whom concerns me more closely than another).
>
> (1797: 452)

But while care and beneficence can be varied, respect cannot (as it is owed to all rational beings possessed of dignity), so in relationships with strangers and mere acquaintances, respect must dominate care. Given the extremely impersonal nature of the "relationship" between the state and its citizens and the lack of personal knowledge inherent in it, citizens should only be shown respect by the state, since care on the part of the state would be overly presumptive and would necessarily involve value substitutions.

Recasting paternalism in terms of care versus respect also highlights its deleterious effect on personal responsibility, as well as the process of choice itself. One aspect of the caring nature of paternalism is that regulators (perhaps sincerely) do not want to force persons to bear the negative consequences of "bad" choices (according to the regulator's judgment). For example, some blame the rise in U.S. home foreclosures near the end of the first decade of the twenty-first century on poor loan choices (perhaps manipulated by unethical mortgage brokers), and subsequently new financial industry regulations were passed to promote simplicity in lending contracts. To be sure, many mortgage borrowers might have made poor decisions by their own judgment, but others made decisions which, while perhaps unfathomable to financial economists or policymakers, made sense to the borrowers when they entered into them. (See Bar-Gill and Warren (2008) for criticism of consumers' financial choices, and White (2011a) for a counterargument.) Since no one wants to see people lose their homes, regulators attempt to restrict the riskier aspects of financial products, which ultimately limits the options of rational, self-aware lenders to structure their financial portfolios according to their

interests (rather than those of a financial planner or economist). Along the same lines, many are uncomfortable with the thought of people becoming ill and not having health insurance, but mandating that everyone has it (and pays for it), motivated by discomfort with holding people responsible for their choices, limits the options of the very people subject to the "care" of policymakers and bureaucrats. Furthermore, it weakens the sense of responsibility among decision-makers, giving rise to moral hazard problems, while holding people responsible for the consequences of their actions promotes better decision-making (Klick and Mitchell 2006), which can reasonably be considered a more caring attitude, and also more respectful of autonomous choice.

4. Markets and Choice

Opposition to state paternalism on the basis of the epistemic and ethical problems with value substitution provides a strong argument for a reliance on private markets to organize much economic activity, especially when it pertains to private and personal choices of individuals. Paternalistic value substitutions motivated by care are usually not present in the realm of market transactions; private suppliers and merchants can generally be relied upon to further their own interests (which may or may not be oriented toward profit), and consider their customers' interests only insofar as that consideration will further the business owners' own. And even if a private business did deign to make decisions in the customers' own interests, it has no legal coercive power to enforce that judgment on the customer, as he or she has other choices regarding where to obtain the desired good or service (or, in the case of monopoly, might have the option of competing with an alternative, nonpaternalistic product or service). Only the possibility of free exchange and choice guarantees respect for the dignity and autonomy of free agents, without co-opting or manipulating their choices in (what is presumed to be) their own interests.

Furthermore, the market, with its basis in voluntary transactions assumed to be in all parties' subjective interests, embodies the respect for the dignity of persons that is violated by paternalistic intervention. Normally, defenses of the market focus on its effect on wealth-creation, but this is contingent on market conditions and circumstances at any given time, and does not reflect on the essential nature of the market as an autonomy-preserving institution. As I have written elsewhere (White 2010: 10), "insofar as the market is comprised of individual choices, freely made in the absence of coercion and deceit as enforced by the background laws of justice, it embodies respect of dignity in an essential, integral way." It also coordinates these individual decisions in a way that respects the dignity of all participants, allowing each to express her autonomy and pursue her goals and ends—as determined by each agent—consistent with all others doing the same (resulting in a limited version of Kant's kingdom of ends; see White 2011b: 116–17).

Again, this argument supporting the market, based on its respect for individual choice, autonomy, and self-determination, is purely procedural, and as such does not take into consideration the particular outcomes of a market-oriented system (in the spirit of Robert Nozick's distinction between procedural and historical justice). This is consistent with the position that respect, rather than care, is the proper attitude for impersonal relationships such as those that exist in the market, and especially that which exists between a state and its citizens. Ironically, this point is more difficult yet important to remember in cases in which care seems to be appropriate, such as the provision of health-care services

(White 2010). Universal health care is often presented as an act of beneficence or care on the part of the state towards its citizens, but it is also associated with reduced choice among health-care options for those same citizens. Often, the judgment of individuals is replaced by the judgment of bureaucrats who must allocate scarce medical resources among those who need or request them. But ironically, it can be said that it is in medical choice situations that personal autonomy and private values should be respected the most, and individuals, their families, and their physicians should make the ultimate choices, based on personal interests, medical judgment, and the resources available to the individual (which also stem largely from his or her previous choices). This provides a striking illustration of the conflict between care and respect, in which the care that seems appropriate conflicts with an even greater sense that respect is needed (and also explains why autonomy is a central concept in the discussions of bioethics).

5. Conclusion

A key element of liberalism is respect for the voluntary, self-regarding choices of competent, mature persons, whether justified by deontological concepts such as autonomy or dignity or consequentialist concerns of self-development and fulfillment. Paternalism offends this central liberal value chiefly, I have argued, through the substitution of the judgments and interests of distinct individuals with those of the choosers, circumventing their own values and concerns and disabling the pursuit of their own interests without any justification in the form of third-party effects. Of course, the possibility of such nonpaternalistic rationales raises the issue of mixed motivation of paternalistic laws and regulations (Feinberg 1986: 16–21): for instance, restrictions on smoking are more often justified in terms of cutting public health-care costs, or even protecting the health interests of nonsmokers, than concern for smokers. Understandably, these alternative rationales raise their own issues within the philosophy of liberalism, such as the propriety of public health-care expenditures themselves, but avoid the problems specific to paternalism. Furthermore, these other-regarding justifications can be used to mask the true reasoning between presumably paternalistic laws; even motorcycle helmet laws, perhaps one of the clearest cases of paternalism, can be defended in terms of the well-being of family members and the increase in public health-care costs. But if the argument of this chapter is accepted, vigilance is required to assess new laws and regulations for terms and justifications that offend Mill's harm principle, and thereby sacrifice persons' autonomy and dignity to the imposed interests of others.

Related Topics

Kant, Mill, Liberalism, Libertarianism, Freedom, Autonomy, Rights

References

Bar-Gill, O. and Warren, E. (2008) "Making Credit Safer," *University of Pennsylvania Law Review* 157: 1–101.

Brock, D.W. (1988) "Paternalism and Autonomy," *Ethics* 98: 550–65.

Devlin, P. (1965) *The Enforcement of Morals*, Oxford: Oxford University Press.

Dworkin, G. (1988) *The Theory and Practice of Autonomy*, Cambridge: Cambridge University Press.

—— (2005) "Moral Paternalism," *Law and Philosophy* 24: 305–19.

—— (2010) "Paternalism," in *Stanford Encyclopedia of Philosophy* (http://plato.stanford.edu/entries/

paternalism/).
Dworkin, R. (1966) "Lord Devlin and the Enforcement of Morals," *Yale Law Journal* 75: 986–1005.
—— (2000) *Sovereign Virtue*, Cambridge, MA: Harvard University Press.
Feinberg, J. (1971 [1980]) "Legal Paternalism," reprinted in *Rights, Justice, and the Bounds of Liberty: Essays in Social Philosophy*, Princeton, NJ: Princeton University Press.
—— (1986) *Harm to Self*, Oxford: Oxford University Press.
Griffin, J. (1986) *Well-Being: Its Meaning, Measurement, and Moral Importance*, Oxford: Clarendon Press.
Hart, H.L.A. (1963) *Law Liberty and Morality*, Stanford, CA: Stanford University Press.
Held, V. (2006) *The Ethics of Care: Personal, Political, and Global*, Oxford: Oxford University Press.
Hill, C.A. (2007) "Anti-Anti-Anti-Paternalism," *NYU Journal of Law & Liberty* 2: 444–54.
Kant, I. (1785 [1993]) *Grounding for the Metaphysics of Morals*, trans. J.W. Ellington, Indianapolis, IN: Hackett Publishing Company.
—— (1793 [1983]) "On the Proverb: That May Be True in Theory, But Is of No Practical Use," in *Perpetual Peace and Other Essays*, trans. T. Humphrey, Indianapolis, IN: Hackett Publishing Company.
—— (1797 [1996]) *The Metaphysics of Morals*, trans. and ed. M.J. Gregor, Cambridge: Cambridge University Press.
Klick, J. and Mitchell, G. (2006) "Government Regulation of Irrationality: Moral and Cognitive Hazards," *Minnesota Law Review* 90: 1620–63.
Mill, J.S. (1859 [1975]) *On Liberty*, ed. D. Spitz, New York: W. W. Norton.
Sartorius, R.E. (1972) "The Enforcement of Morality," *Yale Law Journal* 81: 891–910.
Slote, M. (2007) *The Ethics of Care and Empathy*, London: Routledge.
Sugden, R. (2008) "Why Incoherent Preferences Do Not Justify Paternalism," *Constitutional Political Economy* 19: 226–48.
Sunstein, C.R. and Thaler, R.H. (2003) "Libertarian Paternalism Is Not an Oxymoron," *University of Chicago Law Review* 70: 1159–1202.
Ten, C.L. (1972) "Enforcing a Shared Morality," *Ethics* 82: 321–29.
Thaler, R.H. and Sunstein, C.R. (2008) *Nudge: Improving Decisions about Health, Wealth, and Happiness*, New Haven, CT: Yale University Press.
Trout, J.D. (2005) "Paternalism and Cognitive Bias," *Law and Philosophy* 24: 393–434.
White, M.D. (2010) "Markets and Dignity: The Essential Link (with an Application to Health Care)," in *Accepting the Invisible Hand: Market-Based Approaches to Social-Economic Problems*, ed. Mark D. White, New York: Palgrave Macmillan.
—— (2011a) "We've Been Nudged: The Effects of the Downturn on Dignity and Responsibility," in *Consequences of Economic Downturn: Beyond the Usual Economics*, ed. Martha A. Starr, New York: Palgrave Macmillan.
—— (2011b) *Kantian Ethics and Economics: Autonomy, Dignity, and Character*, Stanford, CA: Stanford University Press.

69
RELIGION IN PUBLIC LIFE

Kevin Vallier and Christopher Eberle

Conflict over the proper role of religion in public life is quite old, yet the problems associated with it are as pressing as ever. Indeed, the philosophical problems raised by the proper public role of religion remain among the most challenging in political theory. Our aim in this essay is to clearly formulate these philosophical problems and to survey both prominent and plausible ways to resolve them. We point out where mainstream approaches to the public role of religion go wrong and how they might be rectified. Consequently, we will survey responses that include either (i) modifying the dominant view or (ii) rejecting the dominant view.

In Section 1, "Narratives," we reflect on the narratives that frame philosophical debate about the proper role of religion in public life. In Section 2, "Public Reason and Religion," we focus on the manner in which the academically dominant form of liberalism—public reason liberalisms—construes the public role of religion. In Section 3, "The Integrity Objection," we articulate one of the most persuasive objections to public reason liberalism. In Section 4, "Convergence," we argue that public reason liberalism can be formulated so as greatly to mitigate the force of the Integrity Objection, although this reformulation requires a rather thoroughgoing revision of its basic elements. In Section 5, "Religious Accommodation," we reflect on the claim that a liberal polity ought to make special accommodation for those citizens who dissent on religious grounds to the law of the land. Section 6 reviews our conclusions and encourages a somewhat revisionist take on these perennially pressing matters.

1. Narratives

Many philosophers and political theorists frame their understanding of the proper public role of religion with an historical narrative that traces how religion has shaped public affairs in the past. John Rawls introduces his *Political Liberalism* by arguing that political theory is best done in light of distinctively modern circumstances, that liberal polities are riven by pervasive pluralism of belief and practice. Rawls illustrates this pluralism with a familiar narrative about the role religion played in the inception of Western liberal democracies.

> What the ancient world did not know was the clash between salvationist, creedal, and expansionist religions. That is a phenomenon new to historical

> experience, a possibility realized by the Reformation. Of course, Christianity already made possible the conquest of people, not simply … to exercise power and dominion over them, but to save their souls. The Reformation turned this possibility inward upon itself.
>
> What is new about this clash is that it introduces into people's conceptions of the good a transcendent element not admitting of compromise. This element forces either mortal conflict moderated only by circumstance and exhaustion, or equal liberty of conscience and freedom of thought. … Political liberalism starts by taking to heart the absolute depth of that irreconcilable latent conflict.
>
> (Rawls 2005: xxv–xxvi)

Rawls here appeals to a familiar "creation myth" regarding the origins of liberal democracy. Once upon a time, there existed a religio-political unity called "Christendom." Almost everyone in Christendom practiced some version of Christianity and took its basic tenets for granted. But once Christians began to argue amongst themselves about which sect had the correct theology, they engendered a century and a half of devastating military conflicts between Lutherans, Calvinists, Anabaptists, and Catholics, the culmination of which was the so-called Thirty Years War. This "holy cataclysm" consumed the lives of a huge number of imperial subjects. Eventually, the Peace of Westphalia terminated the bloodletting, enshrining in Western culture and politics the core principle that religious difference ought never constitute a just cause for war (Philpott 2001: 81).

The storyline continues into colonial America. Due to religious persecution, Christian sects such as the Puritans and the Pilgrims traveled to the new world, but rather than establishing religiously tolerant societies, they created their own persecutory regimes. Religious repression in the United States continued for a century or more.

The problem raised by the master narrative is resolved by the rise of a sensible, secular class of individuals. In the United States, these include the founding fathers, who used secular reasoning to ensure that the United States would never impose religious orthodoxy on recalcitrant unbelievers. This American Republic serves as a guiding light to all nations, but, paradoxically, only by sharply restricting the role of religion in public life.

This narrative powerfully and subtly shapes moral understanding. It construes "religion" as a potentially mortal threat to social order and liberty. The story's lesson is that we must refrain from settling political differences based on our sectarian religious doctrines. However, there is reason to question the narrative. The basic conflict between religious advocates and their secular opponents is overdrawn. Instead, religiously based political action has often provided normative resources that reinforced liberal commitments.

Consider the role played by religious believers in enshrining the right to religious freedom in the early American Republic. Whereas the standard narrative pits the religious defenders of religious persecution and establishment against their secular and skeptical Enlightenment nemeses, alternative accounts, such as that articulated by Thomas Kidd, accord to the devout a decisive role in undermining religious persecution and establishment. According to Kidd, many colonial Americans were motivated by their conviction that "the cause of liberty" was "the cause of God" to resist both religious persecution and domination by imperial, colonial powers (Kidd 2010: 72). Politically effective commitment to religious liberty was decisively shaped by the Great

Awakening of the 1740s—a profound transformation of American culture in which itinerant preachers, such as George Whitefield, transgressed established ecclesiastical—and so political—boundaries in order to awaken socially respectable but spiritually dead congregants. Perceived as a threat to social order and respectable religion, these itinerant preachers were ostracized, vilified, fined, imprisoned, and on occasion flogged. But their numbers grew, as did the conviction that the persecutory establishment that impeded the free preaching of the Gospel was deeply un-Christian. Increasing numbers of citizens concluded that their religious convictions were incompatible with religious coercion. According to Kidd, these citizens provided famous advocates of religious freedom—such as the more secular Thomas Jefferson—with the demographic clout necessary to enact their reforms.

Kidd's counter-narrative does not reinforce the common admonition to "privatize" religious belief. To the contrary: if Kidd is correct, an unlikely coalition of skeptics, deists, religious liberals, and fervent evangelicals were led by their diverse normative commitments to converge on the same transformative achievement: the inclusion of a robust right to religious freedom in the United States Constitution. Kidd writes that

> [t]he evangelicals wanted disestablishment so they could freely preach the gospel; the rationalists and deists wanted disestablishment because they felt that an enlightened government should not punish people for their religious views. The combination of the two agendas would transform America, helping make it both intensely religious and religiously free.
>
> (Kidd 2010: 55)

If so, American religious liberty is both a secular *and* religious success story. The American commitment to religious freedom was no less a consequence of religious reasoning as of secular reasoning.

Another persuasive challenge to the common narrative advanced by the secular liberal tradition has been presented by Bryan McGraw. McGraw agrees with our contention that historical narratives about the role of religion in public life deeply affect our understanding of the philosophical problems related to them. McGraw argues at length that the traditional liberal understanding of the role of religion in politics as disruptive, destabilizing and often violent is incorrect. He counters that religious influence in contemporary public life resembles not the conditions found in the seventeenth century, the period from which most public reason liberals draw their understanding of religion in public life, but those in the nineteenth and early twentieth centuries in Europe, where religious political parties played a much subtler role that was considerably friendlier to liberal democracy than many liberals would expect (McGraw 2010: 46). Cases of religious political parties in the Netherlands, Belgium, Germany and Austria demonstrate that religious communities "organized powerful networks as a means both of protecting their members against the corrosive acids of modernity and of pursuing their social and political goals via electoral contests and parliamentary debate" (McGraw 2010: 64). While McGraw does not claim that religious political parties are always compatible with parliamentary democracy, neither does he claim that they are always incompatible. Careful and empirically informed study of the actual practice of religious believers presents a more complicated picture than that found in familiar liberal narratives.

We believe that the counter-narratives articulated by Kidd and McGraw comport with a very broad understanding of the public role of religion in liberal polities, one that

welcomes religious contributions to public decision and debate about common matters. That broad understanding incorporates the crucial claim that religiously justified interventions in public life need not differ in any morally relevant respect from secular interventions. Both can advance or retard the cause of justice, liberty and the common good. Both have an important and legitimate role to play in the public life of a robust liberal polity. Let us now specify that role in detail.

2. Public Reason Liberalism

Much of the recent dispute about the proper role of religion in public life has concerned the nature of *liberalism*. The dispute has usually focused on one sort of liberalism, "public reason" liberalism. This liberal political theory combines traditional liberal commitments to freedom of the press, speech, assembly, universal suffrage and representative government with a demand that all state coercion be "publicly justified." The requirement of public justification holds that in order to be morally permissible, state coercion must be justified to the citizens of that liberal polity, called "members of the public." Coercion is so justified only when each citizen has sufficient reason to endorse the coercive measures to which each is subject. So all public reason liberalisms affirm some version of what we call the Public Justification Principle (PJP):

> *Public Justification Principle*: A coercive proposal, law or policy L is morally permissible only if each member of the public has sufficient reason to endorse L.

To specify the public role the PJP assigns to religion, we make four clarifying remarks.

First, PJP imposes a moral (not legal) constraint on state coercion that can be formulated more or less narrowly—to include every instance of state coercion or only some subset thereof. So, for example, Rawls's formulation of PJP applies only to matters of basic justice and constitutions essentials (Rawls 2005: 137), whereas Gerald Gaus's formulation applies to specific laws or policies (Gaus 2011: 495).

Second, although PJP requires that state coercion be justified to members of the public, it need not imply that justification is tied down by the gross ignorance, inconsistent reasoning, vicious attitudes, etc. of actual citizens. Indeed, public reason liberals typically *idealize* the commitments and values of members of the public in light of some standard of good information, reasoning, consistency, reasonableness, and the like. Views vary considerably, as Rawls, David Gauthier, Jürgen Habermas and Gaus have articulated competing accounts.

Third, PJP specifies that coercion-justifying reasons must be sufficient. A sufficient reason must have some significant normative weight—evidentially adequate, overriding, or even conclusive. Moreover, justificatory reasons must satisfy some standard of intersubjectivity—"publicity," "sharability," "accessibility," "intelligibility," and the like. Often justificatory reasons must be reasons that can be shared or that are open to external criticism. Reasons that violate this standard lack justificatory force. So, for example, a reason unintelligible to most members of the public fails to be justificatory—no state act or law can be justified by that reason.

Finally, PJP helps to specify what makes for responsible citizenship. Public reason liberalism is partly constituted by an ethic of citizenship according to which a citizen ought to restrain herself from voting for or advocating in favor of any coercive measure that cannot be publicly justified (see Weithman 2002: 37–66) for discussion of such

views). A citizen who violates this *principle of restraint* fails to treat her compatriots as her moral and civic equals and so acts *disrespectfully* (Larmore 2008: 213). Public reason liberals advocate different conceptions of responsible citizenship—they might restrict restraint to legislators and officials or extend them to all citizens.

In light of the foregoing, what does public reason liberalism have to say about the public, political role of religion?

Mainstream public reason liberals typically hold that *no religious reason counts as a public reason*. This is so because religious reasons are taken to fail a plausible test of intersubjectivity. Since religious reasons cannot be endorsed by all members of the public as somehow justified or valid, they fail to count as properly public. From this position, serious restrictions on the justificatory role of religious reasons are derived. Given that due respect requires that citizens restrain themselves from supporting coercive measures that cannot be publicly justified, and given that religious reasons cannot publicly justify anything, citizens are morally required to restrain themselves from supporting coercive measures that cannot be justified other than by some religious rationale. Rawls's proviso is noteworthy in this regard, as (in its final form) it requires that religious public discourse be backed by suitably public reasons (Rawls 1999: 144). Notice here that mainstream public reason liberalism treats religious and secular reasons differently, as secular reasons can serve a justificatory role far more often than religious reasons (Audi 2000: 86–104). This is not to say that all secular reasons can justify state coercion, as many might lack such force. There is, therefore, an asymmetry between religious and secular reasons embedded in much of contemporary political theory. The secular is universal and common, whereas the religious is particular and sectarian. (This asymmetry is typically presupposed by the narrative we discussed in the previous section.)

3. The Integrity Objection

Public reason liberalism has many critics, some of whom object to its restraint on the use of religious reasons. These critics hold that fundamentally *liberal* commitments count against such restraint.

How so? Let us articulate this criticism by first explaining what it is not. No reputable critic of public reason holds that citizens must be allowed to make political decisions on the basis of any reason whatsoever. These critics grant that some religious reasons lack justificatory force—those that are self-contradictory or that are manifestly illiberal, such as claims that some persons lack moral status. Instead, they object to *global* restrictions on religious reasons and principles of restraint that apply to every religious reason as such.

Critics find global restrictions on religious reasons objectionable for a number of reasons (see Neal 2009 for an overview of the dialectical terrain), but we focus on what is called the "Integrity Objection." The Integrity Objection holds, briefly, that principles of restraint impose unjustifiable burdens on the integrity of religious citizens. This objection begins with a familiar self-understanding among religious citizens, specifically theistic citizens.

Theists typically believe that God issues imperatives that enjoin human beings to engage in certain lines of conduct, some concerning their private activities, but many concerning their public responsibilities. These persons take themselves to be morally obligated to "integrate" their faith into every aspect of their lives. Nicholas Wolterstorff articulates the theological orientation of these "integralist" citizens:

> [I]t belongs to the religious convictions of a good many religious people in our society that they ought to base their decisions concerning fundamental issues of justice on their religious convictions. They do not view it as an option whether or not to do so. It is their conviction that they ought to strive for wholeness, integrity, integration, in their lives; that they ought to allow the Word of God, the teachings of the Torah, the command and example of Jesus, or whatever, to shape their existence as a whole, including, then, their social and political existence.
>
> (Wolterstorff 1997: 105)

Suppose that a religious citizen has a compelling theological reason to support some coercive measure C for which there is no shared or accessible reason. In this case, mainstream public reason liberals (those who endorse religious restraint) hold that she should refrain from supporting C, despite her conviction to the contrary. Nonetheless, the religious citizen takes herself to have a strict moral duty to God to do otherwise. Anything less than full compliance constitutes a kind of hypocrisy that precludes genuine friendship with God. Consequently, she seems to have reason to reject PJP precisely because it requires her to disobey God. Restraint is religiously unacceptable.

Perhaps a public reason liberal need not be overly concerned by the fact that the kind of restraint she affirms is unacceptable to religious believers. From her perspective, she should be no more troubled by integralist believers than those who adhere to theological claims that require them to commit all kinds of familiar illiberal practices—human sacrifices, exterminating heretics, and so on. Yet no commitment is more fundamental to liberal political theory than respect for individual conscience. Commitment to liberty of conscience is not merely a legal proposition—that the state should not punish citizens for their conscientiously held beliefs (religious or otherwise). Liberals should expect citizens to do their best to live according to their deeply held principles. What could be a more paradigmatically liberal understanding of what makes for responsible citizenship in a pluralistic polity?

The commitment to restraint therefore seems *illiberal*. Any liberal should be troubled by the fact that she recommends a citizen not make political decisions by relying on his deepest commitments simply because other persons find their reasons inaccessible, unshareable, etc. It is true that she advocates such restraint because she is concerned that coercion only be employed when it is mutually justifiable, but the liberal nonetheless has excellent reason to be wary of restraint. The traditional commitment to respect for conscience militates against the claim that some citizens should be restrained.

Two clarificatory points. First, to argue that religious citizens should be permitted to support coercive measures on religious grounds does not mean that they may be indifferent to public justification. Instead, even critics of mainstream public reason liberalism hold that religious citizens should regard public justification as a great moral good and think it worthy of pursuit. Religious citizens have good reason to exit their own perspective and take that of others in an attempt to appeal to their fellow citizens. According to these critics, religious citizens ought to *pursue* public justification, though should they fail to justify the measure in public terms, they may still support it on religious grounds (Eberle 2002). Second, the Integrity Objection does not apply only to religious citizens. Appeals to freedom of conscience against restraint apply to secular citizens as well, be they Kantians, utilitarians, or whatever. According to the integralist critics of public reason liberalism, all such restraint must be rejected on liberal grounds: secular and

religious citizens have excellent reason to be chary of the kind of restraint advocated by mainstream public reason liberalism.

The Integrity Objection is mostly ignored by public reason liberals, and sometimes derided. Stephen Macedo, for example, has told integralists to "Grow up!" (Macedo 2000: 35). But we believe that the Integrity Objection is a powerful threat to public reason liberalism. In fact, we both think that objection succeeds against mainstream public reason liberalism (those forms of public reason liberalism that endorse restraint). However, it is not often clear how the Integrity Objection succeeds. Kevin Vallier has argued that the Integrity Objection only threatens public reason liberalism via several complex steps (Vallier 2012).

First, PJP must entail a conceptual claim about the kinds of reason that can satisfy it. As we have seen, mainstream advocates of public reason liberalism assume a very strong claim about the nature of justificatory reasons, namely, that they must be accessible or shareable. The first step merely specifies what might count as justificatory reasons, not how they are to be used. The next step involves showing that the proper account of justificatory reasons must entail a principle of *exclusion* which specifies precisely how non-justificatory reasons are to be excluded as bases of coercion (that is, via constitutional amendment, legislative or judicial practice, etc.). So, for example, if PJP can be satisfied only by shareable or accessible reasons, one could argue that such reasons cannot serve as a basis of coercion by a legitimate political order as specified by a publicly justified constitutional principle such as the Establishment Clause of the U.S. Constitution (which expressly rejects religious reasons as bases for the exercise of Congressional power). Third, public reason liberals interested in defending restraint must show that the relevant principle of exclusion entails a principle of restraint. That is, the fact that some reasons should be excluded from law must entail some civic principle that regulates political behavior. Given this complex understanding of the core commitments of public reason liberalism, advocates of the Integrity Objection must show that restraint violates integrity. Let us assume that they can do so. Let us also assume that the relevant degree of integrity violation is sufficient grounds to reject the principle of restraint in question. It is *then and only then* that the Integrity Objection defeats the PJP. Advocates of the Integrity Objection almost never set out their arguments in this much detail.

4. Convergence

We believe that principles of restraint that appear in the literature impose objectionable costs to integrity. However, there are two ways one might respond to the Integrity Objection. First, one might simply reject PJP (as does Eberle 2002). That is the response pursued by integralists. A second reply involves modifying the common interpretation of PJP (Vallier 2012). This response holds that PJP can withstand the Integrity Objection but only via a reinterpretation of its claims.

The second reply holds that public reason liberals should reject shareability and accessibility requirements while still affirming PJP. The literature sometimes refers to conceptions of justificatory reasons that include shareability or accessibility requirements as "consensus" conceptions, as opposed to "convergence" conceptions of reasons. Convergence holds that all of the reasons a citizen affirms (depending on the degree of idealization) can play a role in public justification. Specifically, all of a citizen's reasons can potentially serve to *defeat* coercive laws, which is to say that even private, sectarian and "religious" reasons can demonstrate that a coercive law cannot be justified to each

(appropriately idealized) member of the public. Nonetheless, many coercive laws will still fail to be justified if they depend wholly on unshared considerations, for many citizens might still lack reasons of their own to endorse the coercion in question. Although the convergence conception implies that even sectarian reasons can defeat otherwise morally permissible coercive laws, it is most unlikely, under conditions of pervasive pluralism, that any one sectarian reason can justify a coercive law. Note that it is far more likely that some diverse spread of sectarian reasons justify a coercive law.

By allowing citizens to make political decisions even solely on the basis of their religious convictions, convergence sets restraint aside and in so doing avoids the Integrity Objection. Even so, it is quite restrictive of state power since it permits many more reasons to play the role of defeaters. The crucial question is which political laws remain: does convergence restrict the coercive power of the state so much that it threatens to undermine the legitimacy of morally necessary policies, such as the use of military violence in a just war?

Convergence can be defended via a number of routes. First, one might simply accept standard defenses of PJP but provide arguments against the shareability and accessibility requirements. For example, some have argued that when either requirement is spelled out in detail, it proves to either be far less restrictive than supposed or implausibly restrictive on the public reason liberal's own terms (Eberle 2002: 234–293; Vallier 2011). Second, one can defend convergence by arguing that it better expresses respect for persons under conditions of reasonable pluralism than do consensus conceptions given that convergence permits citizens to appeal to the full range of their reasons when it comes to justifying coercive laws or policies. Further, one could argue that convergence thereby better respects individual liberty by rejecting religious restraint.

Whatever the outcome of this debate, the convergence view deserves further exploration. It is here we believe that the theoretical battle between public reason liberals and their religious critics must turn.

5. Religious Accommodation

As we have seen, critics of mainstream public reason liberalism argue that its doctrines of restraint are easier on secular citizens than religious citizens, as restraint applies to them less forcefully. However, there is another vital domain of philosophical concern where the two positions are reversed. In the law, *religion* is given special preference over non-religious moral and philosophical doctrines with regard to accommodation. The United States Constitution paradigmatically protects the free exercise of religion as distinct from other domains of human experience and this protection has often been applied narrowly. One commonly cited example is the exemption of the Amish from Social Security. Their rejection of state-mandated pensions is based on their interpretation of a biblical injunction to care for one's elderly. They regard state-mandated pension systems as co-opting that duty from the church, and they were exempted on the basis of their religious objections. However, those with moral objections to the same effect are not afforded this exemption.

Philosophers ask whether this preferential treatment can be justified either constitutionally or morally. The constitutional issue arises in the context of the Establishment Clause of the First Amendment which bars the establishment of a particular religion and the Equal Protection Clause of the Fourteenth Amendment which requires U.S. law to treat all persons equally in terms of their constitutional rights and duties. Some

argue that by giving special preference to religion, U.S. constitutional law establishes an inequitable privilege for religious over secular forms of life. Further, many maintain that the preferential treatment of religion is inconsistent with equal protection. The legal issues are acutely raised by conscientious objectors. For a significant period of U.S. history, conscientious exemptions were given only to pacifist Christian denominations. Over time, conscientious objector status was extended to all theistic religions and then to various non-religious and moral doctrines. However, much U.S. law resisted extending conscientious objector status this broadly: such an extension would permit too many exemptions and encourage deception on the part of insincere objectors. A legal question arises as to whether these practical considerations are sufficient to justify what appears to be unequal treatment in the law.

A number of legal theorists have argued that preferential treatment of religion is morally unjustified. Michael Perry has defended a right of "moral freedom" (Perry 2010: 996). If he is correct, the reasons that count in favor of the legal protection of religious conscience apply no less forcefully to moral conscience: the features that make religion specially worthy of protection (for instance, that it plays a central role in structuring an individual's values and/or practices) are on all fours with various moral doctrines. In contrast, Kent Greenawalt has maintained that for a number of practical, legal and moral reasons, religion might deserve its special status. Many judges will have trouble distinguishing "genuine claims of conscience from lesser moral objections" without using religion as a standard of demarcation (Greenawalt 2010: 906). Greenawalt also suggests that legal tradition itself is some reason to continue religion's special treatment, and he also argues that in comparison with non-religious moral claims there is far more "at stake for the typical believer; God may punish wrongdoers when this life is over" (Greenawalt 2010: 914). In the end, though, even Greenawalt holds that if non-religious claims of conscience are sufficiently similar to religious claims and "there are not strong independent reasons to prefer religious claimants, a court should rule that the Equal Protection Clause and the Establishment Clause require equal treatment" (Greenawalt 2010: 917).

Other legal theorists have objected to the preferential treatment of religion by reflecting on the claims of *conscience*. As we have seen, a commitment to freedom of conscience is at the heart of a genuinely liberal sensibility. Conscience is typically taken to be a faculty or capacity that renders important or weighty moral and religious judgments and so generates weighty reasons for action. We have reason to respect conscience by virtue of its capacity to produce plans and pursuits worthy of respect. Martha Nussbaum has argued that conscience is worthy of respect by virtue of the fact that its deliverances are the natural consequence of a morally admirable search for meaning. Human beings are meaning-seeking creatures; this fact about human beings is most excellent, and so we have excellent reason to respect the claims of conscience generally (Nussbaum 2008: 169). Of course, the convictions of conscience that derive from the search for meaning can be either religious or secular, and as a consequence we have excellent, reason to accommodate both religious and secularly grounded convictions of conscience. As with Perry, Nussbaum denies that a liberal polity may give preferential treatment to religious forms of life alone.

Brian Leiter has articulated a distinctive objection to the claim that religiously and secularly grounded claims of conscience should be accorded equal protection in law. He articulates his argument by employing Stephen Darwall's distinction between "appraisal" and "recognition" respect (Darwall 1977). A person is treated with

recognition respect when she is treated in accord with her status as a person, full stop. In contrast, a person is treated with appraisal respect when she is treated in accord with her contingent features, such as her character traits or accomplishments. But Leiter has argued that religious judgments of conscience are not worthy of appraisal respect due to the fact that they are rooted in bad epistemic practices (Leiter 2010: 945). While Leiter's argument is controversial and arguably pejorative, he raises an important point. It is certainly possible that an individual who sincerely employs her conscience in order to lead a meaningful life nevertheless forms systematically bad judgments. It is not obvious that the convictions of conscience she forms would merit appraisal respect. If Leiter is correct in his estimation of the manner in which believers form their religious convictions, then it is at least doubtful that we have reason to respect religiously formed convictions of conscience—much less preferentially accommodate religiously grounded convictions of conscience.

Perhaps a better strategy—one that more firmly grounds a commitment to accommodating religiously and secularly grounded convictions of conscience—roots that commitment in respect for *persons* on the grounds that respecting the deliverances of conscience is part and parcel of treating someone with respect. Consequently, a right of conscience can be rooted in human dignity (arguably the basis of the right in the Universal Declaration of Human Rights). In turn, whether conscience is worthy of respect is not based on its particular judgments or its role in a search for meaning but, rather, upon the fact that persons as such are due (recognition) respect (a point with which Leiter largely agrees). This strategy obviates a worry about Nussbaum's position. Many individuals' significant conscientious commitments of conscience develop when they stop searching for meaning. Instead, their conscience convicts them because they believe they have already developed deep convictions about the meaning of life. A recognition respect interpretation of the right of conscience makes sense here. And since this form of respect applies to all persons, we can arguably avoid worries about unequal treatment. Anyone who makes a claim of conscience would have at least some claim on the state for accommodation, religious or no. This right would not show partiality as it would be open to all in virtue of their status as persons.

The philosophical issues raised by religious accommodation can be boiled down to two questions: In constitutional law, should religion be given preferential treatment? If not, how should the law change? We have reviewed different responses to these questions, some of which hold that religion is special for pragmatic and epistemic reasons, whereas others hold that religion is not special because the considerations motivating respect for religious convictions apply equally to non-religious, moral convictions. Some try to ground this equal treatment in a right of conscience, though many are skeptical that the idea of conscience threatens to make the law partial or excessively biased in favor of conceptions of conscience rooted in a particular moral or religious tradition.

6. Conclusion

This essay maps the philosophical territory concerning the role of religion in public life. While we have not covered every issue, we have focused on three significant sources of concern: common narratives, public reason and religious accommodation. Worries about religion in public life arise out of our inherited and morally orienting but flawed historical narratives. These narratives raise the question of whether privatization is an appropriate social norm. They also raise concerns about whether religion deserves special treatment in the law.

In one sense, the philosophical issues raised herein can be understood as asking one broad question: how and to what extent should religious and non-religious contributions to public life be treated equally? Our inherited narrative often places special blame on religion and these biases are frequently transmitted into the public reason debate, where religious reasons are seen as private, divisive and sectarian and secular reasons as peaceful and neutral. Similarly, due to the religiosity of Americans in decades and centuries past, religion has acquired a special place in the law. Thus, religious claims are accorded greater legal respect than non-religious claims.

We believe that religious and secular reasons should be treated symmetrically in political theory and in the law. The arguments for unequal treatment seem weak with respect not only to biases against but also in favor of religion. In our view, many of these issues can be framed as questions within the framework of liberal egalitarianism where there is a presumption in favor of treating religious and secular reasons on an epistemic and moral par in political life. Equal treatment of the religious and the secular is particularly appropriate to a pluralistic liberal polity in which citizens are deeply committed both to religious creeds and their denials.

Related Topics

Contractualism and Political Liberalism, Natural Law and Rights Theory, Rights, Toleration

References

Audi, R. (2000) *Religious Commitment and Secular Reason*, New York: Cambridge University Press.

Darwall, S. (1977) "Two Kinds of Respect," *Ethics* 88: 36–49.

Eberle, C. (2002) *Religious Conviction in Liberal Politics*, New York: Cambridge University Press.

Gaus, G. (2011) *The Order of Public Reason*, New York: Cambridge University Press.

Greenawalt, K. (2010) "The Significance of Conscience," *San Diego Law Review* 47: 901–918.

Kidd, T. (2010) *God of Liberty: A Religious History of the American Revolution*, New York: Basic Books.

Larmore, C. (2008) *The Autonomy of Morality*, New York: Cambridge University Press.

Leiter, B. (2010) "Foundations of Religious Liberty: Toleration or Respect?," *San Diego Law Review* 47: 935–959.

Macedo, S. (2000) "In Defense of Liberal Public Reason: Are Slavery and Abortion Hard Cases?," in R. George and C. Wolfe (eds.), *Natural Law and Public Reason*, Washington, DC: Georgetown University Press: 11–49.

McGraw, B. (2010) *Faith in Politics: Religion and Liberal Democracy*, New York: Cambridge University Press.

Neal, P. (2009) "Is Political Liberalism Hostile to Religion?," in S. Young (ed.), *Reflections on Rawls: An Assessment of His Legacy*, Burlington: Ashgate: 153–176.

Nussbaum, M. (2008) *Liberty of Conscience*, New York: Basic Books.

Perry, M. (2010) "From Religious Freedom to Moral Freedom," *San Diego Law Review* 47: 993–1013.

Philpott, D. (2001) *Revolutions in Sovereignty: How Ideas Shaped Modern International Relations*, Princeton, NJ: Princeton University Press.

Rawls, J. (1999) *The Law of Peoples with "The Idea of Public Reason Revisited,"* Cambridge: Harvard University Press.

—— (2005) *Political Liberalism*, New York: Columbia University Press.

Vallier, K. (2011) "Against Public Reason's Accessibility Requirement," *The Journal of Moral Philosophy* 8(3): 366–389.

—— (2012) "Liberalism, Religion and Integrity," *Australasian Journal of Philosophy* 90(1): 149–165.

Weithman, P. (2002) *Religion and the Obligations of Citizenship*, Cambridge: Cambridge University Press.

Wolterstorff, N. (1997) "The Role of Religion in Decision and Discussion of Political Issues," in R. Audi and N. Wolterstorff (eds.), *Religion in the Public Square: The Place of Religious Convictions in Political Debate*, Lanham, MD: Rowman & Littlefield: 67–120.

Further Reading

While *Political Liberalism* is a crucial work, Rawls's final statement on religion in public life is J. Rawls (1997) "The Idea of Public Reason Revisited," *The University of Chicago Law Review* 64: 765–807. Another important public reason or deliberative democrat contribution to the public reason literature that addresses religion in politics is A. Gutmann and D. Thompson (1996) *Democracy and Disagreement*, Cambridge: Harvard University Press, along with S. Macedo (1991) *Liberal Virtues: Citizenship, Virtue, and Community in Liberal Constitutionalism*, Oxford: Clarendon. A classic treatment of religion and politics problems can be found in R. Audi (2000) *Religious Commitment and Secular Reason*, New York: Cambridge University Press. Some prominent criticisms of public reason liberalism's standard treatment of religion can be found in C. Eberle (2002) *Religious Conviction in Liberal Politics*, New York: Cambridge University Press. Article-length criticisms include P. Quinn (1997) "Political Liberalisms and Their Exclusion of the Religious," in P. Weithman (ed.) *Religion and Contemporary Liberalism*, Notre Dame, IN: Notre Dame University Press and G. Gaus (2010) "The Place of Religious Belief in Public Reason Liberalism," in M. Dimova-Cookson and P. Stirk (eds.), *Multiculturalism and Moral Conflict*, Oxford: Routledge: 19–37. Many traditionally secular-leaning philosophers have lost their hard edge as of late. See R. Audi (2009) "Natural Reason, Natural Rights, and Governmental Neutrality Toward Religion," *Religion and Human Rights* 4: 157–175 and J. Habermas (1996) "Religion in the Public Square," *European Journal of Philosophy* 14: 1–25. For defenses of the convergence view besides *The Order of Public Reason*, see J. Stout (2004) *Democracy and Tradition*, Princeton, NJ: Princeton University Press and K. Vallier (2012) "Consensus and Convergence in Public Reason," *Public Affairs Quarterly*, 25: 261–279.

INDEX

'Abbasids 206, 207
Abelard-Heloise complex 283
ab initio problem 578
abolitionist movement: and anarchism 221–2; punishment 771
absolutism, and conservatism 247–9
Abu Bakr 205–6
Achenwall, Gottfried 103
achievement gap, education 723
Ackerman, Bruce & Fishkin, James 613
acquisition, just 75–76, 428, 430, 530
Action Française 170
action-space 413–14
Active Intellect, and human happiness 211
activism: environmental 293, 296, 297–8; political 156, 614; women 156
Adam Smith Problem, The 97, 99, 100
adaptive preference formation 581–2
Adkins, Arthur 17
Adorno, Theodor 165
adversity, usefulness of 49
Aeterni Patris 32
Agathon 170
agency: autonomous 576; direct/oblique harmful 545–6; human 150–1, 162, 448; and human rights 500–1; moral and sentiment 287–8; and universal rights 64; valuable autonomous 350
agential direction/governance, and autonomy 576
agents, rational 118
aggregation of health benefits 739
aggregative views, democracy 609–10
aggression, resisting 540
agnostic ethic of care 356
agnostic pluralism 353, 356–7
agonism 356
agorism 225
agreement: reasoning from 321–5; and social practices 325
Albert the Great 25
Alexander VI 55
'Ali 206, 207
'Alids 207
alienation: and human flourishing 141; and meaningful work 757–9; subjective/objective 141
allies, Saptanga theory of the state 197
all-pervading principle 644
alternating offers models 701
Althusser, Louis 176, 267–8
altruism, and civilization 221
altruistic punishers 652–3
Ambedkar, B.R. 200, 201
ambivalence, and second-order volitions 579
America: Disneyland as 671; power in 585 *see also* United States
American Civil War 222
American culture, and France 172
American exceptionalist conviction 781
American Paradox 300
American Revolution 136
amour de soi 84
amour propre 85, 88
Anabaptists, and anarchism 218
Analects 188
Analytical Marxists 275
Ananga Ranga 194
anarcha-feminists 226
anarchism 217–27
anarchist militias 224
anarcho-capitalism 224–6
anarcho-communist ideals 224
anarcho-primitivism 226, 298
anarcho-syndicalism 224
Anarchy, State and Utopia 77, 225, 422, 423, 425
Anderson, B. 483
Anderson, Elizabeth 394, 398, 460, 461, 464–5, 665, 725
Andrews, Stephen Pearl 220
animal motion 42
animal rights, and Confucianism 185
animals, status of 555, 556
al-Ankam al-sultaniyya 209
Annas, Julia 279
Anonymity, and Arrow's Theorum 682
Anscombe, Elizabeth 364
antagonism, and coercive dictatorship 356
anti-elitism, and Daoists 188
anti-globalization, and anarchism 226
anti-imperialism, Hobhouse 155
anti-whaling 297

Apartment, The 696
Apology, The 4, 5
Aquinas, Thomas 25–35, 217–18, 376, 382
Aquinus: Moral, Political, and Social Theory 34
Arab Spring, as terrorism 778
arbitrary interference, and republicanism 259
arbitration networks 225
Archard, David 744–6
architecture, and postmodernism 667
Arendt, Hannah 162, 163, 164
Arguing About War 786
aristocracy, as form of government 29, 89
Aristotle 14–23; and distributive justice 703; and ethical theory 379; and Medieval thinkers 40; metaphysics/epistemology of 26–7; perfectionism 342–3; and reasoning 357; role of state 368; and the self 575
Aristotle: Political Philosophy 21
Armand, Émile 223
Arneson, Richard 391, 398, 457, 460, 461, 582, 757, 758
Aron, Raymond 175, 176, 177
Arrow, Kenneth 610, 679
Arrow's Theorum 679–87
art: and politics 161; psychological effect of/Plato 7
artha 193, 194
Arthasastra of Kautilya 201
Ashoka, Emperor 199
ashrama dharmas 194–5
Aspemont, Claude d' & Gevers, Louis 684
Asquith, Herbert 239–40
atheism, and natural law 383–4
Augustine 25, 217–18
Aulard, Alphonse 170
Aumann, R. 698
Aurobindo, Sri 201, 202
Austin, J.L. 707
autarchy 574
authenticity 161
authenticity claim, ethical-political discourse 710
authenticity condition, autonomy 577
authoritarian conservatives 244
authoritarian personality, and anti-democracy 165
authority, and legitimacy, 596–605 *see also* political authority
automatic sanctions 769
autonomous agency, motivational/valuational systems 576
autonomous self-determination 790
autonomy 573–83; individual and political authority 599; individual/value plurality 361; Kant 109; liberal and nationalism 491–2; Machiavelli 49; and paternalism 789, 791, 798; as perfection 347–8; and perfectionism 350; personal 563; personal and health care 798; and societal culture 491; and structuralism 175; three senses of 347
autonomy-based state perfectionism 350
autonomy-transparency thesis 576
avoidable responsibility 392
Axelrod, Robert 652
axiomatic bargaining theory 700

Badiou, Alain 178
Baglioni, Giovampagolo 55
Bai, Tongdong 181–90
Baker, J. & Jones C. 449
Bakunin, Mikhail 223
Balinski, Michel & Laraki, Rida 689
al-Banna, Hassan 213
Barbarism with a Human Face 176–7
Barbusse, Henri 171
bargaining analysis 700, 701
bargaining problem 699
Barker, Colin 273
Barker, Ernest 147
Baron, Hans 255
Barrès, Maurice 170
Barry, Brian 485, 508–9, 511, 590
Barthes, Roland 175
Basic Needs Approach 454
basic needs model, and distribution 474–5
Basic Rights 453
Battle of Siffin 207
Baudrillard, Jean 667, 670–1, 674, 675–6
Bauman, Zygmunt 672–4, 676
Bayesian rationality 693
Beauvoir, Simone de 173, 174, 285
behavioualism, and political science 586
beings and doings 457
Beitz, Charles 440–1, 502, 503
belief(s): and cultural difference/toleration 637; and experience 658; Fixation of Belief 665; and toleration/intolerance 632; well-settled 660–1
Bellegarrigue, Anselme 221
Benda, Julien 171
beneficence 99, 796
benefit: greatest benefit of the least advantaged 403, 406; mutual 403–5
benevolence: and justice 94–5; and morality 93
Benhabib, Seyla 288
Benn, Stanley 574
Bensaïd, Daniel 266–7
Bentham, Jeremy 311, 498, 601
Bergson, Henri 170
Berlin, Isaiah 235, 255, 256, 353–7, 359–60, 562–71
Bertram, Christopher 82–90
"best in itself" 126
Between Facts and Norms 706
Bhagavadgita 194, 195
Bhandarkar, R.G. 201
biased transmission 652
Binmore, Ken 652, 696
biological evolution, and Idealism 147
biological science, and political thinkers 147
bioregionalism 296
Black, Duncan 682, 686
Blackorby, Charles & Donaldson, David 684, 685
Blackstone, William 772
Bloch, Marc 173

blueprints, Utopian 143–4
Blühdorn, Ingolfur 299
body metaphor, Middle Ages 45
body politic, Medieval thinkers 37, 38, 39
Bohman, James 158–68
Bolshevik Revolution 170, 171
bombings, US marine base/Beirut 780
bonds, political/interpersonal 676–7
Bookchin, Murray 226, 296
Borda rule 686
borders, and justice 526–35
Borel, E. 696
Borgia, Cesare 55
Bosanquet, Bernard 147, 148, 152–3, 153–4
Bosanquet, Helen 155–6
Bossert, Walter 685
Bouglé, Célestin 169
bounded justice 530
Bourdieu, Pierre 178
bourgois philosophy/society, France 170, 171, 172
bourgois state 271
Boutroux, Emile 170
Boyd, Robert & Richerson, Peter 651–2
Brahmacharya 194
Brahmin 198
Braithwaite Bargaining Problem 700
Braithwaite game 698–9
Braithwaite, Richard 696, 699, 702
Brake, Elizabeth 749
Brasillach, Robert 173
Bratman, Michael E. 576
Braybrooke, David 447, 448
Break Through 300
Brentano, Franz 580
Brethren of Purity 211
Brighouse, Harry 721–31
British Idealism 152–3, 154
British political thought, late 19c/early 20c 146–56
Brock, Dan 789
Brock, Gillian 444–54
Brooks, Thom 517–24
Broome, J. 736
Brower, David 297
Brown, John 222
Brundtland Definition of Sustainability 478
Brundtland Report 294
brute luck 391–2, 393, 465, 737, 752
Brutus 54
Buchanan, Allen 605
Buchanan, Allen & Keohane, Robert 604
Buchanan, James M. 422, 425, 426
Buchanan, J.M. & Tullock, G. 426, 429–30
Buckle, Stephen 92–101
Buddhism, and caste system 198–200
Burke, Edmund 219, 243, 468

caliphate, the 205–7
callings 755
Callinicos, Alex 266–76
Camus, Albert 174–5
Caney, Simon 485
capabilities: aggregating 462–3; cental human 461–2; and human rights 520; as metric of justice 457–8; of people 519–20; and power 591; v. social primary goods 458–60
capabilities approach: and distribution 454, 474; as liberalism 369; and positive duties 519–20; and social justice 456–65; and the state 368; well-being 332
capability metric, of well-being freedom 459
capability theories of justice 463–5
capacity, and moral status 556–7
Capital 135, 139, 267, 270, 271, 274
capital: human and money 76–7; and the state 269
capitalism: and the autonomous state 270; and environmental issues 299; and exploitation 139, 272, 274; and general commodity production 271; industrial/anarchism 218; late/free market regulation 166; Marx 138–40; and performativity 669; and postmodernism 674; transnational 672–4
capitalist anarchism 223 *see also* anarcho-capitalism
capitalist economy, Hegel 121
capitalist exploitation 274
capitalist mode of production, and the state 270, 271
capitalist societies, and entitlement theory 430
capital logic school 268, 270
capital-relation, as domination 272
care, and respect 796, 798
caste system 197–8
Castoridis, Cornelius 162
Categorical Imperative, and Universal Principle of Right 105–6
cave image, Plato 9
Cavendish, Margaret 282, 283
censorious tolerator 634
censorship: Mill 125; state/and public sphere 164
Centesimus Annus 33
Cercle Proudhon 170
chain connection 407
chance, deserving a 435–8, 440
character, and desert 439–40
charity, and moral rights 414
Charity Organisation Society 153, 155–6
Charles VIII of France 47
Chatterjee, Partha 202
childhood, sociology of 729
children, moral status of 553
children with disabilities, education of 723
China, political philosophy 181–90
choice: and avoidable responsibility 392; and customs 126; and evaluation of good life 127; free/autonomous 577; and paternalism 796, 797–8; and responsibility 397; voluntary 791
choice-directed action, and freedom 564, 568
choice theory of rights *see* will theory of rights
choiceworthiness, human life 343–4
choicism, and luck egalitarianism 389–91
Christian Church, and misfortune of his times/Machiavelli 51–2

Christian Democracy, Aquinas 29
Christian Democrat parties 33
Christian faith, and British Idealism 152
Christianity: Hobbes 69; Machiavelli 49, 52; Mill 125
Christiano, Thomas 551–60, 603
Christian thought, medieval/anarchism 217–18
Christman, John 578, 582–3
Church, Machiavelli and the 52
Cicero 41
circumstances of justice 697
citizenship: global/human rights 503; Marsiglio 42; and political equality 611; and republicanism 259–60; responsible and PJP 803–4; restriction of/Aristotle 20, 23
citizen virtues 21, 365
city(ies): Aristotle's ideal 20; non-ideal/wicked/virtuous/ignorant 210–11; Plato's just 7–10
civic humanism 255, 258, 262
civic republicanism 613
civic virtues, and republicanism 259
civil collective reasoning 616
civilians, killing of in war 545, 546
civilization, and altruism 221
civil law 68–9, 621
civil liberty 573
civil society: in capitalist economy 121; and democracy 614; everday communication in 714; France 177; global/human rights 503; Gramsci 162–3
civil state, and securing rights 105
Civil War, American 222
civil wrongs 774
claims: and duties 154, 618, 620; and liberties 618–19; and rights 618
class, and education 723
class divisions, Marx 137
classes, Plato's just city 7
class states 86
class struggle: and anarchism 219; history of 137–8
clean hands argument 542
Clement VII 47
Cleyre, Voltairine de 222–3
climate change 299, 300
close-knitness 403
Club of Rome 292
coaction, property of/rights 378
Coady, C.A.J. 777–86
codified legal systems 36
coercion: and autonomy 577, 581; extra-economic/Marx 272; and faith 30; and freedom 565, 569; and human rights 498; and religious citizens 805; state 313, 349, 366, 372, 803, 804; state/and deliberation 614; and Universal Principle of Right 105; and unmet needs 448
coercive laws 806–7
coercive social institutions 313, 314
Cohen, G.A. 138, 275, 460, 756
Cohen, Joshua 602–3, 613, 759, 761
Cohen, Joshua & Sabel, Charles 604
cold war era: France 174–5; negative libertarianism of 154; and power 585
collaboration, intellectual/WW II 173
collateral deaths of innocents 545
collective identities 492
collective self-determination 490
collectivist anarchism 223
colonialism: and Indian political theory 200–2; and liberalism 233
colonization 505
Combat 175
combatant/non-combatants 782 *see also* soldiers
commensurability, and need 451
commerce, modern and feudal system 101
Commission on the Social Determinants of Health 735
commitment, and existentialism 174
commodities, people as 673
commodity fetishism 271
common good: and democracy 610, 611; and justice 40; Medieval thinkers 37, 38, 44; and natural law 380; and personal good 151; theory of 151
common knowledge 698
Common Law, The 659
commonwealth, Hobbes 60–1
communal functionalism 45–6
communal life, and mutual advantage 41–2
communal order, organic principle of 44–5
communication: everyday in civil society 714; and institutionalized discourse 713–14; media-based mass 714; non-rational/deliberation 615
communicative action 706–7
communicative rationality 166
communism: Marx 137, 142–3; Plato 7, 8
Communist Manifesto 137
communitarianism, and multiculturalism 178
communities 430
community: growth/improvement 662–3; and organism 40
Community of Citizens 178
community standards, and intergenerational distributive justice 467–8
comparative equality 558–60
comparative justice 529, 533, 534
comparison, and liberalism 237
compassion: cultivating 184–5; as social glue 183–4
compensation: and civil wrongs 774; and desert 436–7; and luck egalitarianism 398; and rights 626; and territory 533; and unowned natural resources 415, 416, 522 *see also* rectification
competition, and natural law 381
competitiveness, and social/political ills/Daoism 188
competitive stability 649
complacency, overcoming 54
complexity, modern societies 165, 166, 167
compossible rights 620
comprehensive doctrines, and conceptions of the good 636
Comte, Auguste 130
conceptual approach, pluralism 359–61

Conciliar Movement 29
Conciliation With the Colonies 219
Condorcet jury theorum 689
Condorcet's Paradox 679, 682
confession, obligatory annual 30
Confessions 82
conformity, and autonomy 579
Confucians, and modernity 183–7
Confucian teachings, and large states 189
Confucius 183, 186
congress of states, Kant 112
connection theory of remedial responsibility 523
Connolly, James 484
Connolly, William 356
Connor, Walter 483
conscience 808–9
conscientious objectors, US law 808
conscription, and individual rights 540
consensus, formation 715, 716
consent: and contractarianism 306, 307; and democratic decision-making 601; individual/legitimate authority 600; normative 601; and political authority 601; sexual 747; and State coercion 313
consent theories, and legitimate authority 600–1
consequentialism: and democracy 603; and moral status 552; and utilitarianism 329–39
consequentialist justifications, for punishment 769–70
conservation duties, rights-based 417
conservatism 243–52; and family life 751; and sexual morality 746–7
conservative pluralists 357
conservatives, Confucian 183–7
consistent Manchesterism 222
Constant, Benjamin 177, 255
constitutinal monarchy, and natural law 383
constitutionalism, Sunni 207–10
constitutional patriotism 715
Constitution of Equality 603
Constitution of the Lacadaemonians 20
constructive argument 403
consumerism: and postmodernism 668, 674, 677; and transnational corporations 672
Consuming Life 672–3
consumption: and distributive justice 473; and postmodernism 672, 674
contestation, popular 613
contestatory democracy 164, 261, 262
contextual pluralism 353, 357–9
contingency strategies 698
contingent pacifism 542
Contract and Domination 280
contractarianism 282, 305–15
contractarian jurisprudence 313–14
contract, notion of 309
contract rights, Kant 110
contract tradition, feminist critique 282
contractualism, and political liberalism 317–27
Contribution to the Critique of Political Economy, A 137
Contribution to the Rectification of the Public's Judgement of the French Revolution 219
convention, a 698
convergence, and public justification 806–7
cooperation: and anarchism 226; and contractualism 323; evolution of large-scale human 652–3; and Rational Contractarianism 311; voluntary/compulsory 149
cooperative game theory 703
co-operative practices, and justice 526–7
cooperatives, worker 758
correlated equilibrium 698
corruption: of clergy 51–2; republican democracy/France 170
cosmopolitanism: and global citizenship 503; and human rights 494–504; and multiculturalism 510
cost-effectiveness analysis (CEA) 739, 740
cost principle 220
Council of Lyons 30
Council of Trent 32
counterfinality 144
courts, and inquiry 659–60
craving, as cause of suffering 198
crime, protection against/anarchism 221
criminal law: contractarian approach 314; and will theory of rights 621
criminal wrongs 774
crisis of civilization 172
criteria route, capabilities for justice 461
critical analysis, and discourse theory 713–14
critical-level utilitarian rule 685
critical reason, constructive/destructive sides of 159
Critique of Dialectical Reason 174
Crito, The 5, 307–8
Crone, Patricia 206
Cronin, M. 382
Crowder, George 353–62
Crow nation 505, 511
cruelties well used 53, 54–6
cultural attachments, and liberal multiculturalism 512
cultural discrimination 262
cultural diversity, and individual rights 511
cultural evolution 651–3, 703
cultural groups: national groups as 492; self-determination for 490, 492
culturalist approach, pluralism 357–8, 359
cultural minorities: and group-differentiated rights 506–7; and patriarchy 510
Cultural Revolution, and new humanity 176
cultural social fabric, Hegel 119–22
cultural transmission 653
culture: and genes/co-evolution of 652; and incommensurable values 358; minority/dominant 506; and nation 492; public 488; societal and autonomy 491; societal and nationalism 490–2
customs, and choice 126

Dahl, Robert 588, 759

danda 193
Dandekar, R.N. 192, 195, 197
Daniels, Norman 732–42
Daoists 187–9
Darwall, Stephen 343, 372
Darwin, Charles 147, 149, 643, 645, 646
Darwinism and Politics 646
Darwinsche Theorie und die Sprachwissenschaf, Die 643
Dasein 574
Dasgupta, P. & Heal, G. 477
Dasgupta, Partha 468–70
Dawkins, Richard 653
death, for thought-crime 12
Death of Environmentalism, The 300
Debray, Régis 177
decision making: collective 679, 689; under uncertainty 693
decision theory, normative/descriptive 691
De Cive 59, 62
decoupling 294
deep ecology 295–7
defeasible presumption 713
Defense of Anarchism, In 227
Defense of Idealism 156
defense of persons, and war 539–40
Defensor pacis 40, 41, 42–3
definitional derivations, Hobbes laws of nature 65
Déjacque, Joseph 223
deliberation: and non-rational communications 615; public/discourse theory 714; real-world/and power 616; and state coercion 614
deliberative democracy 168, 602–3, 611–12, 613, 614–15, 616, 657
deliberative politics 167
democracy 608–16; agnostic 356; Aristotle 20; Athenian 4, 9; and common nationality 489; and consequentialism 603; and contestation 164; deliberative 168, 602–3, 611–12, 613, 614–15, 616, 657; discourse theory of 712; and end of politics 160; and the Enlightenment 158–9; and famines 520; and governing economic enterprises 759–60; Madisonian view 610; majoritarian 429–30; and modern society 165; and nationalism 485, 488–9; participative/deliberative 611–13; Plato 9, 11; pluralist view 610; and political autonomy 573–4; and positive totality 166; and power 585; and pragmatism 657; proceduralist social epistemology of 603–4; radical 160–1, 614, 715; real 165; republican 170, 260–1; Rousseau 89; schoolyard view 608–9; and scientific inquiry 662; and Second Vatican Council 33; and subjective will 122; substantive 167; and toleration 635; two-track model 167–8; Weber 165; Western/and environmental issues 299; workplace 761
Democracy and Reaction 155
democratic decision-making, and consent 601
democratic deliberation 663
democratic governance, political sphere 759
democratic ideals, and legitimacy 320
democratic iterations 614
democratic legitimacy 602–4
democratic politics 165–8, 166, 613
democratic principle, and laws 713
democratic procedures, authority/legitimacy of 689
democratic republics, Machiavelli 57
democratic theory: of multiculturalism 513; and public value 739–40; rationalist 714–16; and voting power 592–4
De moneta 43, 45
De motibus animalium 42
De officiis 41
deontological justifications, toleration 632–3
deontological theories 307
deontology, Hobbes laws of nature 65
dependence thesis 597
Derrida, Jacques 176
Descent of Man, The 645
descriptive realism 538
desert 433–42; and distribution 450; as promissory notion 437; and punishment 771
deserving a chance 435–8, 440
de-Shalit, A. 468
desire based derivations, Hobbes laws of nature 65
desire-fulfillment conceptions, well-being 331
desires: first/second order 577, 578, 580, 581; and needs 449–50
desire, satisfaction of/and righteousness 194
despotism, Kant 111, 112
destiny, taking control of 49
detachment, freedom of/capitalism 139–40
deterrent, punishment as 767, 770
de Tocqueville, Alexis 177
de-universalization, of liberalism 235
Deutscher, Penelope 278–88
developmental thinking: and evolution 644
development, understanding 519
Devin du village, Le 83
Dewey, John 657, 661–4
dharma 192–3, 194
dialectical method of enquiry 144
dictatorship of the proletariat 143
difference principle 401–10; and the disabled 458; and justice 324; and principles of desert 438–9
difference, recognition of 513
Differend, The 669, 675
differentia specifica 140–1
Different Senses of "Freedom" as Applied to the Will and to the Moral Progress of Man, On the 151
Diggers 218
dignity: and capabilities 520; and desert claims 439; and human rights 494; Kant 110, 558; and liberalism 512; and markets 797–8; and paternalism 791–4, 798; and rights 558; and status of persons 552, 553
Dimova-Cookson, Maria 146–56
Dionysius II 10
Dionysius the Areopagite 26
Diprose, Rosalyn 288
direct action politics, and environmental activism 297

direct consequentialism 330
direct intent 545
dirty hands problem 336–7, 785–6
disability: and health 734; and rights 619
disagreement: face-to-face/dynamics of 615; reasonable 346, 360, 362
disapproval, and toleration 630
Discipline and Punish 175
discourse: genres of 675; political 716; and political legitimacy 712–13; rational 708–10; as system of social meaning 715; types of 710–12
Discourse on the Origin of Inequality 83, 84, 86, 88, 89, 90, 218–19, 573
Discourse on the Sciences and the Arts 83, 97
Discourse Principle (D) 712–13
Discourses 48, 49, 51, 53, 54, 57
discourses: ethical/moral 711; pragmatic 710
Discourses on Political Economy 489
discourse theory 706–16
discrimination: caste system 198; principle of/and terrorism 783
disenchantment of disenchantment 159
dislike, and toleration 630
dissidents, Soviet Bloc 176
distance, and moral significance 518
distribution: contribution/needs principles 143; and desert 441; of goods/services 412; of income/weath 427–8; justice in 756–7 *see also* distributive justice; and need 445, 446, 451–3
distribution problem with absolute shares 692
distributive justice 325; Aristotle 703; and bounded world 527; in education 722, 728–9, 730–1; intergenerational 467–78; liberal theories of 468; moral rights 412; and needs 444–54; and rational choice theory 695; sufficientarian/egalitarian 559; and terrorism 784, 785; theories 418
distributive models 473
distributive principles, health 738–41
distributive rule, capability theories of justice 464
divergent preferences 758
diversity: cultural and rights 511; and liberalism 241; and multiculturalism 506; and toleration 636–7
divine command doctrine, law of nature 65, 73
Divine Law 28
divine natural order, and the political/India 192, 193
divine providence 28
doctrinal differences, and toleration 636–7
Doctrine of Double Effect (DDE) 30–1, 544–6
Doctrine of Right 103, 105, 109
doctrine, regulation of/Hobbes 67
domestic analogy, national self-defense 539
domination: capital-relation as 272; and freedom 570, 583; identifying 261; and interference 254; and power-relations 273; and republicanism 257; and structural relations 257–8
Dominick, Saint 52
Doppelsatz 115
Dowding, Keith 589
Doyal, L. & Gough, I. 447–8
Dreyfus Affair 169–70
dual systems theory 287
Dunn, John 485
Durkheim, Emile 169
duties: and claims 154, 618, 620; and meeting needs 449; and moral rights 414; positive/negative 517–22; and severe poverty 524; and theory of justice 465; unperformable duties of non-trespass 415
duty based derivations, Hobbes laws of nature 65
duty, immanent doctrine of 119–22
Dworkin, Gerald 577–8, 581, 582
Dworkin, Ronald: and distributive justice 695; envy test 462–3; hypothetical insurance model 399; and justice/equality 464–5; and legitimate authority 600; and majotitarianism 239; and option luck 392–3; and paternalism 788, 789; and rights 337

Earth First! 297, 298
Earth Liberation Front 297, 298
eco-authoritarianism 293–4
ecocentrism 295–7
ecofeminism 296–7
ecological crisis, and libertarian municipalism 296
ecological modernization theory (EMT) 294
ecologism 299
ecologists 291
ecology: deep 295–7; and republicanism 263
Ecology Party 297
economic activity, Aristotle 22–3
economic conditions, role of the state 155
economic enterprises, worker participation in management/governance 759–61
economic game, as positive sum game 76–7
economic growth: and environmental issues 299; and health investment 736
economic health, of community 43–4
economic life, and self-interest 97
economics: Aristotle 16–17; and natural law 381; and politics 16–17; and Rational Choice Contractarianism 310–13
economic socialism, and moral socialism 153
economic stability, and morality 186
economic structure, and legal/political institutions 137
economic systems: and civil society 121; and conservatism 252
economic theories, and psychological egoism/utility 311
economic theory of rational agency 310–11
economy, and natural law/rights theory 379–82
education 721–31; Aristotle 20; and inequalities 402; Plato 7, 9, 10, 12; and regulation of doctrine/Hobbes 67; Spencer 150; of women 281, 283
educational achievement, and socioeconomic disadvantage 723
educational egalitarianism 723, 724
educational standards, and liberalism 235
efficacy claim, and pragmatic discourses 710

egalitarianism: and conservatism 251; as constraint on beneficial outcomes 408; educational and meritocracy 723; and equality 559–60; liberal and work 759–61; and luck egalitarianism 389–91; and multiculturalism 508, 510; and natural law 383; in politics/Confucianism 186
egalitarian justice 390, 391
Ego and Its Own, The 221
electoral democracy, and republicanism 261
electoral participation, and environmental activism 297
Elements of Law 61, 64
Elements of the Philosophy of Right 115
elite discourse, media based 714
elite(s): Chinese intellectual/ruling 181; Confucians 186; Daoists 188; and deliberative democracy 615
elitist concept, of democracy 610
Elster, Jon 144, 579, 581–2
emancipation: homosexual 223; and liberalism 232
embodiment, and sexual differation 288
Emerson, Ralph Waldo 292
emigration, and left libertarianism 416
Émile 83, 84
emotional experiences, and normative force/pragmatism 665
emotion, and liberalism 235
Empire, and the nation-state 273
empirical-theoretical truth, validity claims 710
empiricism, and crisis of Christian faith 152
employment, and women 284
emptiness objection, moral law 108
enabling, people to meet needs 449
ends, of society 468
enforcement, of global human rights 497
Enfranchisement of Women, The 284
Engels, Friedrich 135, 757
English Civil War 59
engulfment, liberation from/capitalism 139–40
Enlightenment, the 114, 116, 158–9
"enough, and as good" proviso 76, 77
Enquiry concerning Political Justice 220
Enquiry concerning the Principles of Morals, An 93, 96
entitlement, and desert 442
entitlement theory 428–9, 430
entrapment 186
entrenched needs 447
environmental activism 293, 296, 297–8
environmental crisis, discourse of 299
environmentalism 291–301; and Confucianism 185; death of 298–300; France 177; moral 348–51
environmentalists 291
environmental justice, and republicanism 263
environmental movement 297
environmental philosophy, and environmental practice 300
environmental preferences 300
environmental skepticism 298–300
envy test 462–3
Epicureans, and anarchism 217
Epicurus 96
epidemiology of representations 653
epidemiology, social 735
epistemic proceduralism 712
equality 551–60; and Buddhism 199; conditions for 115; educational 722–5, 726–7, 729; and freedom 155; gender/Confucianism 185–6; and hedonistic pleasure 390–1; and idealism 115; and justice 464–5; and liberal deliberativism 612–13; and liberalism 232; and libertarianism 418; and luck egalitarianism 390, 391; and Moral Law/Kant 110; of opportunity 401, 534, 729; political and citizenship 611; principle/minority groups 491; and toleration 634–6
Equality of What? 458
equal status 551–2, 554–8
equal treatment 551–2, 558–60
equilibrium selection 699
Espirit 172
Essay Concerning Human Understanding (ECHU) 74
Essay on the Principle of Popultion, An 292
Essays on the Law of Nature (ELN) 73, 74
Estlund, David 601, 603, 615–16, 689
eternal France 170
Eternal Law 28
ethical-existential discourses 711
ethical hierarchism 295
ethical ideas, social institutions as 153
ethical individualism 422
ethical life, Hegel 120
ethical-political discourse 710
ethical, the and the good 513
ethical theory, and political theory 371
ethics: of care 366, 795–6; of complementarity 296; family/state 149–50; Islamic socio-political 205; and philosophical liberalism 240; and politics 106, 370–1; and respect 795; of responsibility 175; from revolution to/France 177
Ethics of Nationalism 487
ethnic cleansing 498, 532
eudaimonism, and human flourishing 367–8
eudaimonist approach, political philosophy 367
European Community, human rights 498
European Union: border crossings 533; voting arrangements 593
Europe, Green Party 297
evaluation: and consequentialism 329; and difference principle 438; of intergenerational social choice 473–5
Evelyn, John 292
evil, and conservatism 249, 250, 251
evolution: competitive/cooperative 646–7; cultural 651–3; and game theory 649–51; mental 647; social 643–54; and spontaneous order 647
Evolution and the Theory of Games 649
evolutionarily stable strategy (ESS) 650
evolutionary game theory 650, 651
evolutionary sociology 647
excellence, and distribution of educational resources 727–8

executive power, civil society 121
existentialism 173, 174
Existentialism and Humanism 173
exit, right of 424, 426, 427, 509
expected utility theory 685
experience, recalcitrant 658
experimentation, and pragmatism 665
exploitation: and capitalism 139, 272, 274; and free markets 417; and justice in production/distribution 756–7; Marx 272, 275, 276
Expression of the Emotions in Man and Animals, The 645
extended self-defense 540
external rights, and benevolent system 94
extra-economic coercion 272

Fabianism 155
Fabian Society 148
Fable of the Bees 95
Faintest Passion, The 579
fairness: justice as 131, 322, 409, 459, 598, 693; and Moral Law/Kant 108; principles 689; and punishment 315
fair process, and healthcare 740
family: abolition of/Plato 8; Aquinas 27; and cultivation of compassion 184–5; and educational inequality 729; and justice 282; as liberation 120–1; marriage, sex, and family 744–6; and natural law/rights theory 379
family life, and privacy 745
family state/state of families 750
Fanon, Frantz 176
al-Farabi, Abu Nasr 210
Farrelly, Colin & Solum, Lawrence 368
fascism, virile 172
Fascist state, and anti-democracy 165
faultless disagreement 346
feasibility condition, personal welfare 307
Feinberg, Joel: and autonomy 347, 348; and desert 434, 435, 436, 439; and paternalism 789, 794; and voluntary choice 791–2
Feldman, Fred 338
felons, moral status of 553
Felsenthal, Dan & Machover, Moshé 592
feminism: and anarchism 226; and autonomy 583; ecofeminism 296–7; and the family 751; France 177, 285; and history of political philosophy 278–88; and marriage 744, 750; and multiculturalism 509–10; and power 589; and rape 747; and sex/family 744
feminist issues, and republicanism 263
Feminist Readings of Aristotle 287
Feminist Readings of Immanuel Kant 287
Festenstein, Matthew 663, 664
feudal structure, China 183
feudal system, and modern commerce 101
Fichte, Johann Gottlieb 219
Fifth Republic, French 175
Filmer, Robert 75
finders keepers 702
Finkelstein, Claire 305–15
Finnis, John 33–4, 749
First Civil War, Muslim 206
Fisher, R.A. 649
Fixation of Belief, The 658, 665
Fleurbaey, M. 689
Flies, The 173
Flikschuh, Katrin 562–71
Florence 47, 53, 56
folk epistemology 665
Foot, Philippa 364
Foreman, Dave 297
Forget Foucault 674
formative arguments 761
Forms, the 7–8
forts, Saptanga theory of the state 196
fortune, conquering 50
Foucault, Michel 175, 176, 273
Fouillée, Alfred 169
foundational rights, and libertarians 414–15, 418
framing characterization 323
France: and American culture: 172; bourgois philosophy/society 170, 171, 172; cold war era 174–5; feminism 177, 285; general strike 170; globalization 178; liberalism 177; May 1968 demonstrations 176; New Philosophers 176; occupation of 173; Resistance 175; social provision 178 *see also* French
Francis, Saint 52
Frankfurt, Harry 447, 450–1, 577, 578, 579, 580, 582
Frankfurt School 165, 167
Fraser, Nancy 490
Frederick II 25
free choice, and free reasonable choice 239
free contract 152
Freeden, Michael 231–41
freedom 562–71; and authority/Rousseau 87, 90; and autonomy 583; and the Enlightenment 114; and equality 155; of expression 125; Hegel 116–19, 120, 121–2; individual 254; and individuality 127; and intellectually active people 125–6; juristic/free 152; and the law/Locke 71–2; Marx 143; and Moral Law/Kant 109; natural/civil/moral 90; negative/positive 150, 152, 563–6, 567, 570, 571; as non-domination 569–71, 583; as non-interference 564, 568–9, 570; personal 149; and power 592; and property/Locke 72; pure negative 562–3, 567–9, 571; and right of exit 424; and rule of law/individual responsibility 175; and state authority 149; and state role/modern society 159; theory of positive freedom 152; true/positive 150
freedom physicalism 564
free human activity 117
free-love movement: and anarchist movement 220, 222; Armand 223
free market(s): and anarcho-capitalists 225; environmentalism 294; in late capitalism 166; libertarian movement 224, 227, 417
free moral action, Hegel 118

free trade, Adam Smith 100
free way of living 421, 422–3, 431
free will: and autonomy 573, 577; as causality 108; Hegel 114–15, 116–17; Machiavelli 49, 50
French Communist Party (PCF) 171, 176
French political thought 170–1, 175–7, 176
French Revolution 90, 177, 243
Fried, B. 429
Friedman, David 224–5
Friends of the Earth 293, 297–8
Frohlich et al. 696
fulfillment, natural law 377
Fumifugium 292
functionings 591
Furet, François 177

Galston, William 361
game theory: and anarchism 226; and evolution 649–51; and social contract 696–703; theory of repeated games 702
Gandhi, Mahatma 201, 202
Garrison, William Lloyd 221–2
Gatens, Moira 283, 287
Gauchet, Marcel 177
Gaudium et Spes 33
Gaus, Gerald 337, 603, 803
Gaus, Gerald & Thrasher, John 643–54
Gauthier, David 305–6, 433, 698, 699
gay rights, France 177 *see also* homosexuality; same-sex marriage
GDP as measure for development 519
Gellner, E. 484
gender, and deliberative democracy 615
gender equality: American anarchists 222; Britain late 19c/early 20c 155–6; and the Qur'an 205
General Idea of the Revolution 220
general strike, France 170
general will: and democratic theory 602, 611; Rousseau 86–9; theory of 153; and will of all 88
genes, and culture/co-evolution of 652
Genetical Theory of Natural Selection, The 649
genetic information, as natural resource 418
genetic inheritance, and human rights 419
Geneva Conventions 542
Genevan state 82, 89
genocide 498, 538
George, Robert P. 368
Geras, Norman 275
German Idealism 159, 160
Germany, Green movement 297
germ-line genetic information, as natural resource 418
Gewirth, Alan 501–2
al-Ghazali, Abu Hamid 208–9
Gibbard, Allan & Satterthwaite, Mark 686
Gide, André 172
Gilbert, Margaret 600
Ginsberg, Morris 647
Gintis, Herbert 650
Glaucon, social contact argument 5
global citizenship, and cosmopolitanism 503
global civil society, and human rights 503
global environmental crisis 293
global fund, left libertarianism 416, 418
global governance 273, 500, 604
global institutional order, and severe poverty 521
globalization: France 178; and the nation-state 273; and political authority 604
global justice 453, 517–24, 527, 529
global legitimacy, and human rights 605
global poverty 453, 521
global public policy, and basic needs 454
Global Resources Dividend (GRD) 522
Glucksmann, André 171, 177
God, authority of/Locke 73
God's will, as source of normativity 600
Godwin, William 219–20
Golden Rule 702
Goldie, M. 74
good: choice of 360; conceptions of/comprehensive doctrines 636; conceptions of/contractarianism 306; concept of the 377; and consequentialism 331–2, 552; and the ethical 513; and evil/Hobbes 68; and marriage 749; and the right 502
Goodin, Robert E. 336, 338, 778
good life: libertarian framework of 425–7; and pluralism 369; and republicanism 255, 258, 259; and role of state 371; and toleration 632
goodness of ways of life, assessing 126–7
good will, and morality 150
Gorgias, The 4, 5
governance: agential/autonomy 576; democratic/economic enterprises 759; global 273, 500, 604; global institutions 604; over area of land 532; self-governance 577
government: absolute authority of/Hobbes 60; by absolute monarch 59; and allocation of power 592; best forms of/Aquinas 29–30; as freedom-curtailing 565; Locke 78–80; and moral responsibility/meeting needs 448–9; and natural law 383; philosophical foundation for 27; Plato 4; republic as/Kant 111; self-government/minorities 507–8; and sovereign/Rousseau 89–90; and traditionalist conservatives 246
Government of Poland, The 489
Gramsci, Antonio 162–3, 267
Gray, John 359, 567
Great Awakening 801–2
Great Betrayal of 1927, The 171
greatest happiness for the greatest numbers 94
Great Society 648
Great War (WW I), and French political thought 170–1
Greenawalt, Kent 808
Green, Leslie 597
Green Party, The 297
Greenpeace 293, 297–8
Green, T.H. 147, 150–2
Gregory, Saint 49

Grhastha 194
Griffin, James 500–1
Grotius, Hugo 93, 383, 384
group-conscious policies 512
group-differentiated rights 506–7, 508, 511
group identity, and violence 510
group membership, liberal model of 512
group protection, justification 513
group selection 645
Grundrisse 270
Guesde, Jules 170
guided variation 652
Gulag Archipelago, The 176
Gutmann, A. & Thompson, D. 613
Gutmann, Amy 750

Habermas, Jürgen: and democracy 162, 163, 164, 166, 168, 603; and discourse theory 706–16; and substantive democracy 167
hadith-reports 207
Halsey, A.H. 676
Hammond, Peter 684
Hampshire, Stuart 336
Han Fei Zi 189
happiness: and free way of living 431; human and Active Intellect 211; human and reasoning 17; as immortality of the soul 211; libertarian framework 425; pursuit of 74, 75
Hardt, Michael & Negri, Toni 273
harm: direct/collateral 545–6; by misrecognition 490; and negative duties 521–2; notion of 349; to people and property rights 424; and punishment 773; third-party 747; and unmet needs 447
harmony and equilibrium, classical Indian thought 194
harmony principle 236, 240, 350
harm principle 349, 383
Harm to Self 789
Harsanyi, John 685–6, 691, 692, 694, 695
Hart-Devlin debates 795
Hart, H.L.A. 129
Hashimite Shi'ism 207
hawa 205
Hawk-Dove game 650–1, 701–2
Hayek, F.A. 427, 433, 647–9
Haymarket trials 222
Hayy, Ibn Yaqzan 211
health 732–42
health investment, and economic growth 736
health needs, meeting 738–41
health promotion 738
hedonism 331, 338
Hegel, Georg W.F. 90, 114–23, 123, 159, 288, 571
Heidegger, Martin 158, 161, 574
Held, Virginia 780, 781, 784
hereditary monarchy, Hegel 121
heredity, and natural law 383
heresy, Aquinas 30
Hershovitz, Samuel 602
Hicks, Edward 507
hierarchy of preferences/desires 577
higher-order self-governing policies 576
Hill, Claire 791, 793
Hinduism 200, 202
Hindu theories, and caste system 200
Hirsch, Joachim 271, 272
historical claims, to territory 532–3
historical conception, autonomy 581–2, 582–3
historical entitlement, justice 558
historical materialism 275
historical texts, destruction of 49
history: and conservatism 243; Marx's theory of 136–8; of philosophy/new interpretations of 286; speculative/Rousseau 84–5
History of Women Philosophers 288
Hobbesian inheritance rule 702
Hobbes, Thomas 59–69; equal status 554; and freedom 564, 570; and human sociability 93–4; moral virtues 365; and ownership of indivisible goods 702; and political authority 599; and power 588; Rational Choice Contractarianism 310; and sexual equality 279–80; and state of nature 218
Hobhouse, L.T. 148, 152, 154–5, 236, 647
Hobson, J.A. 147, 236
Hodgskin, Thomas 220
Hohfeldian incidents 619
holism, and pragmatism 656
Holloway, John 272
Holmes, Oliver W. 659–61
Holmgren, Margaret 438
homosexuality 223, 746, 795 *see also* gay rights; same-sex marriage
Hooker, Brad 332, 335
Hooker, Richard 32
Hook, Sidney 663
Horkheimer, Max 165
household, managing and economics 16–17
households, and *polis* 16
Hsieh, Nien-hê 755–62
human agency 150–1, 162, 448
human, as moral concept 184
human beings: function/power of 377; intrinsic purpose of 376–7; as naturally good 86; as political animals 17, 18, 19; as self-promoting 311; as sensual creature/Hegel 116; status of/and animals 555; unsocial sociability of 104
human capital, and money 76–7
human capital development, and education 723
human conduct 151
human development, four stages of/Smith 100–1
human dignity: and capabilities 520; and desert claims 439; and liberalism 512; and right of conscience 809
Humane laws 79
human fallibility, and skeptical conservatives 247
human flourishing: and basic needs 448; and environmental issues 299; and eudaimonism 367–8;

Marx 141–2; and modern industrialism 359; as moral standard 371; and private ownership 370; and state coercion 372; welfare as 234
human good: and perfectionism 342–4; and respect for persons 346–7; Rousseau 84
human happiness: and the good life 423; and liberty 421; and moral sense 94
human interdependence, and liberalism 234
humanism, teleological 296
humanitarian intervention 539, 541
humanity, as social glue 183
Human Law 28
human misery 246–7
human nature 151; and conservatism 248; and the human good 342; and law of nature 74; and moral law 376; and reason/Aristotle 19; universality of/Marx 140
human-nature relationship 293
human/non-human relationship 291
human rights: Aquinas 33; and capabilities 520; and closing borders 534; and cosmopolitanism 494–504; and global legitimacy 605; to healthcare 741; as moral rights 497, 605; and natural rights 500; proper 502–3; protecting 153; rival conceptions 500–3; and severe poverty 521; talk 500; violations and terrorism 784–5
human rights movement, and natural rights tradition 494
humans, as rational agents 64
Hume, David 34, 92–101, 601, 702
Humphrey, Mathew 291–301
Hursthouse, Rosalind 367
Husserl, Edmund 580
Hutcheson, Francis 94, 577
hybrid government 187, 190
Hyndman, Henry 147, 155
hyperreality 667, 670–1

Ibn Bajja 211
Ibn Rushd 211
Ibn Sina 211
Ibn Tufayl 211
Iceland, medieval legal system 225
ideal deliberation, and democratic politics 613
Idealism: absolute 115; and biological evolution 147; British 152–3, 154; German 159, 160; and positivism 155; rejection of/return of politics 160
idealist metaphysics 154
idealized acceptance, and justifiability 317
idealized conditions, Rawls 317
ideal speech situation 167
Identification and Wholeheartedness 579
identifying, with natural world 295–6
identity(ies): collective 492; group and violence 510; and group membership 512; national and global justice 523; shared sense of/and just society 88–9; substantial 120; and toleration/recognition 637; and well-being 490
identity politics 356
If I Were King 692
illiberal components, liberalism 239
illiberal minorities, rights of 508, 509
imagined utopia 425–6
imama 210
Imamate, the 205–8
Iman 205
immanent doctrine of duty 119–22
immigrant people: and group-differentiated rights 511; and multiculturalism 507, 508; priorities of 513
immigration: and comparative justice 534–5; and left libertarianism 416
immoralists, Machiavelli as 48
immunity, and rights 619
impact definitions, terrorism 777, 778–9
impartial spectator, theory of 100, 691–2
imperialism: and liberalism 233; Marxist theory of 274, 276
impious hypothesis 383
implantation, and territorial rights 532
inborn constitution program, law of nature 73–4
inclusion: and Buddhism 199; India's move towards 202
inclusivity condition, and moral discourse 711
income, distribution of 402–3, 427–8
income inequality, and education 723
incommensurable values 353, 355–6, 357, 358
incompossibility of rights 415
independence, and Moral Law/Kant 110
Independence condition 575
Independence of Irrelevant Alternatives 681, 682, 684
Indian political theory 192–202
indifference curves, social welfare function 473
indigenous peoples: and cosmopolitanism 511; and cultural survival 513; and multiculturalism 507; and patriarchy 296–7; and territorial rights 532–3
indirect consequentialism 330, 337
individual autonomy, and value plurality 361
individualism, methodological/ethical 422
individualist anarchism 219, 220–3
individuality: development of/Mill 126, 132; and freedom 127; and general welfare 129; Mill 131–3, 147; oppression of 129
individual, new era of/May '68 protests 176
individual rational choice, and social contract 691–6
individual rationality, theory of 312
individuation 322
indubitable introspectables 61
industrial capitalism, and anarchism 218
industrialism, and human flourishing 359
industrial production, and oppression 172
inequality(ies): and desert 441; and difference principle 406; and distribution of public goods 45; and education 722, 723, 724, 729; and the family 752; health 735; and luck egalitarianism 390, 393; and political upheaval 44; and public health 402; and the Qur'an 205; sexual 279; sources of 401; in starting points 406; of status 553–4; and unmet needs 451; of wealth 155
inheritance, Hobbesian inheritance rule 702

initial just holdings 76
injustice(s): and caste system 198; economic 44; and the family 751–2; and people's conduct 6–7; Plato 4, 6; as violation of rights 100
innocence: return to/Daoism 188; and terrorism 781–3
inquiry: and truth 658; unified theory of 661
Inquiry into the Original of our Ideas of Beauty and Virtue, An 94
institutional virtues 365
institutions: free/choice criterion of value 128; global/international and political authority 604; just/free 130; power-allocating 592
Institutio Traiani 38, 43
insurance model, Dworkin 399
Integrity Objection 804–6
intellectually active people, and freedom 125–6
intellectual property, and anarchism 226
intellectual, role of 171–2
intent, direct/oblique 545
intentionality 580
intention, and punishment 773
interest-group pluralism, and democracy 610
interest(s): and anarchism 220; and paternalism 791–4
interest theory of rights 413, 621–2, 623
interference: arbitrary and republicanism 259; and domination 254; and freedom 563–4
intergenerational distributive justice 467–78
intergenerational saving 468, 469
intergenerational social choice 471–5
intermediate needs 447
internal freedom 570
international commitment/law, and human rights 502
international court, Kant 112
international distributive justice, and severe poverty 517–24
international economic bodies, and severe poverty 521
Internationale Marx-Engels-Stiftung (IMES) 135
international institutions, lack of and war 538, 539
international law: Grotius 384; and human rights 497; and natural law 383; refugee status 534; and war 537, 539
international legitimacy 604–5
International Monetary Fund 604
international relations: and Confucianism 185–6; and human rights 503; and left libertarianism 416; and principle of subsidiarity 384; and war 537–8
international right, Kant 111–12
international rule of law 503
international wealth transfers 416
international world, and toleration 636
internet, and anarchism 226
interpersonal morality 325–6
interpersonal utility comparisons 694, 695
intersectionality theory 280
In the Shadow of the Silent Majorities 676
intolerance 630
intranquility, and the political organism 41
inviolability: founded on justice 153; of natural rights 378
invisible hand 97, 101, 219, 225, 233
Irigaray, Luce 285–6
iron law of oligarchy 165
Islam: and Indian political theory 200; as realistic utopia 213
Islamic legal order, and Sunnism 208
Islamic modernism 211–14
Islamic political philosophy 210–11
Islamic political thought 204–14
Islamic revivalism 212–14
Islamic socio-political ethics 205
Islamic thinkers, and anarchism 217
Italy, division of/Machiavelli 51
ius ad bello 540
ius ad bellum 537, 538–42
ius in bello 537, 542–4

Jacobin project 177
Jainism, and caste system 198
James, Aaron 317–27
Jameson, Fredric 266
James, Susan 287
James, William 657
Jamieson, Dale 300
janpada, Saptanga theory of the state 196
Jasay, Anthony de 427
Jayaswal, K.P. 201
Jeanson, Francis 175
Jefferson, Thomas 219, 802
Jencks, Charles 667
Jennings, Ivor 485
Jennings, Jeremy 169–78
Je Suis Partout 173
Jeunes Gens d'aujourd'hui, Les 170
Jews, and Aquinas 30
John of Salisbury 38–40
John Paul II 33
Jones, Peter 629–38
Jordan, Jeff 749
Jowett, Benjamin 152
judgement aggregation 688
judges, role of 659–60
Julius II 55
jurisprudence: contractarian 313–14; virtue-based theory 368
juristic freedom 152
jury system 689
jus ad bellum 384, 781
jus in bello 781, 782
jus in bellum 384
just acquisition 430
Just and Unjust Wars 785
justice: administration of/civil society 121; Aristotle 22; and borders 526–35; bounded 530; and capabilities 457–8, 460–2, 463; and common good 40; comparative/non-comparative 528–9, 533, 534; concept of 324–5; and consequentialism 337–8; and contactualism 317–27; and

convention 95, 96; in distribution of income/wealth 428, 431; distributive *see* distributive justice; in education 722; egalitarian/and benefits 450; entitlement theory of 428–9; and equality 464–5; as fairness 131, 322, 409, 459, 598, 693; and the family 282; global 453, 522–3, 527, 529; as hands-off virtue 93; and health 732; historical entitlement 558; as human product 101; Hume 94–6, 99; of institutions/rightness of actions 366; intergenerational distributive justice 467–78; and legitimacy 320, 598; and liberalism 240; and luck egalitarianism 395; Marxist theory of 274–6; minimal theory of 461; mixed concept of 409, 410; and needs 446, 454; as negative virtue 99; normative theory of/will theory 620; normativity of 327; between persons/generations 475–8; Plato 4, 5–7; and political philosophy 326; principles of 407; in production/distribution 756–7, 762; and protection of opportunity 737; public conception of 321; Rawls 401; and realistic stability 324; republican theory of 263; and rights 115; rules of 95; scope-restricted 526–7; screening/ranking of principles 408–9; and social development/evolution 93, 96–7; and toleration 636; and universality 104; and virtue 365; as virtue-ethical 366; and womens rights 284
Justice as Fairness 131, 409, 410
Justice as Fairness: A Restatement 2001 598
justifiability: and contractualism 320, 407; and idealized acceptance 317
justification, public 598, 806–7
just revolution 778
just society, social preconditions for 88
just war 30, 384, 537, 539–42, 545, 781–2
al-Juwayni, Abu'l-Ma'ali 208

Kalai, E. & Smorodinsky, M. 699
Kalinga war 199
Kali Yuga 192
Kalyana Malla 194
kama 193, 194
Kamasutra 194
Kamtekar, Rachana 14–23
Kantian Constructivism in Moral Theory 321
Kant, Immanuel 103–12; and autonomous choice 790; and dignity 558; and the Enlightenment 159; feminist readings of 287; and freedom 566; and human rights 494, 534; and justice 620; moral autonomy 573; and natural rights 499; and normative Contractarianism 309; and personhood 555; political legitimacy 597; as positive retributionist 770; and respect 795; and Rousseau 90; social contract 599–600, 601; war 539
Kautilya 193
Kavka, Gregory S. 65
Kekes, John 243–52, 357
Kellogg, Frederic R. 660
Keyt, D. 18–19
Kharijites, the 207
Khomeini, Ayatollah Ruhollah 213
Kidd, Thomas 801–2
killing of enemies: intended/foreseen killings 544; permission to 543–4
kings, qualifications of/Saptanga theory 195–6, 197
Kleinig, John 435
Klosko, George 3–12
knowledge: common 698; Machiavelli 50; and postmodernism 668–9; sociology of 283
Konkin, Samuel E. III 225
Kornhauser, Lewis & Sager, Lawrence 688
Kouvelakis, Stathis 266
Kramer, Matthew 563
Kraut, Richard 21, 23, 371–2
Kristeva, Julia 176
Krita Yuga 192
Kropotkin, Pëtr 223–4
Kshatriya 198
kufr 204
Kukathas, Chandran 422, 424, 430, 505–14
Kymlicka, Will 487, 489, 491, 492, 506–8, 511

Labadie, Joseph 222
labor-exchanges 221
laboring 755
labor movement, and anarchism 224
labor, rights over own 75–6
labor theory of value 271
labor unions 222
Labour Party, UK 148
Lacan, Jacques 175
Laclau, Ernesto 270
Lafargue, Paul 170
laisez-faire politics 146, 147, 190
Lamarck, Jean-Baptiste 147, 149
land: ownership/borders 530–1; private ownership/anarchism 221; symbolic/religious value 532 *see also* territory
Language of Post-Modern Architecture, The 667
language philosophy 585; and power 587–8, 589
Laozi 188
Laqueur, Thomas 278–9
Laqueur, Walter 777
La Rochelle, Pierre Drieu 172
Las Casas, Bartholomé 505
Latin America, Christian Democratic parties 33
lawgiver, the 88
Law of Equal Freedom 221
law of nations, and natural law 383
law of nature: Hobbes 64, 65–8; two programs/Locke 73–5
Law of Peoples, The 636
Law of Politics, The 323
law(s): Aquinas 28; chosen by general will 87–8; civil law 68–9, 621; coercive 806–7; criminal/contractarian approach 314; and democratic principle 713; discourse theory of 712; divine positive laws 69; and freedom/Locke 71–2; humane 79; international 383; of liberty 81;

law(s) (*cont.*):
moral 106–8; municipal (local) 96; paternalistic 788, 798; Plato/compliance with 5; political/Kant 106; positive 382; pragmatist account 659–60; as product of will of law-giver/Locke 73; prostitution 747; public coercive 112; rape 747; and religion 807, 809; and rule 676; US Constitutional 807–8; of war 384, 542–3; and will theory of rights 621
Laws of Ecclesiastical Polity, The 32
Laws, The 10–11, 20
league of nations, Kant 112
LeBar, Mark 372
Lecture on "Liberal Legislation and Freedom of Contract" 151, 152
Lectures in the History of Political Philosophy 130
Lectures on Jurisprudence 98, 100
Lectures on the History of Political Philosophy 128
Lectures on the Principles of Political Obligation 151
Le Doeuff, Michèle 283, 286
Lefort, Claude 162, 164
Left, France/non-Jacobin/second 176, 177
left libertarianism 227, 412–19
legal domination 165
legal inquiry 659, 660
legal institutions, as consensual 313
Legalists 189–90
legality, and morality 106
legal moralism 795
legal rights, and natural rights 498–9
legal systems, emergence of codified 36
legislative authority, Kant 111
legislative power, civil society 121
legitimacy: and authority 596–605; and justice 320, 598; principle of liberty 317–18
legitimate authority 596–8
Leiter, Brian 808–9
Lenin, Vladimir 267
Leopold, David 135–45
Leo X 47
Leo XIII 32
Letter Concerning Toleration, A (LCT) 79, 80
Letters from the Mountains 89
leveling down 559
Leviathan 59–69, 218, 588
Levinas, Emmanuel 177
Lévi-Strauss, Claude 175
Lévy, B.-H. 176
Levy, Jacob 484, 487
Lewis, D. 698
leximin rule 684, 700
lex talionis 772
li 183
liability, and rights 619
liberal autonomy: and culture 491; and nationalism 491–2
liberal deliberativism 612–13
liberal democracies, and environmental crisis 293, 294
liberal democratic theory, and liberal deliberativism 612–13
liberal dogmatism 236
liberal egalitarianism, and work 759–61
Liberalism 154
liberalism 231–41; and conservatism 252; economic 146; France 177; Hegel 122; and human dignity 512; Locke 71; and moral rights 414; and multiculturalism 508–11, 514; and natural law 383; new 146, 150, 154; and paternalism 798; and pluralism 359, 360–1; political 801; political and contractualism 317–27; and positive freedom 563; public reason 803–4; Rawls/Mill 128, 131; and religion in public life 803; and republicanism 256, 258; and sexual morality 747; and social complexity 167; and socialism 155; and social rationalization 166; Spencer 146; with state intervention 154; and virtue ethics 371; as winning ideology 239
liberal multiculturalism 512
liberal paradox 687
Liberal Party 150
liberal pluralists, and personal autonomy 361
liberal principle of state neutrality (LSN) 344–5, 351
liberals, and nationalism 484–5
liberal society, and private/public 744
liberal state, and family life 751
liberal theories of justice, and intergenerational justice 468
liberal theory, and minorities 507
libertarianism 223, 421–31; free-market libertarian movement 224; left libertarianism 227, 412–19; and meeting needs 449; minarchist libertarians 225; right libertarianism 414, 418; and Will Theory 413
libertarian municipalism, and ecological crisis 296
libertarian paternalism 791, 794
liberties: and claims 618–19; Hobbes true liberties of subjects 66–7; non-utilitarian elements 131
Liberty 222
liberty: defense of/Mill 125–7; and freedom 562; Green 152; harmless/Hobbes 66; Hegel 121; Hobhouse 154; and human happiness 421; individual and harm condition 745; and liberalism 231–2, 233, 234, 237; and libertarianism 422–3; negative/positive 255–6; as non-domination 255, 257, 258; and personal development 154; political/republicanism 254, 255, 256; principle of 132; rational system of 101; Rawls/Mill 128; rights 496; Spencer 148, 149; and state power 149; and utility 127–31, 131–2
Liberty, Equality, Fraternity 339
lifestyle choices, and health 738
limitation, of rights 378
Limits to Growth 292–3
linguistic struggle, and postmodernism 675
Liquid Love 673
liquid modernity 672–3, 676, 677
living wage, right to 155
living well: Aristotle 17; Marsiglio 41

livre de politiques, Le 43, 45
Lloyd, Genevieve 286, 287
Lloyd, S.A. 59–69
Lockean proviso 531
Locke, John 71–81; and American Revolution 136; and anarchism 218; and civil government 32; and equal status 554, 558; and land ownership 530–1; natural law theory 375; natural rights 425; and political authority 601; property rights 430; proprietarianism 751; and sexual equality 279–80; and state of nature 62
Lomborg, Bjørn 298–9
Long, Roderick T. 217–27
Lorenzo the Magnificent 47
love, universally and unequally/Confucianism 185
Lucas, J.R. 699
Lucifer the Lightbearer 222
luck: and desert 434; and responsibility 391–3
luck egalitarianism 389–99, 412, 465, 737–8
luckism, and luck egalitarianism 389–91
luck vs. choice filter 390, 397
Lucretius 95
Lukes, Steven 588, 589
Lum, Dyer 222
luxury: and morality 97; and society 101
Luxury, Of 97
lying, Aquinas 31
Lyotard, Jean-François 667, 668–70, 674–5

MacCallum, Gerald 562, 566, 567
Macedo, Stephen 510, 511, 806
Machiavelli, Niccolò 47–57, 258
MacIntyre, Alasdair 34, 357, 364
Mackay, John Henry 223
Mack, Eric 71–81, 425, 430
Madisonian view, democracy 610
Magnesia 11
magnus opus, Being and Nothingness 173
Mahabharata 192, 193, 197
Maha-Parinibbana-Sutta 199
Maistre, J. de 249
majority rule 167; pairwise 682; in political society 78–9
Malthus, Thomas 292, 646
man: and natural law/rights theory 376–8; as political/social animal 27, 40; virtues of ideal 195
Man and the State 33
Manchester Guardian 154
Manchesterism, consistent 222
Mandeville, Bernard 95, 97
Mandragola 48
Manent, Pierre 177
Manifesto of the Communist Party, The 757
mankind, unsocial sociability of 159
Man of Reason, The 286
Manu 192, 193, 197
Man Versus the State, The 149
March, Andrew F. 204–14
Maritain, Jacques 33, 34
market anarchism 225
market capitalism: and work 755, 756; and worker cooperatives 758
market economy, resistance to/France 178
market(s): as allocator of resources 592; and liberalism 233, 234; and paternalism 797–8
market system, and John Paul II 33
marriage 748–50; marriage, sex, and family 744–6; monogamous/Aquinas 27
Marsiglio of Padua 40–3
Marsilius *see* Marsiglio of Padua
Martin, Rex 401–10
Marx-Engels-Gesamtausgabe 135
Marxism: and alienation 757–8; and contemporary political thought 266–76; of French Communist Party 176; and French political thought 170–1, 175, 177; revival of 267; and Sartre 174; and social anarchists 223; wage-labor relationship 755, 756
Marxism and Politics 266
Marxists, and nationalism 484
Marxist tradition, feminist reworking of 287
Marx, Karl 135–45; and alienation 757; capitalist exploitation 274; capitalist mode of production 270, 271; freedom by self-organization 162; and Hegel 123; and metaphysics 246; politics of 158; rejection of modern politics 160–1; and republicanism 258
Massis, Henri 170
mass media, and discourse theory 714
material consumption, and environmental issues 299
materialism, historical 275
material sustainability 468–70
Maurras, Charles 170, 173
al-Mawardi, Abu'l-Hasan 208
Mawdudi, Sayyid Abu al-A'la 213
maximin argument 409, 410, 696
maximin rule 693
May, Todd 667–77
McGraw, Bryan 802
McKeown, Thomas 735
McLean, Iain 299
McMahan, Jeff 543, 786
Meadowcroft, John 421–31
mechanism design 703
media based elite discourse 714
median voter theorum 686
media, rise of and loss of representation 670
Medici, Giulio d' 47
medicine, modern and health 734–5
medieval Christian thought, and anarchism 217–18
medieval Icelandic legal system 225
medieval political thought 36–46
Meech Lake Accord 490
Meikle, S. 16–17
Mele, Alfred 582
meme 653
Memories of Youth 170
Mencius 183–4, 186

meritocratic educational equality 722, 723–4
meritocratic principle, distribution of educational resources 728–9
meritocrats, Confucianism 186
Merleau-Ponty, Maurice 174
metanarratives, and postmodernism 668, 669, 674, 675
metaphor, and history of philosophy 286
Metaphysical Club 656, 659
metaphysics: and conservatism 246–7, 248; ecocentric/environmental activism 296; and Plato 279
Method of Ethics 311
methodological individualism 422
Middle Ages, body metaphor 45
middle path, Buddhism 198
middle way strategy, Confucianism 185–6
Miliband, Ralph 266, 267, 268, 269, 270
Millar, John 98, 99
Miller, Dale E. 329–39
Miller, David 526–35; and desert 436; distribution and need 451, 452; and divergent preferences 758; and nation 484, 492; and nationalism 487, 488–9; and needs, desert and equality 450; and severe poverty 522–3
Miller, F.D. 19
Miller, Seumas 780
Mill, John Stuart 124–33, 147; and happiness 332–3; and human development 581; and liberalism 233–4, 239; and nationalism 489; paternalism 789–90, 791; personal autonomy 573; and rights of women 284; and utilitarianism 234
Mills, Charles 267, 280
minarchist libertarians 225
minimal marriage 749
minimal theory of justice 461
minister, qualifications of/Saptanga theory 196
minorities: cultural assimilation of 513; and equality principle 491; and group-differentiated rights 506–7; and healthcare 735; and multiculturalism 506; national/self-government 507–8; and patriarchy 510
minority rights, external protections/internal restrictions 508
Misak, Cheryl 656–66
Mises, Ludwig von 224
misfortune, as opportunity 48–51
Miskawayh 211
misrecognition, harm by 490
Mitra, R.L. 201
Miyoshi, Masao 672
model of preferences and needs, intergenerational justice 471–3
modernism, and metanarratives 668
modernity 114; and antiquity in China 183; and Confucians 183–7; and cosmopolitanism 511; and German Idealism 160; and India 201; Islamic 211–14; liquid 672–3, 676, 677; as philosophical problem 158; and politics 162
modern societies, complexity of 165
modern society, as a totality 166
modus tollens 93
moksha 193, 194
Molinari, Gustave de 221
monarchical rule 101
monarchy: Aquinas 29; as form of government 89; hereditary/Hegel 122; and natural law 383
monetary policy, Oresme 43
money, and human capital 76–7
monism, moral 354
monopolies, and natural law 380–1
Montesquieu, C-L de S. 98
Moore, G.E. 332
Moore, Margaret 483–92
moral action, free 117
moral agency, and sentiment 287–8
moral agents 120
moral autonomy 4, 573, 574
moral character, and republicanism 258
moral concept, human as 184
moral conception, republicanism 259
moral conscience, Hegel 122
Moral Contractarianism, and Political Contractarianism 305–7
moral costs, failure to act 518–19
moral cultivation, Confucians 185
moral dimension, *Rta* 192
moral duties, Green 151
moral economy 676
moral environmentalism 348–51
moral environments 350
moral equality of soldiers 543, 554
moral freedom 808
moral goodness 86
moral gravity of offense, and punishment 773–4
moralistic legal paternalism 794–5
morality: and anarchism 226; Aquinas 31; and benevolence 93; and conservatism 251; and economic stability 186; and freedom 566; and good will 150; Hegel 119; human 354; Hume 94; interpersonal 325–6; and legality 106; and luxury 97; Machiavelli 52–3; and *pitié* 86; political 342; public and consequentialism 335–7; public/and politics (Sunni) 209–10; and rights 627; same-sex relations 749; as system of laws 376; of terrorism 781–3; as universal/necessary 107; of war 537, 542–3
moral judgment(s): and sympathy 97; and war 538–9
moral justification: deaths of civilians 546; for punishment 769–72; of terrorism 783
moral law 68–9, 106–8, 376
moral monism 354
moral obligations 349
moral order, and conservatism 246, 247
moral philosophy: and rights 627; role of sentiment/moral agency 287–8
moral prohibition, on terrorist acts 786
moral psychology, Rousseau 83–6
moral reasons, for toleration 635–6
moral rightness, validity claims 710

moral rights: constraint view 624; human rights as 497, 605; left libertarianism 412
Morals By Agreement 305
moral sense, and human happiness 94
moral sentiments, Confucians 184
Morals from Motives 366
moral socialism, and economic socialism 153
moral status: children/adults/felons 553; and quality of status 551; restrictive conception of 552–3
moral theory(ies): consequentialist 333–5; utilitarian 312
moral universalization principle (U) 711
moral values, and republicanism 263
moral weighting 518–19
Moriarty, Jeffrey 760
Morriss, Peter 585–94
Morrow, Glenn 11
Moscow show trials 174
motivational system, autonomous agency 576
motive, and punishment 773
Mouffe, Chantal 356, 714–16
Moulin, Hervé 686, 688
Mounier, Emmanuel 172
Mueller, Dennis 472–3, 474
Muir, John 292
mulk 210
Müller, Wolfgang & Neusüss, Christel 270
Multicultural Citizenship 507
multiculturalism 505–14; France 177–8; and pluralism 358
Multiculturalism and the Politics of Recognition 512
multi-layered society 379–80
multinational corporations (MNC) 672
mu'min 205
Mu, Qian 181
Murphy, Mark 380
music, and virtue/Aristotle 20–1
Muslim Brotherhood 213
Muslim political/economic weakness 211–12
mutual advantage, and communal life 41–2
mutual benefit 403–5, 408, 410
mutualism, Proudhon 220
mutuality, and social practices 325
mutual recognition 326
Myth of Sisyphus, The 175

Nadeau, Christian 254–63
Naess, Arne 295–6
Nagel, Thomas 429, 527, 528
naivety, French political thought 171
Najdiyya sect 217
narratives, public role of religion 800–3
Narveson, Jan 424, 760
Nash Equilibrium 650, 651, 697–8
Nash existence theorum 650
Nash, J. 697, 699
Nash problem 701
Nathanson, Stephen 780, 783, 786
national defense 540
national groups, as cultural groups 492
national identity, and global justice 523
national interest, as over-riding good 784
nationalism 170, 483–92, 489
national self-defense 539, 540, 544
nation-building 489
nation(s): and culture 492; as primary agent of global justice 522–3; and states 484, 486; term 483–4; and territorial acquisitions/retentions 416
nation-state: and Empire 273; the liberal/cultural communities 507; rightful/Kant 111; and transnational corporations 672
Native Americans, and cosmopolitanism 511
natural goodness 377
natural inclinations, of rational creatures 28–9
naturalism, and pragmatism 656
natural law: and civil law 68–9; context/content 28–9; convention 702; movement 368; theory(ies) 25, 34, 149, 375–84, 701
Natural Law and Natural Rights 34
natural resources: foundational rights to 418; unowned and compensation 415
natural rights: and equal dignity of humans 378; Green 151; and human rights 500; and legal rights 498–9; Locke 72, 425; theory classical/contemporary 377–8
natural rights tradition, and human rights movement 494
natural selection 645; theory of 147
natural talent, and education 724
nature: and Daoists 188; and patriarchy 296–7; and the *polis* 18–19; state of/Hobbes 62–5
Nature of Things, Of the 95
Naturphilosophie 643
Nausea 173
Nederman, Cary J. 36–46
needs: competition among claims 452; and desires 449–50; and distributive justice 444–54; relative/distribution of goods 338; and responsibilities 448–50; satisfying basic/social 141; substitutable 447; types of 447; as universal preconditions 447–8; and wants 445
negative duties, severe poverty 521–2, 524
negative freedom 567, 570, 571
negative libertarianism, cold war era 154
Negotiating Nationalism 487
neo-Kharijites 207
neo-liberalism 178, 273
neo-Platonic Christian cosmology, and organic discourse 37
Neo-Platonism, and Aquinas 26
neo-republicanism 259, 261, 262
neo-Roman republicanism 255
Neo-Thomist social thought 33
Neutrality, and Arrow's Theorum 682
new French feminists 285
New Left, and minarchist libertarians 225
New Liberalism 146, 147, 150, 154, 155
New Natural Law 34, 375
New Philosophers, France 176
Nickel, Jim 496–7

Nicomachean Ethics 14, 16, 17, 22, 28, 40, 368
Nietzsche, Friedrich W. 184, 288
Nizan, Paul 171–2
No Child Left Behind 723
noema 580
noesis 580
non-action, Daoists 187–9
nonagreement point 699
nonautonomous preference formation, and procedural independence 581
non-combatants, status of 782
non-comparative equality 558–60
non-comparative justice 528–9, 533, 534
non-consequentialist justifications, punishment 770
non-depletionary use, and intergenerational justice 470
non-derivative rights 414–15
nondictatorship 681
non-discretionary rights 378
non-domination: freedom as 569–71, 583; and republicanism 260, 262, 563
non-interference, freedom as 564, 568–9, 570
nonintervention, right of 503
non-public reason 745
non-trespass, unperformable duties of 415
Nordhaus, Ted & Shellenberger, Michael 300
normal justification thesis 597
Norman, Wayne 484, 487
normative consent theory 601
Normative Contractarianism 308–9
normative decision theory *see* rational choice theory
normative realism 538
normative theory of justice, and will theory of rights 620
normativity, and new pragmatists 665
Nostri Aetatis 33
Novarum 32
Nozick, Robert: and anarcho-capitalism 225; and autonomy 575; and distribution 427, 428–9; entitlement theory 430; and the good life 422; happiness 425–6; and hedonism 331; historical entitlement/justice 558; and rights 338, 423–4; and rights modelling 687
Nudge 794
Nussbaum, Martha: capabilities approach 332, 368–70, 463, 464, 517, 519–20; and conscience 808; and social justice 461–2

Oakeshott, Michael 421
objective consequentialism 330, 334
objective-list conceptions, well-being 331
objective-list route, capabilities for justice 461
objective spirit 115
objectivist/paternalistic position, autonomy 583
obligation(s): associative/nations 488; and authority/legitimacy 598–600; eudaimonist account 372; to future generations 468; positive/moral 349
oblique intent 545
Occupation, French 173
Occupy movement 226
Oderberg, David S. 375–84
Offen, Karen 285
Of Property 76
Of the Original Contract 601
Okin, Susan Moller 282, 509–10
Olsaretti, S. 441
O'Neill, Onora 465
On Kingship 29, 30
On Nationality 488
On Representative Government 489
On the Laws and God, the Lawgiver 32
On the Social Contract 599
Open Marxism 272
Open Society and Its Enemies, The 9–10
Operation Backfire 298
opium of the intellectuals, Marxism as 175
Oppenheimer, Franz 223
opportunity, misfortune as 48–51
oppression 129–30, 172, 490
Oppression and Liberty 172
optimism, and conservatism 249–51
option luck 392–3
options, plurality of/autonomy 574
order of actions 647–8
ordinary language philosophy 585
Oresme, Nicole 43–5
organicism 147
organic metaphor, and priority of reciprocity 45
organic model, and political society 41
Orientalism 201
original position: and contractarianism 309, 315, 317; and difference principle 402, 407–10
Origin of the Species 149, 643
Origin of the Work of Art, The 161
orthogenic evolution 155
Oshana, M. 583
outcome power 589

pacifism, absolute/contingent 542
Pagans, and Aquinas 30
Paine, Thomas 219
pairwise majority rule 682
Palmer Raids 224
panarchy 221
Pankhurst, Sylvia 156
Pansardi, Pamela 589
paradox of toleration 631
parallel case argument 759
paralogical science 670
paralogy 669–70, 674, 675
Parekh, Bhikhu 358
parents/parenting: and education 721–2; and life prospects 752; repressive rearing and autonomy 581; rights/duties of 750
Pareto efficiency, and difference principle 405, 407
Pareto Indifference, welfarism 684
Parfit, Derek 332, 461, 478, 559, 685
participant interpretation, protestant/catholic approaches 323

participatory democracy 611–12
particularization of the state 271–2
Pascal, B. 528
passion, and reason 716
Pateman, Carole 279–80
paternalism: and autonomy 583; as care v. respect 796; and distribution according to need 446, 454; moralism, and markets 788–98; and perfectionist politics 351; and prostitution 747; and republicanism 259; soft/hard 790; and the state 349, 350
Path Independence, Arrow's Theorum 681–2
patriarchy: and cultural minorities 510; marriage, sex, and family 744; and oppression 296–7
Pazzi conspiracy 47
Peaceable Kingdom, The 507
peace, and toleration 631
Peaceful Revolutionist, The 220
Peace of Westphalia 801
Péguy, Charles 170
Peirce, Charles S. 656, 657, 658–9
Pence, Gregory 755
Penitentials 30
PEOPLE 297
perfectibility, Wollstonecraft 283
perfection 555; autonomy as 347–8
perfectionism 342–51
perfectionist capability theory 461
perfectionist republicanism 259
performativity: and capitalism 669; and postmodernism 674
permanent interests of man, and just/free institutions 130
Perpetual Peace 494
Perry, Michael 808
persecution, and intolerance 630
personal autonomy 563, 574, 798
personal development, and liberty 154
personal flourishing, and social welfare 154, 155
personal good, and common good 151
personalism, doctrine of 172
personalist civilization, A 172
personhood: and freedom 564; moral status of 555–6
person, the 119–20
pessimism, and conservatism 249–51
Pétain, Marshal 173
Peter, Fabienne 596–605
Petit Conseil 83
Pettit, Philip 255, 258–9, 261, 262, 263, 613
Phaleas of Chalcedon, constitution 21, 22
philosopher-kings 11, 12
philosophical anarchism 227
philosophical anthropology 140–2
philosophical discussion, bringing the masses into/India 202
philosophical imaginary 286
Philosophical Letters 282
philosophical liberalism 236–7, 239–41
Philosophical Rudiments 64
Philosophical Theory of the State, The 153
philosophy of biology 733
philosophy of education 263
philosophy of science 586–7
philosophy of self, and deep ecology 295–6
philosophy of toleration 633–4
philosophy, sexual bias in 283
Philpott, Daniel 485
phronesis 357
Physics 18, 19
pitié 84, 85–6
Pitkin, H.F. 589
pity, as a virtue 184
Pius IX 32
Pius XI 32
Plato 3, 3–12, 210, 279, 365, 750
Plenty Coups 505, 511
pluralism 353–62; and conservatism 247–9; and good life 369; of large states 189; moral 190; reasonable/unreasonable 345; value/reasonable 346
plurality, and pluralism 354–5
plural subject theory 600
plural values, and conflict 355
plurism: and social structure 236; and toleration 235
PMP (minimum necessary level of provision) 474
Pogge, Thomas 453, 521–2
poleis 19
Policraticus 38, 39, 40, 44
polis: and households 16; and nature 18–19
politeia 17, 20
political activism: and radical democratic deliberativism 614; women 156 *see also* activism
political activity, Aristotle 17
political agreement, bargaining approach 308
political arrangements: favoured by conservatives 250, 251; preserving 243; and social practices 325
political authority: and individual autonomy 599; legitimate 604–5; Medieval times 36; and political legitimacy 599–600
political autonomy 573 *see also* autonomy
Political, circumscribing/valuing the 14–18
political communication, and institutionalized discourse 713–14
Political Contractarianism 305–7
political decision making: Confucian 186; Mouffe 715
political deliberation 713, 716
political difference principle 187
political discourse 716
Political Discourses 93
political equality 553, 611 *see also* equality
political freedom 570
political institutions, and global justice 527
political law, and Categorical Imperative 106
political legitimacy: and authority 597, 598–600; Hobbes 60; person-centered approach 605; Samaritan account 601; and the state/globalization 604–5
Political Liberalism 323, 597, 636, 800–1
political liberalism 801
political liberty, and republicanism 254, 255, 256, 260

political morality 342, 345, 350
political naturalism 40
political obedience, and stability 59
political obligation 5, 151, 366, 372–3
political order, organic/Medieval thinkers 37–8
political organism: constitution of 39; and tranquility/intranquility 41
political parties, religious 802
political philosophy, continental 158–68
political power: coercive and the private 745; and contractualism 318; as second-order power 592
Political Power and Social Class 268, 269
political processes, India 202
political realism, and morality of warfare 784
political resistance, and postmodernism 675
political rule, Aristotle 15, 16
political science 14, 60–2, 115, 586
political society: and government structures 78; organic model 41
political sphere, democratic governance 759
political state proper 121
political status definitions, terrorism 777–8
political struggle, and postmodernism 675
political system, Hegel 122–3
political theory: and ethical theory 371; Indian 192–202; and value pluralism 361–2
political thought: British late 19c/early 20c 146–56; French 20c. 169–78; Islamic 204–14; medieval 36–46
political toleration 635
political utopia/utopianism 210–11, 354
political weakness, Machiavelli 49
political writings, Chinese 182
Politics 15, 16, 18, 19, 21, 44, 368
politics: and Aquinas 27; and art 161; deliberative 167; direct action/environmental 297; and economics 16–17; effective/harsh punishments 54; egalitarian nature of Confucian 186; end of/democracy 160; and ethics 106, 370–1; gender equality in Confucian 185–6; and global justice 517–24; and identity 356; idependent realm for 163; of imperfection/conservatism 249; of inclusion 42; and judging actions 53; and modernity 162; and need 445, 446; non-coercive/rational 159; perfectionist 351; philosophical foundation for government 27; and postmodernism 667–77; and power 674; and public morality/Sunni 209–10; as public space 162–3; of recognition 490, 512; rejection of modern 160–2; and religion/Sunni vision 209–10; return of/Continental philosophy 162–5; virtue 53, 365, 366; women in 156
Politics as a Vocation 354
politics of difference 236, 512
Politics of Friendship 176
Politics of Recognition, The 490
politikê (art of politics) 15, 16, 17
politikos (statesman) 15, 18
Polsby, Nelson 586–7
polyamory 223
polyethnic rights 507
poor, and Adam Smith 100
pope(s): claims to earthly jurisdiction/dominion 40; power of 29, 43; secular incursions by 42 *see also* names of individual popes
Popper, Karl 9–10
popular will 186, 610
population health, sources of 734–6
population size, and food production 292
positive duties: and remedial responsibilities 523; and severe poverty 517–20, 524
positive freedom 152, 562, 563, 565–6; theory of 152
positive law 382
positive obligations 349
positivism, and idealism 155
Posner, Richard 610
possession, and property ownership 220
possibility theorums 689
post-anarchism 226
post-ecologism 299–300
post-left anarchism 226
post-material values, and environmentalism 291
Postmodern Condition, The 668, 669, 675
postmodern discourse theory 706
postmodernism 176, 667–77
poststructural anarchism 226
Poulantzas, Nicos 268, 269, 270, 273
poverty: capabilities approach 520; global 453, 521; severe and international distributive justice 517–24; and survival of the fittest 149
power 585–94; advanced capitalist societies 267; coercive political 597; and deliberation 615, 616; executive/legislative/civil society 121; imbalance of West/Islam 211; legitimizing colonial 201; papal 29, 43; political and contractualism 318; and politics 674; state/and liberty 149; state/perfectionism 351
power, a 619
power-allocating institutions 592
power-knowledge 273
powers: as rights 377; second order 620
practical rationality, and human rights 500–1
practical reasoning 357, 360
practical will 116–19
practice, and pragmatism 656
pragmatic discourses 711
pragmatism 656–66
pre-emptive thesis 597
preference aggregation, and social welfare function 679
preference formation, adaptive 581–2
preferences: autonomous 582; divergent 758; first/second order 581; and needs/intergenerational justice 471–3
preference utilitarianism 695
priests, function of 42
Prime Mover 27
primogeniture convention 702
Primoratz, Igor 779

Prince, The 48, 49, 50, 51, 52, 53, 54, 55, 56
principalities 56
principle of exclusion, and priciple of restraint 806
Principle of Individuality and Value, The 153
principle of justice, and principle of legitimacy 320
principle of legitimacy 317–18, 320
principle of liberty, Mill 124–5, 129–30, 132, 133
Principle of Precedence 449–50
principle of restraint 804
principle of subsidiarity 382, 384
Principle of Utility 130
prioritarianism 333, 450, 559; triage 475, 478
priorities problem, healthcare 739
Prison Notebooks 267
private acquisition, and "enough and as good" 76
private and public: and Confucianism 185, 186; marriage, sex, and family 744–6
private life, and public authority 745
private ownership, and human flourishing 370
private property: Aquinas 31; Locke 75–8
private right, Kant 110
private, three senses of 744–5
procedural independence, nonautonomous preference formation 581
proceduralist/historical conception, autonomy 581–2
proceduralist social epistemology, democracy 603–4
procedural justice, and healthcare 740
procedure of construction 403
production: justice in 756–7; relations 137, 138
production possibility frontier (PPF) 471, 472, 474
productive activity, types of 755
productivity, Marx 136–7
profit, and anarchism 220
progress, and conservatism 249
Prolegomena to Ethics 150
Proper Sphere of Government, The 148
property: Locke 72–3; ownership 77, 220; private and anarchism 220
property rights: Adam Smith 100–1; and freedom 569; initial 75–6; Kant 110; and libertarianism 413–14, 415, 424; and natural law 381; natural rights to 380; Nozick 430; private/Locke 75–8
Prophet's *sunna* 207
proportionality, and *ius in bello* 543
proprietarianism, parental 751
prostitution, and paternalism 747
Protagoras, The 4
Proudhon, Pierre-Joseph 217, 220, 226
provision, PMP (minimum necessary level of provision) 474
Psichari, Ernest 170
psychoanalysis, Sinclair 156
psychological conflict, political solution/Plato 6
psychological egoism 311, 313
psychological support argument 761
psychology: Hobbes 62; of toleration 633–4
public coercive laws 112
public culture, and nationalism 488
public good, and individual interest 154
public health 402, 734–5
publicity, and public sphere 163–4
publicity criterion, capability approach 460
Public Justification Principle (PJP) 803, 805, 806
public life, women in 156
publicness, and truth 161
public opinion: and the public sphere 163; and social order 220
public order, Medieval thinkers 45
public/private sphere, and liberalism 233
public reason 745; liberalism 803–4
public right, Kant 110–11
public role, of combatants 783
public sector, and natural law 382
public space, politics as 162–3
public sphere: and democracy 614; and democratic politics 166; political construction of 514; and postmodernism 677; space for politics 163–5
Pufendorf, Samuel von 104
punishers, altruistic/rule-following 652–3
punishment 765–74; criminal/contractarian approach 313–14; death/for thought-crime 12; and fairness 315; harsh/overcoming complacency 54
pure negative freedom 562–3, 567–9, 571
purusharthas 193–4
Puydt, Paul Emile de 221

Qin, state of 189
quality adjusted life year (QALY) 739
quantum physics, and paralogy 669–70
Qudaragesimo Anno 32
quietism, Daoists 187
Quong, Jonathan 618–27
Qur'an, the 204, 205, 212
Qutb, Sayyid 213–14

race, and inequalities in healthcare (US) 735
Rachels, James 436, 437, 439
Racial Contract, The 280
racial privilege, and deliberative democracy 615
Radhakrishnan, S. 200
radical democracy 614
radical democratic deliberativism, and political activism 614
radical environmental thought 295–6
al-Raghib al-Isfahani 211
Raghuramaraju, A. 192–202
Ramayana 195–6
Rand, Ayn 224
ranking, and difference principle 409
rape, and consent 747
rational agency, and psychological egoism 311; theory of 310–11
rational agents 118
Rational Choice Contractarianism 309–10
rational choice theory 691–703
rational consensus, discourse theory 714–15

rational discourse 713
rationalist conception, autonomy 575–7
rationalist democratic theory 714–16
rationalist discourse theory 706
rationality: individual 312, 313; and moral status 556
rationalization, modern 165, 166
rationally required end interpretation, Hobbes laws of nature 65
rational system of liberty 101
rationing, healthcare 739, 740
Rawls, John: and boundaries 527–8; contractualism/political liberalism 317–27; and desert 433; difference principle 401–10, 438–9, 458, 693, 695; and distribution 691, 692; equality of opportunity 729, 737; and the family 752; and human rights 502; intergenerational justice 468; and justice 337–8; and liberalism 236, 239, 240, 241; and meaningful work 758; on Mill 127–31; and nationalism 486; and normative Contractarianism 309; pluralism 362; political legitimacy 597, 598; and property rights 429; Public Justification Principle (PJP) 803; public reason 745; and religion 800–1; self-imposed obligations 315; social primary goods 458; theory of justice 90, 274–5, 459–60, 462, 527; and toleration 636
Raz, Joseph 358–9, 512, 581, 596–7, 598, 601–2
Reading Capital 267
real democracy 165
realism: normative/descriptive 538; and war 538–9, 784
realistic pessimism, and conservatism 250
realistic stability, and justice 324
realistic utopia 213, 240, 326
reality, absence of/postmodernism 670, 671
Real Self condition 575
reason: and human nature/Aristotle 19; modern 165; and passion 716; public/non-public 745–6; public use of 159
reasonable array of services, healthcare 741
reasonable disagreement 346, 360, 362
reasonable pluralism 362
reasoning: from agreement 321–5; civil collective 616; and freedom 566; practical 357, 360; substantive 323
Rebel, The 175
reciprocity: and difference principle 410; direct/and evolution of cooperation 652
reciprocity theorum 64, 66
recognitional accounts, autonomy 583
recognition, dialectic of 117
Reconstruction in Philosophy 663
rectification, of land unjustly taken 430 *see also* compensation
recycling, and consumerism 673
redistribution: and taxation 430; of wealth 155
Re-enchanting Humanity 296
refinement, benefits of 97, 101
"Refinement in the Arts, Of" 97, 98
Reflections on Violence 169
reflective equilibrium 542, 623
reflective self-management 576
reflexivity thesis 576
Reformation, the 801
reformist-nationalist response, Islamic modernity 211–12
refugees, and border crossings 534
regime of contract/status 149
regulation: moralistic 795; paternalistic 788, 792, 793, 798; of people by laws/institutions 189–90
Rehg, William 706–16
Reidy, David A. 494–504
Reiff, Mark R. 765–74
Reign of Terror 177
relevance filters, luck egalitarianism 393, 397
religion: and anarchism 226; and conservatism 252; Hobbes 67, 69; and multiculturalism 510; Plato 12; in politics 802; and politics/Sunni vision 209–10; preferential treatment of 808; in public life 800–10; as threat 801
Religion of Humanity 130
religious accommodation 807–9
religious citizens, and public justification 805
religious conceptions, value pluralism 358
religious conflict, and de-politicization of religion 81
religious freedom, and Second Vatican Council 33
religious fundamentalism, and republican thought/France 177
religious liberty 124
religious marketplace 81
religious persecution 801
religious reason 804
religious sectarianism, Islam 206
religious significance, of land 532
religious toleration 80–1, 632, 637
remedial responsibilities, severe poverty 522–3, 524
Renan, Ernest 170
ren, as social glue 183
Renault, Alain 176
rent, and anarchism 220
repeated games, theory of 702
replicator dynamic 651
repression: injustice of 365–6; religious 801
repressive rearing, and autonomy 581
republic: as ideal government/Kant 111; as ideal society/Buddha 199
republican freedom 90, 563, 569–71
republicanism 254–63; civic 613; classical 48; and liberalism 256, 258; martial/Machiavelli 56–7; and multiculturalism 177; revival of/France 177
republican liberty, and deliberative democracy 613
republican theory of justice 263
republics 56–7
Republic, The 4, 5–10, 11, 12, 15, 20, 21, 210, 279, 365, 750
Rerum Novarum 33
research, performative requirements of 669
resistance, and postmodernism 674–5, 677
Resistance, French 175

Resnik, M. 695
resource allocation, axiomatic models of 688–9
resource egalitarianism 695
resources, non-depletionary use 468–70
respect: and care 796, 798; and ethics 795; Kant 109; and markets 797–8; opacity respect 557; for others rights 425; for persons/human good 346–7; recognition/appraisal 809
responsibilities, and needs 448–50
responsibility: and capability theory of justice 464; and choice 397; and luck 391–3; personal and paternalism 796
responsibility-sensitive theories 412, 413
Responsibility to Protect (R2P) doctrine 498
responsible agents, and desert 442
ressentiment 288
restraint, and integrity 806
restricted neutrality principle (RNP) 345–6
retributive punishment 767, 768, 770–1, 773
revivalism, and Islamic modernity 212–14
revolt of individuals, and revolution 175
revolution: to ethics/France 177; Kant 111; and terror 175, 177
revolutionary socialism 147–8
revolutionary syndicalist movement, and revolt of individual 175
rewards/punishments, and regulation of people 189–90
Richter, M. 152
Ricœur, Paul 177
right: concept of 114–15; as moral power 378
right intention, and just war 541
right libertarianism 414, 418
Rightly-Guided Caliphs 205
right of conscience, and human dignity 809
right of nature, Hobbes 63
right(s) 618–27; of association/multi-layered society 380; civil/political 496; and consequentialism 337–8; to cross borders 534; deontic logic of 495; and dignity 558; to do a wrong 414; economic/cutural/social 496; essential unity of 496; of exit 424, 426, 427, 509; first/second order 78, 619; foundational 414–15, 418; framework for 618–20; of freedom/Kant 109; and free way of living 431; and global justice 453; group-differentiated 506–7, 511; to health/healthcare 741–2; human/moral 605; human rights 495–8; and intergenerational distributive justice 467–8; and justice 115; left libertarianism 412, 413–17; libertarian framework of 423–5; liberty 496; moral 414, 624; natural 378; to natural resources 412; non-acquired/Locke 72; non-discretionary 378; to nonintervention 503; over children 751; of ownership 424; polyethnic 507; positive/negative 453, 496; powers as 377; to privacy 744–5; of private judgment 79; private/Kant 110; private property/Locke 75–8; private/public/international/Kant 110–12; property 429, 430; to religious freedom 801; respect for others rights 425; of self-ownership/Locke 72; of self-proprietorship 75–6; and taxation 430; territorial 532–3; theory of 151–2; and toleration 635–6; to unowned natural resources 414–15; as valid moral claims 499; violation of/Locke 79; welfare 496; and welfarist social choice 687; to work/living wage 155
Rights of Man 219
Rights of Man and the Natural Law, The 33
Rights of War and Peace, The 93
rights-protection industry, state monopoly of 224
rights talk 500
rights theory, and natural law 375–84
right to a living wage 155
right to work 155
Rig Veda 192
Riker, William 610
Ripstein, Arthur 597
risks, unconsented-to/health 736
Ritchie, D.G. 147, 150, 646–7
road-pricing 294
Robespierre, Maximilien 177, 246
Robeyns, Ingrid 456–65
Rock Edicts 199–200
Rocker, Rudolf 224
Rodin, David 540
Roemer, John 757
Rolland, Romain 170, 171
Roman Catholicism, and Thomism 33
Romulus 54
Rorty, Richard 657, 664
Rosanvallon, Pierre 177
Rose, David 114–23
Ross, W.D. 785
Rousseau, Jean-Jacques 82–90; and anarchism 218–19; and democracy 602, 611; and nationalism 489; and political authority 599, 601; and political autonomy 573; role of women 281, 283
Rta 192–3
rule consequentialism 334–5
rule-egoist variant, desire based derivations 65
rule-following punishers 652–3
rule of law: international/human rights 503; and libertarian theories 414
rule(s): and justice 526; and law 676; true kings/tyrants 44
rule utilitarianism, and justice/rights 337
ruling, as privilege 22
ruling class, China 183
Russell, Daniel C. 364–73
Russia: Bolshevik Revolution 170, 171; and Marxist/anarchist ideas 223 *see also* Soviet Union
Russian Revolution, anarchist militias 224
R. v. Brown 747

Saari, Donald 681
Said, Edward 201
Salafi movement 213
Salle, David 667
Samaritan account, political legitimacy 601
same-sex marriage 749

Sample, Ruth 282
Samyukta Nikhaya 198
Sanders, Lynn 615, 616
sangha 198
sanyasi 195
Saptanga theory of the state 193, 195–7
Sartre, Jean-Paul 173–4, 176
sastras 193
Sastry, Shama 201
satisfaction, and autonomy 579, 580
Satz, D. 725
saving, intergenerational 468, 469
Savonarola, Girolamo 47
Say, Jean-Baptiste 219
Scanlon, T.M. 317, 319–20, 325–6
Scheffler, Samuel 398, 433, 779–80
Schleicher, August 643
Schmidtz, David 433–42
Schmitt, Carl 161, 356
Schnapper, Dominique 178
schooling, and education 721–2
school of British idealism 150
Schumpeter, Joseph 610
Schwartz, Adina 761
science, and God 656
science of politics *see* political science
scientific discourses 711
scientific inquiry, and democracy 662
scientific method, of inquiry 657
scientific thought, and philosophy 661
screening, and difference principle 408
Searle, John 707
Sea Shepherd Conservation Society 297
secession, and libertarianism 416
second-order powers 620
second-order volitions 578, 579
second-person reasons 372
second principle of justice 401–2
Second Sex, The 173
Second Treatise of Civil Government 32
security: competitive market in 225; state monopoly of 222, 224
Security Council 497–8, 537–8
seduction, Baudrillard 675–6
self: modern 116; and other/capitalism 140
self-consciousness, and moral status 555
self-defense 539–40; strict 540
self-delusion, French political thought 171
self-determination: for cultural groups 490, 492; and paternalism 789
self-governance 577
self-government, national minorities 507–8
self-help, and international relations 538
selfhood 85; and deep ecology 295–6
self-interest 99
Selfish Gene, The 653
self-knowledge, and paternalism 791
self/other regarding acts, and liberalism 233
self-ownership: as foundational moral right 414; left libertarianism 412, 415; right of/Locke 72–3
self-ownership thesis 757
self-preservation: agricultural mode of 76; as basic purpose 41; and law of nature 65, 74, 75; natural right of 75
self-protection, and liberty 124
self-realization: and knowledge 669; through work 142
self-reflection, and autonomy 577
self-reliance, and needs satisfaction 449
self-rule, by public opinion 163
self-satisfaction 151
self-sufficiency: Aristotle 20; as citizen-virtue 365
self-understanding, and freedom 121–2
Sen, Amartya: aggregating capabilities 462; and Arrow's Theorum 681; capabilities 591, 738; capabilities approach 517, 519–20; and capability account/justice 461; and distributive justice 695, 700; and justice 464; and primary goods 458–9; Sen's Theorum 687; social welfare function 683; and utilitarianism 685
Sensen, Oliver 103–12
Sen's Theorum 687
sensual enjoyment, supervision of 194
sentiment, and moral agency 287–8
Seventh Epistle 3
sex 746–7; marriage, sex, and family 744–6
sexual and racial contract 279–80
sexual bias, in philosophy 283
Sexual Contract, The 279
sexual difference: revisiting 285–6; and women's rights 278
sexual freedom, American anarchists 222
sexual morality 31, 746–7
sexual relations, Aquinas 31
Shar'i/non-Shar'i duties 208
Shaw, G.B. 148
Shi'ism 206–7
ship of state parable 9
Shue, Henry 453, 496
side-constraints, and libertarian rights 424
Sidgwick, Henry 147, 311, 532, 543
Sierra Club 297
Sigmund, Paul 25–35
silent majority, and seduction 676
Simmons, John 599, 600
Simon, Julian 293, 298
sincerity claims, validity claims 710
Sinclair, Mary 156
Singer, Peter 517, 518–19
sin(s), obligatory annual confession 30
Sittlichkeit 119–22
al-Siyasa al-shar'iyya 209
Skeptical Environmentalist, The 298
skepticism: and conservatism 246–7; environmental 298–300; and human rights 498–500; and toleration 632
Sketch of a Theory of the Emotions, A 173
Skinner, Quentin 255, 258–9
slavery: Aristotle 15–16, 22, 23; Hegel 115; and human rights 500; Locke 72; Plato's just city 7; and property rights 424; and socialism 149

Slote, Michael 365–7
Smart, J.J.C. 335
Smith, Adam 97–101, 219
Smith, John Maynard 649–50
sociability, and liberalism 234
social anarchism 219, 223–4, 226
Social and International Ideals 153
Social Catholicism 32
Social Choice and Individual Values 679
social choice function, strategy-proofness 686
social choice, intergenerational 471–3
social choice model, and distribution 474
social choice theory 679–89
social complexity 167
social condition, reconciliation to 326
social conflict, and law/obligation 103–4
Social Contract 219, 573
social contract: agreeing 426; and cultural evolution 703; and decision theory 691; fraudulent/Rousseau 86; and game theory 696–703; Glaucon 5; impossibility of 95; and individual rational choice 691–6; and political authority 601; theory(ies) 59, 82, 309; tradition 279–80
Social Contract, On the 82, 83, 84, 88, 89, 90
social Darwinism, Spencer 148
Social Democratic Federation 155
social development, and justice 96
social ecologism 296
social egalitarianism, and luck egalitarianism 393–6
social epidemiology 735
social evolution 149, 643–54
social form, of the state 271
social functioning, and basic needs 447
social glue, of society/China 183, 184
social inequality 155, 240
social institutions, as ethical ideas 153
socialism: British/Fabianism 155; and conservatism 252; and evolution 646–7; and liberalism 155; Marx's vision of 142–4; moral/economic 153; and natural law 383; and Plato 8; revolutionary/reformist 147–8; and slavery 149; Tucker 222
socialist ideas, and state intervention 147
social justice: capability approach 456–65; and environmental activism 298; and health 732; as intellectual error 427; and needs, desert and equality 450; and opportunity 737; and population health 735; Rawls 429; and the state 431; and sufficiency 725; and taxation 429; theory of 461
social life, organic view of 155
social minimum baseline 409, 410
social order, and public opinion 220
social ordering 680, 681
social order of actions 647–8
Social Organism, The 644
social power 589
social practice, realistic 326–7
social practices, and political philosophy 325–7
social primary goods 458–60, 462
social problems, Islamic solutions to 213
social progress: theory of 221; and utilitarian public morality 336
social provision, France 178
social rationalization, and liberalism 166
social reform, and liberalism 234
social relations, and exchange of commodities/capitalism 271
social rules: and refined condition 98; and social utility 313
social solidarity, and nationalism 488–9
Social Statistics 148
social structure, and plurism 236
social struggles, and republicanism 262
social thought, Neo-Thomist 33
social union 95
social unity 326
social utility: and individual welfare 312; and social rules 313
social wefare orderings, variable-population 685
social welfare: and individual rationality 313; and personal flourishing 154
social welfare function 679, 680, 694
social welfare functionals 683–5
societal cultures, and nationalism 490–2
society(ies): and luxury 101; modern as a totality 166; multi-layered 379; order of/safety of people 197; as perfectly functioning organism 148; prosperous 97; status hierarchies 490; of strangers/social glue for 184; well-ordered 326
socioeconomic class, and education 723
socio-economic position, and life prospects 752
socio-economic status (SES) 735
sociology of knowledge 283
sociorelational conception, autonomy 582–3
Socrates 3–5, 307–8
Soderini, Piero 56
soldiers, permission to kill 543–4 *see also* combatant/non–combatants
solidarism, and liberalism 232
solidarité 169
Solzhenitsyn, Alexander 176, 177
Sorel, Georges 169–70, 170
soul: and Medieval thinkers 39, 42; needs of 172–3; and Plato's city 6
South African government, state terrorism 780
sovereign authority 64
sovereign(s): and Hobbes laws of nature 66; role of/civil society 121
sovereignty: of individual 220; by institution 59; Rousseau 89; Saptanga theory of the state 197; transnational 274
Soviet Union: and French political thought 176; and Marx 136; and oppression 172; Stalin's 172 *see also* Russia
Spanish Civil War, anarchist militias 224
Spanish conquest, justification of 554
Sparta, constitution 21–2
Spector, Horacio 573–83
speech acts/theory 707
Spelman, Elizabeth 279

Spencer, Herbert 146, 147, 148–50, 221, 643–5
Sperber, Dan 653
Spinoza, Baruch (Benedict De), feminist interpretations 287
spirit, and Plato 6
Spirit of the Laws, The 98
spoilage, Locke 76
Spooner, Lysander 221–2, 225
Spring and Autumn period 181, 182–3
standards, external/non-accidental 660, 661
State and Revolution, The 267
state censorship, and public sphere 164
state coercion 364, 366; and deliberation 614; justification of 803, 804; and moral claims of citizens 371–2; and paternalism 790–1; and perfectionism 349
state-derivation debate 270
State in Capitalist Society, The 267
state intervention: liberalism with 154; and liberty 154; limits of/Bosanquet 153; and new liberals 147
state involvement, economic/social infrastructures 150
state legitimacy, and satisfaction of people's interests/Confucians 186
statelessness, and anarchism 218, 220
stateless people, and failure of politics 164
state neutrality 342, 344–5, 346, 347
state of families 750
state of nations, Kant 112
state of nature: Hobbes 62–5, 218; Kant 104; Rousseau 219; as state of war 106
State, Power, Socialism 269, 273
state(s): affluent/poor and poverty 521; authority of/consent 308; autonomy 269; and coercion *see* state coercion; and contractarianism 306; and cultural identity/membership 512–13; domination 613; and family life 751; and free markets 417; functions of 272; and good life 371; and human rights 502; justification of the/Hegel 122–3; large/ social glue for 189; and marriage 748–9; Marxist theory of 267–74, 275; the modern/Hegel 119; monopoly of force 537; and national self-defense 539; and nations 484, 486; and natural law/rights theory 383–4; particularization of the 271–2; paternalism 796; political authority of 604; and political liberty 254; power/and liberty 149; role/eudaimonia 368–9; role of/economic conditions 155; role of/freedom 159; role of/Spencer 148; Saptanga theory of the state 195–7; size of and political structure 183; social form of 271; and social justice 431
Statesman, The 10, 15
state terrorism 778, 780
statist view, just borders 531–2
status hierarchies, societies 490
status rights, Kant 110
Steiner, Hillel 412–19, 562–3, 564, 568
Stemplowska, Zofia 389–99
Stephen, James Fitzjames 339
Stirner, Max 220–1
Stoics, and anarchism 217
Stoic thought, and freedom 566
Strange Defeat 173
strategic action 707
Streeten, Paul 454
structuralism and after, French political thought 175–7
structuralist/hierarchical conception, of autonomy 577–81
structural relations, and domination 257–8
Structures of Agency 576
student demonstrations, May 1968 176
Suarez, Francisco 32
Subjection of Women 284
subjective act consequentialism 334
subjective consequentialism 330
subjective probability distribution 693
subordination: and law of nature 74; of women 281
subsidiarity, principle of 32
substantial identity 120
substantive reasoning 322, 323
substantive responsibility 392
sub-state terrorism 786
Sudrahas 198
suffering: compensable/incompensible 774; universality of individual 198
sufficiency, and social justice 725
sufficientarianism: and benefits 450; and equality 559
Suffragettes movement 156
Sugden, Robert 698, 794
Sullivan, Vickie B. 47–57
Summa contra Gentiles 26
Summa Theologiae 26, 28, 30, 32, 376
Sumner, William G. 646
Sunni consensus 207
Sunni constitutionalism 207–10
Sunstein, Cass 613
supererogatory acts 634
supreme emergency, as justification for terrorism 785, 786
survivalists 293
survival of the fittest, moral implications 149
sustainability, material 468–70
sustainability requirement, intergenerational justice 470
sustainable development 294
symbolic/religious value, of land 532–3
sympathy, and moral judgments 97
syndicalist movement 169, 170
Syracuse, and Plato 10
System of Economic Contradictions 220
System of Synthetic Philosophy 148

Taine, Hippolyte 170
Taittiriya Aranyaka 192
takfir 208
Talisse, Robert B. 608–16
Taoists, and anarchism 217
Tarde, Alfred de 170
Tasioulas, John 605

taxation: decedents' estates/natural resource values 416, 419; and environmental issues 294, 299; and government 79; and natural law 381–2; as rights violation 430, 431; as slavery 424; and social justice 429
Taylor, Charles 358, 487, 489, 490, 492, 512
Taylor, Harriet 284
Taylor, M. 698
technoanarchists 226
technology, and anarchism 226
teleological approach, Aquinas 31
teleological humanism 296
Ten, C.L. 124–33
territorial rights, peoples as holders of 532–3
territory: annexed 530, 532; collective rights to 531–2 *see also* land
terror, and revolution 175, 177
terrorism 224, 777–86; tactical definitions 779, 780–1
Tesón, Fernando R. 537–47
Tessman, Lisa 366
texts: Chinese classical 182; non-Christian/Medieval study of 36
Thalberg, I. 577
Thaler, Richard & Sunstein, Cass 794
theft, natural law against 380
theists 804–5
theology, and political philosophy 25
theories of the good, and consequentialism 331–2
theorizing, ideal/non-ideal 722
Theory and Practice of Autonomy, The 577
Theory of Communicative Action 706
theory of evolution 147, 152
Theory of Games as a Tool for the Moral Philosopher 696
theory of individual rationality, and theory of value 312
theory of justice 90, 459–60, 527
Theory of Justice, A 128, 130, 274–5, 324, 337–8, 401, 403, 409, 410, 486
Theory of Knowledge, The 154
theory of liberty, Mill 147
Theory of Moral Sentiments, The 97
theory of political justice, and social primary goods 459
Theory of Property 220
thesis of public unavailability 318–19, 321
Thierry, Augustin 219
Third Republic 169, 170
Thirty Years' War 93, 801
Thomistic natural law 28
Thomists 246
Thomson, William 688
Thoreau, Henry David 292
thought-crime, death for 12
Three Fundamental Theorums in Aristotle's *Politics* 18
Tilak, B.G 201
tit-for-tat 652
tolerance, and traditionalist conservatives 245
toleration 629–38
toleration: Mill 124, 127; and plurism 235; political 635; religious 80–1, 632, 637
toleration principle, liberalism 236
Tolstoj, Lëv 223–4
Tooley, James 725
totalitarianism, and positive freedom 563, 565–6
"Toward Perpetual Peace" 106
trade, and poor states 521
trade unionism, militant 172
traditionalist conservatives 244–6
traditional polis 3
tradition, and incommensurable values 358
traditions, and well-being 244, 245, 252
Trajan, Emperor 38
tranquility, and the political organism 41
transcendentalism 292
transnational capitalism 667–8, 672–4
transnational corporations, and power of nation-states 672
transnational sovereignty 274
treasury, Saptanga theory of the state 196
treaties, human rights 497, 502
Treatise of Human Nature, A 92–3
Treatise on Law 28, 376
Treatise on Law of the Summa 29
Treta Yuga 192
triage, distribution according to 451–2
Trout, J.D. 794
true freedom 152
True Levellers 218
trust, and nationalism 489
truth: and freedom 566; and inquiry 658; and publicness 161; seeking and holding/Mill 125; theory of 656–7; understanding 661
Tuana, Nancy 278–9
Tucker, Benjamin 222
Two Concepts of Liberty 562, 563, 564, 571
Two Treatises of Government 71, 72, 74, 75, 78, 218
tyranny: Locke 72; of pain and suffering 198; and power 585

'Ulama 208
Ultimate Resource, The 293
Umayyad dynasty 206
unanimity requirement 407
Unconventional History of Western Philosophy 282–4, 288
under-determination by value judgements 576
unfreedom, and passion 716
Union Carbide disaster 780
union sacrée 170
United Kingdom *see* Britain
United Nations: Convention (1051) 534; and natural law 384; Security Council 497–8, 537–8; Universal Declaration of Human Rights (UDHR) 33, 494, 497, 502, 521, 809
United Nations Charter 537
United States: anarchism 221–2; conscientious objectors 808; healthcare 734–5; religious freedom/repression 801, 802, 807; repression of anarchism 224 *see also* America

United States Constitution: and deliberative democracy 613; and liberalism 232; and religious freedom 802, 807; and slavery 222
Universal Declaration of Human Rights (UDHR) 33, 494, 497, 502, 521, 809
universalism, liberal 233, 237
universality, and justice 104
Universal Principle of Right 104, 105–6, 109
universal rights, and agency 64
unowned natural resources: and compensation 415; rights to 414–15
unsocial sociability, of mankind 159
untouchability 198
upward mobility, Confucians 187
USS Cole 780
usury, Aquinas 31
'Uthman 206
Utilitarianism 132
utilitarianism: and act consequentialism 333, 338; act utilitarianism 337; axiomatized 684; and consequentialism 329–39; Harsanyi 685–6; and health 736–7; Mill 147, 234; normative theory of 312; preference 695; and psychological egoism 311; and public morality 336; and radical poverty 339; and Rational Choice Contractarianism 310–13; of rights 338; and terrorism 784; total 685; and welfare 332–3
utilitarian moral theory 312
utilitarian normative theory, and psychological egoism 313
utilitarian theories 307; and liberalism 233
utilities, distribution problems with 694
utility: and desert 439; in economics 333; interpersonal utility comparisons 694; and liberty 127–31, 131–2; Mill 131; and political authority 601; principle of 130; social/individual 311
utopia, imagined 425–6

Vaggians 199
Vaishya 198
validity claims, discourse theory 707, 708, 710, 714
Vallentyne, Peter 460, 465
Vallier, Kevin 806
Vallier, Kevin & Eberle, Christopher 800–10
valuational system, autonomous agency 576
Value and Destiny of the Individual, The 153
value, choice criterion of 128
value pluralism 353–6, 358, 360, 361–2
values: and conservatism 247–8, 250; and cultural difference/toleration 637; and democratic theory 739–40; human 354; incommensurable 353, 355–6, 357, 358; labor theory of 271; Plato 4; plural 354, 355; plurality of 359; private and health care 798; and well-being 248
value substitution, and paternalism 788, 791–4, 795–6
value theory 272, 275, 312
vanaprastha 195
Vanderschraaf, Peter 691–703
Vatican Council, Second 25, 33
Vatsyayana 194
Vedas 198
Vedic traditions, and Buddha 200
veil of ignorance 407, 408, 685, 692, 693, 695
veto, popular and democracy 610
Vichy regime 173
Vindication of Natural Society 219
Vindication of the Rights of Women, A 281
violence: and group identity 510; renouncing of/ Emperor Ashoka 199–200
Virtue and Politics 372
virtue ethics 364–73
virtue politics 365, 366
virtue(s): Aristotle 17, 20; artificial 99; benevolent 99; citizen 21, 365; civic and republicanism 259; economic 99; and eudaimonia 368; faith/hope/charity 28; institutional 365; and justice 93, 94, 99, 365; lack of citizenly 88; pity as 184; Plato 4, 7, 12; and politics/Machiavelli 53; and self-sufficiency 365; and state coercion 372
virtuous character 372
Virtuous City, Farabi 210–11
Vitoria, Francisco de 505, 558
Vivekananda, Swami 201, 202
Vlastos, Gregory 279
Von Neumann, J. & Morgenstern, O. 696
votes, allocation of 594
voting power, and democratic theory 592–4
voting procedures 679

wage-labor relationship 755, 760–1
wage-labor system, and work content 757
Wahhabi movement 213
Waldron, Jeremy 779
Wall, Steven 342–51
Walzer, Michael 336, 439, 527, 528, 759, 785, 786
wants, and needs 445
war crimes 498
Warren, Josiah 220
Warren, Karen 288
Warren, Mary 282
Warring States period 182–3, 189
war(s) 537–47; American Civil War 222; Aquinas 30–1; civil 63; cold war era 154, 174–5; and defense of states/persons 541; defensive/ offensive 541; English Civil War 59; First Civil War, Muslim 206; Great War (WW I) 170–1; just 30, 384, 537, 539–42, 545, 781–2; Kalinga war 199; killing in 543–4; and morality 537, 538–9; and rights of individuals 111; in self-defense 539–40, 544; Spanish Civil War 224; state of 63; Thirty Years' War 93, 801; Wars of Apostasy 206; World War II 173–4, 496, 502, 778, 783, 785
Watchdogs, The 171
Watson, Gary 576, 578–9
Watson, Paul 297

wealth: distribution of/libertarianism 427–8; distribution of/Rawls 402–3; and the Qur'an 205; redistribution of 155; social/individual factors 155
Wealth of Nations, The 93, 97, 99, 100
Webb, Beatrice 148, 155
Webb, Sidney 148, 155
Weber, Max 162, 165, 175, 354
Weil, Simone 172–3
Weithman, Paul 259, 598
Welchman, Jennifer 662
welfare: conceptions of 332; conditions of/Buddha 199; and liberalism 234; personal/contractarianism 306–7; Rational Choice Contractarianism 310; and utilitarianism 332–3, 736–7
welfare of society 153
welfare rights 496
welfare state, and liberalism 234
welfarism 331, 683–4, 685
well-being: capability approach 456; and conservatism 246; cultural aspects 358; and difference principle 410; and eudaimonia 367; and health 732–3, 736; human 343; and identity 490; and liberalism 240; negative changes in/punishment 767, 768; and paternalism 789, 795; pure conceptions of 331; and state coercion 371; and traditions 244, 245; and values 248; and virtue ethicists 367
well lived human life 343, 350
Wellman, C. 601
well-ordered society 326
well-settled beliefs 660–1
Wells, H.G. 148
Western colonialism/imperialism, Islamic answers to 211–12
Western Zhou dynasty 182
Weymark, John A. 679–89, 685
What is Property? 220
Whitehall studies 735
White, Mark D. 788–98
White, Morton 664
Wiggins, David 447
wilderness, and American environmentalism 292, 298
Wildlands Project 298
will: people's/and what is right 660; philosophy of/Hegel 114–16; practical/Hegel 116–19
William of Ockham 375
will theory of rights 413, 416, 620–1, 623, 626
Winstanley, Gerrard 218
Wolf, Clark 467–78
Wolff, Robert Paul 227, 598–9
Wollstonecraft, Mary 281, 283, 284
Wolterstorff, Nicholas 804–5
women: and Buddhism 198–9; and employment 284; equality of/Plato 8, 12; inferiority of 279, 284; mental incapacity of/Aristotle 22; and patriarchy 296–7; subjugation of 280; subordination of 281
Word, The 222
work 142, 755–62
worker cooperatives, and market capitalism 758
worker participation, managment/governance of economic enterprises 759–61
Workers' Socialist Federation 156
workmanship 755
workplace, democracy 761
World Health Organization (WHO) 732–3, 735
World Trade Organization 521, 604
World War II 173–4; allied bombings as terrorism 778, 783, 785; and human rights 496, 502
Wretched of the Earth, The 176
Wright, Sewall 647
wrongs, civil/criminal 774

Xenophon 20
Xun Zi 183, 186, 189

Young, Iris 490, 512, 615, 616
Young, R. 578
yugas 192

Zermelo, E. 696
Zhuangzi 189
Žižek, Slavoj 266
Zola, Emile 169

Taylor & Francis

eBooks

FOR LIBRARIES

ORDER YOUR **FREE 30 DAY** INSTITUTIONAL TRIAL TODAY!

Over 23,000 eBook titles in the Humanities, Social Sciences, STM and Law from some of the world's leading imprints.

Choose from a range of subject packages or create your own!

Benefits for you

- Free MARC records
- COUNTER-compliant usage statistics
- Flexible purchase and pricing options

Benefits for your user

- Off-site, anytime access via Athens or referring URL
- Print or copy pages or chapters
- Full content search
- Bookmark, highlight and annotate text
- Access to thousands of pages of quality research at the click of a button

For more information, pricing enquiries or to order a free trial, contact your local online sales team.

UK and Rest of World: **online.sales@tandf.co.uk**

US, Canada and Latin America: **e-reference@taylorandfrancis.com**

www.ebooksubscriptions.com

ALPSP Award for BEST eBOOK PUBLISHER 2009 Finalist

Taylor & Francis eBooks
Taylor & Francis Group

A flexible and dynamic resource for teaching, learning and research.