Second Edition

Handbook *of* Univariate *and* Multivariate Data Analysis *with* IBM SPSS

Robert Ho

CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

Second Edition

Handbook *of* Univariate *and* Multivariate Data Analysis *with* IBM SPSS

Second Edition

Handbook *of* Univariate *and* Multivariate Data Analysis *with* IBM SPSS

Robert Ho

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business
A CHAPMAN & HALL BOOK

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Version Date: 20130830

International Standard Book Number-13: 978-1-4398-9022-6 (eBook - PDF)

Contents

Preface

Most statistics textbooks focus on theoretical discussion and mathematical formulas of the concepts presented. The result of this approach is that students often have little understanding of how to *apply* statistical tests to their experimental findings, or how to *interpret* their findings. The advent of the SPSS statistical package has gone some way towards alleviating the frustration that many social sciences students feel when confronted with the task of analyzing data from research projects. Nevertheless, the problems of test selection, execution of a specific statistical test, and interpretation of results from computer printouts still remain daunting for most students.

In preparing this second edition, the major focus of this book has remained the same, that is to provide clear guidelines to both the execution of specific statistical tests and the interpretation of the findings by (1) explaining clearly the purpose of specific tests and the research designs for which they are relevant, (2) identifying the assumptions underlying these tests and how to test for them, (3) demonstrating clearly the instructions for executing univariate and multivariate statistical tests, and (4) explaining how obtained results should be interpreted.

This second edition has been written for use with version 20 of the SPSS statistical package for Windows. As with the first edition, this second edition focuses on both the Windows method and the syntax method of analysis, which has proven popular with both experienced researchers and students. The popularity of SPSS syntax files rests with the fact that the ability to write and edit command syntax is advantageous, such as (1) when the researcher wishes to repeat an analysis multiple times with minor variations, and (2) manipulate data in a complex design.

In preparing this second edition, the following changes/additions have been made to the book.

- Three new chapters have been added. These chapters deal with the multivariate techniques of:
    1. **Multiple discriminant analysis—**A statistical technique that allows the researcher to determine which continuous variables discriminate between two or more naturally occurring groups.
    2. **Logistic regression—**A statistical technique that allows one to predict a discrete outcome, such as group membership, from a set of variables that may be continuous, discrete, dichotomous, or a mixture of any of these.

3. **Canonical correlation—**A statistical technique that facilitates the study of interrelationships among sets of multiple dependent variables and multiple independent variables.

- **Tests of assumptions** (e.g., linearity, outliers, normality, homogeneity of variance-covariance matrices, multicollinearity) underlying the statistical tests covered in this book have been included.
- The calculation of **Type I error** has been included in Chapter 1.
- A section on **missing data** has been added to Chapter 2. Topics covered include:
  - Patterns of missing data
  - Test for patterns in missing data
  - Dealing with missing data
  - An example of the expectation-maximization method for handling missing data
- The procedure for testing statistical significance between two correlation coefficients obtained from two samples has been included in Chapter 10.
- For the topic of **factor analysis** covered in Chapter 12, two additional methods of factor extraction (i.e., in addition to the use of the **eigenvalue** and the **scree plot**) have been included. These are **Velicer's minimum average partial (MAP)** test and **parallel analysis**.
- For the topic of **path analysis** (measurement variables only) via multiple regression analysis covered in Chapter 14, the test of significance of the **mediation hypothesis** has been included. Developed by Preacher and Hayes (2008), the technique allows the researcher to request **significance tests of total and specific indirect effects via bootstrapping**.
- For the topic of **structural equation modeling** covered in Chapter 18, a section on **second-order confirmatory factor analysis** has been included. Additional information has been provided to aid the interpretation of the **RMSEA** fit index. A section on "problems with estimation" has been included to explain the model estimation problems of **nonpositive definite covariance component matrices** and **Heywood cases**.

I believe that the expanded coverage and presentation format of this second edition, like its predecessor, provide a clear guide to students in the

selection of statistical tests, the execution of a wide range of univariate and multivariate statistical tests via the Windows and syntax methods, and the interpretation of output results.

Robert Ho

NOTE: The data sets and SPSS macros employed in the examples in this book can be accessed and downloaded from the following internet address: **http://www.crcpress.com/product/isbn/9781439890219**.

Author Bio

Robert Ho received his DPhil from Waikato University, New Zealand, in 1978. He is presently Associate Professor in the Graduate School of Psychology at the Assumption University of Thailand. His research interests include quantitative methods, health psychology, and social psychology.

1

Inferential Statistics and Test Selection

1.1 Introduction

This book deals with statistical analysis, which involves both *descriptive statistics* and *inferential statistics.* The major concern of descriptive statistics is to present information in a convenient, usable, and understandable form. For example, once the data have been collected, some of the first things that a researcher would want to do is calculate their *frequency*, *graph* them, calculate the *measures of central tendency* (means, medians, modes), calculate the *dispersion* of the scores (variances, standard deviations), and identify *outliers* in the distribution of the scores. These procedures are called descriptive statistics because they are aimed primarily at describing the data. Inferential statistics, on the other hand, is not concerned with just describing the obtained data. Rather, it addresses the problem of making broader generalizations or inferences from the sample data to the population. This is the more complicated part of statistical analysis, and this chapter will focus on the role that inferential statistics plays in statistical analysis.

1.2 Inferential Statistics

As stated above, descriptive statistics is used to describe a set of data in terms of its frequency of occurrence, its central tendency, and its dispersion. Although the description of data is important and fundamental to any analysis, it is not sufficient to answer many of the most interesting problems researchers encounter. Take an experiment where a researcher is interested in finding whether a particular drug can improve people's memory. The researcher offers the drug to one group but not to a control group, and then compares the means of the two groups on a memory test. Descriptive statistics will not tell the researcher, for example, whether the difference between a sample mean and a hypothetical population mean, or the difference between two obtained sample means, is small enough to be explained

by chance alone or whether it represents a true difference that might be attributable to the effect of the experimental treatment, that is, the drug. To address these issues, the researcher must move beyond descriptive statistics into the realm of inferential statistics, and particularly, on to the statistical procedures that can be employed to arrive at conclusions extending beyond the sample statistics themselves. The basic aim of inferential statistics is to use the sample scores for *hypothesis testing*.

1.2.1 Hypothesis Testing

Fundamental to the strategy of science is the formulation and testing of hypotheses about populations or the effects of experimental conditions on criterion variables. For example, in an experiment designed to investigate gender differences in IQ, the researcher hypothesizes that first grade girls have higher IQ scores than first grade boys. She administers an IQ test to four boys and four girls in a first grade class. The results showed that the mean IQ score for the girls (110) is higher than the mean IQ score for the boys (103). Based on these findings, is the researcher justified in concluding that her hypothesis is supported? The answer is obviously "we don't know." That is, although the results clearly showed that there is a difference between the sample means, with the girls scoring higher on average than the boys, there is the possibility that the observed difference could have been due to the chance variability of intelligence among first graders. In other words, given the variability of intelligence among first graders, and the smallness of the sample size, some difference in the means is inevitable as a result of the selection procedures. In order for the researcher to draw valid conclusions about her hypothesis, she needs to employ inferential statistics to test the reliability of the finding of apparent difference in intelligence among first graders. The critical questions that must be answered by inferential statistics then become "does this difference represent a reliable and meaningful difference, or is it due purely to chance variation and therefore without consistency?" A prime function of inferential statistics is to provide *rigorous* and *logically sound* procedures for answering these questions.

1.2.2 Types of Hypotheses

Before discussing the procedures underlying hypothesis testing, it is necessary to distinguish between the research hypothesis and the null hypothesis.

1.2.2.1 Research Hypothesis

Hypotheses derived from the researcher's theory about some social phenomenon are called research hypotheses. The researcher usually believes that his research hypotheses are true, or that they are accurate statements about the condition of things he is investigating. He believes that his

hypotheses are true to the extent that the theory from which they were derived is adequate. However, theories are only suppositions about the true nature of things, and thus hypotheses derived from theories must also be regarded as just tentative suppositions about things until they have been tested. Testing hypotheses means subjecting them to *confirmation* or *disconfirmation*.

1.2.2.2 Null Hypothesis

Null hypotheses are, in a sense, the reverse of research hypotheses. They are also statements about the reality of things, except that they serve to *deny* what is explicitly indicated in a given research hypothesis. For example, if the researcher states as his research hypothesis that "the average grade of Assumption University psychology students is 90%," he may also state a null hypothesis that can be used to evaluate the accuracy of his research hypothesis. The null hypothesis would be "the average grade of Assumption University psychology students is *not* 90%." If the researcher can demonstrate that the average grade of these students is at or near 90%, then he concludes that the null hypothesis is refuted or regarded as not true. If the researcher rejects the null hypothesis, then logically the statement that "the average grade of AU psychology students is 90%" is supported. In other words, the researcher constructs a situation that contains two contradictory statements, namely "the average grade of AU psychology students is/is not 90%." These statements are worded in such a way that they are mutually exclusive, that is, the confirmation of one is the denial or refutation of the other. *Both cannot coexist simultaneously*. In social research, when we test a hypothesis, we test the null hypothesis—the statement that there is no difference between groups or no relationship between variables. Whether our research hypothesis is supported or refuted depends on the outcome of the test of the null hypothesis.

1.2.3 Testing Hypotheses

Testing hypotheses mean subjecting them to some sort of empirical scrutiny to determine if they are supported or refuted by what the researcher observes. Suppose a researcher has collected data pertaining to her hypotheses and recorded what she has found. The question then is, what are the bases upon which she concludes that her hypotheses are supported or refuted? What would prevent two researchers working independently to arrive at quite different conclusions after studying identical information? To avoid, or at least reduce the amount of subjectivity that exists when social scientists interpret what they have found, they employ a decision rule that specifies the conditions under which the researchers will decide to refute or support the hypotheses they are testing. This decision rule is called the *level of significance*.

1.2.4 Level of Significance

When a difference in characteristics (e.g., IQ, verbal test, personality traits) between two groups is observed, at what point do we conclude that the difference is a *significant* one (i.e., not due to chance)? We are usually going to observe differences between various people regarding commonly held characteristics, but how do we decide that the differences mean anything important to us? That is, how do we judge these differences? What may be a significant difference to one researcher may not be considered as such by another researcher. In order to introduce greater objectivity into our interpretations of observations, we establish the *level of significance*. To state a level of significance is to state a probability level at which the researcher will decide to accept or reject the null hypothesis. How do levels of significance operate to enable researchers to make decisions about their observations? To answer this question, we need to look at *probability theory*.

Probability in social research is concerned with the possible outcomes of experiments, that is, the likelihood that one's observations or results are expected or unexpected. For example, if we were to flip a coin 100 times, we would expect that we would get 50 heads and 50 tails. This expected outcome assumes that the coin is not biased. A biased coin however, would reflect a greater proportion of either heads or tails compared with an unbiased one. Now, if we actually flip the coin 100 times, does the distribution of heads and tails differ from what would be expected? Determining what is expected is based on the possible outcomes associated with our observations. In the flip of a coin, there are only two possible outcomes, heads or tails, and if the coin is unbiased, then we would expect heads to come up 50% of the time and tails to come up 50% of the time. In other words, the probability of getting heads is $P = 0.5$, and the probability of getting tails is $P = 0.5$. This is what would be expected from flipping an unbiased coin. However, if we flip the coin 100 times, and it comes up heads 60 times and tails 40 times, will we say that this distribution is so different from our expected distribution as to conclude that the coin is biased? How about an outcome of 70 heads and 30 tails, or 75 heads and 25 tails? The question here is, at what point do we decide to regard an outcome as significantly different from what we would expect according to probability? In social research, we set the probability or significance level at ≤ 0.05. In other words, using our coin example, only when the coin comes up heads (or tails) 95 times or more out of 100 throws will we say the coin is biased. In statistical terms, we say that the observed outcome is *statistically significant* if the probability of the difference *occurring by chance* is less than five times out of a hundred (i.e., we conclude that something else other than chance has affected the outcome).

1.2.5 Type I and Type II Errors

Type I error. We use the level of significance to help us to decide whether to accept or reject the null hypothesis. In the coin example, when we set the level of significance at 0.05, we will reject only the null hypothesis of

neutrality if the coin turns up heads 95 times or more (or tails 5 times or less) in 100 throws. If, for example, the coin turns up heads 96 times out of 100 throws, and we reject the null hypothesis of neutrality, isn't it still possible that we could be wrong in rejecting the null hypothesis? Is it not possible that we have, in fact, obtained a statistically rare occurrence by chance? The answer to this question must be "yes."

If we reject the null hypothesis when it is true and should not be rejected, we have committed Type I error. In testing a hypothesis, the level of significance set to decide whether to accept or reject the null hypothesis is the amount of Type I error the researcher is willing to permit. When we employ the 0.05 level of significance, approximately 5% of the time we will be wrong when we reject the null hypothesis and assert its alternative. It would seem then that in order to reduce this type of error, we should set the rejection level as low as possible. For example, if we were to set the level of significance at 0.001, we would risk a Type I error only about one time in every thousand. However, the lower we set the level of significance, the greater is the likelihood that we will make a Type II error.

Type II error. If we fail to reject the null hypothesis when it is actually false, we have committed Type II error. This type of error is far more common than a Type I error. For example, in the coin experiment, if we set the level of significance at 0.01 for rejecting the null hypothesis, and if the coin turns up heads 98 times out of 100 throws, the researcher will not be able to reject the null hypothesis of neutrality. That is, based on these observations, the researcher cannot claim that the coin is biased (the alternative hypothesis) even though it may very well be. Only if the coin turns up heads 99 times or more will the researcher be able to reject the null hypothesis.

It is clear then that Type I and Type II errors cannot be eliminated. They can be minimized, but minimizing one type of error will increase the probability of committing the other error. The lower we set the level of significance, the lower the likelihood of a Type I error, and the higher the likelihood of a Type II error. Conversely, the higher we set the level of significance, the higher the likelihood of a Type I error, and the lower the likelihood of a Type II error.

1.2.5.1 Calculating Type I Error

As stated above, when testing a hypothesis, the level of significance set to decide whether to accept or reject the null hypothesis is the amount of Type I error the researcher is willing to permit. The 0.05 level of significance is the conventional level set for rejecting the null hypothesis, which means that approximately 5% of the time we will be wrong when we reject the null hypothesis and assert its alternative. If the test of the hypothesis, requires multiple runs of the same test, then the rate of Type I error increases. For example, if a researcher is interested in investigating whether there are gender differences for 10 personality variables, then a series of 10 independent *t* tests are required to test each personality criterion variable separately. As each *t* test is run and

tested for statistical significance, the rate of Type I error increases. Assuming each hypothesis is independent and a traditional Type I error rate of 0.05, then the Type I error rate could be estimated as:

$$\begin{aligned}\text{Type I error rate} &= 1-(1-\text{Type I error rate})^{k}\\ &= 1-(1-.05)^{10}\\ &= .40\end{aligned}$$

A Type I error rate of this size substantially exceeds the conventional and accepted rate of 0.05.

1.3 Test Selection

When a researcher is ready to test a specific hypothesis generated from a theory, or to answer a research question posed, the researcher is faced with the task of choosing an appropriate statistical procedure. Choosing an appropriate statistical procedure to test a specific hypothesis is important for a number of reasons. First, for any hypothesis posited, the statistical procedure chosen must offer a legitimate test of the hypothesis; otherwise no meaningful interpretation of the results can be made. For example, if a study hypothesizes *differences* in mean scores between groups, it will make no sense for the researcher to choose a test of *relationship* between pairs of variables. Second, choosing an inappropriate statistical procedure can also mean choosing a less-than-robust test that fails to detect significant differences between group-scores or significant relationships between pair of variables (i.e., increasing the probability of committing Type II error).

While the researcher is faced with a multitude of statistical procedures to choose from, the choice of an appropriate statistical test is generally based on just two primary considerations: (1) the nature of the hypothesis and (2) the levels of measurement of the variables to be tested.

1.3.1 The Nature of Hypotheses: Test of Difference versus Test of Relationship

In choosing an appropriate statistical test, the first issue that the researcher must consider is the nature of the hypothesis. Is the intention of the hypothesis to test for differences in mean scores between groups, or is it testing for relationships between pairs of variables?

1.3.1.1 Test of Difference

Testing for differences means that the researcher is interested in determining whether differences in mean scores between groups are due to chance factors

or to real differences between the groups as a result of the study's experimental treatment. Say, for example, a company whose main business is to help people increase their level of IQ wants to market a new drug called IQADD that is supposed to raise IQ scores. To test the effectiveness of this drug, the company conducts an experiment where 64 random people are randomly assigned to two groups of $n = 32$ people each. The first group is then administered a specific dosage of IQADD while the second group is given a placebo. The company then measures the IQ scores (obtained from a standardized IQ test) for all 64 people to see which group has the higher mean score. Suppose that the IQADD group has a mean IQ score of 134 and the placebo group has a mean IQ score of 110. The question that this company must now ask is whether this difference of 24 IQ points in favor of the IQADD group is so big that it can rule "chance" out as an explanation (and therefore, must be due to the effect of the IQADD drug), or whether it is so small that chance variation in IQ between the two groups is still a legitimate explanation. In order to answer this question, the company must choose a statistical test that will appropriately test for the *difference* in mean IQ scores between the two groups.

1.3.1.2 Test of Relationship

Testing of relationships among two or more variables asks the question, "Are variations in variable X associated with variations in variable Y?" For example, do students who do well in high school also perform well in university? Do parents with high intelligence tend to have children of high intelligence? Is there a relationship between the declared dividend on stocks and their paper value in the exchange? Is there a relationship between socio-economic class and recidivism in crime? All these questions concern the relationships among variables, and to answer these questions, researchers must choose statistical tests that will appropriately test for the *relationships* among these variables.

1.3.2 Levels of Measurement

In addition to considering the nature of the hypothesis to be tested (difference or relationship), the researcher must also consider the measurements of the variables to be tested. This is because the levels at which the variables are measured determine which statistical test is used to analyze the data. Most typically, variables in the behavioral sciences are measured on one of four scales: *nominal*, *ordinal*, *interval*, or *ratio* measurements. These four types of scales differ in the number of the following attributes they posses: *magnitude*, *an equal interval between adjacent units*, and *an absolute zero point*.

1.3.2.1 Nominal Scale

This is the lowest level of measurement and involves simply categorizing the variable to be measured into one of a number of discrete categories. For

instance, in measuring "ethnic origin," people may be categorized as American, Chinese, Australian, African, or Indian. Once people have been categorized into these categories, all people in the same category (e.g., those categorized as Americans) are equated on the measurement obtained, even though they may not possess the same amount of the characteristics. Numbers can be assigned to describe the categories, but the numbers are only used to name/label the categories. They have no magnitude in terms of quantitative value.

1.3.2.2 Ordinal Scale

This level of measurement involves ordering or ranking the variable to be measured. For example, people may be asked to rank-order four basketball teams according to their skills. Thus, a rank of 1 is assigned to the team that is the most skillful, a rank of 2 to the team that exhibits the next greatest amount of skill, and so forth. These numbers allow the researcher to quantify the magnitude of the measured variable, by adding the arithmetic relationships "greater than" and "less than" to the measurement process. While ordinal scales allow one to differentiate between rankings among the variable being measured, they do not permit determination of how much of a real difference exists in the measured variable between ranks. The basketball team that is ranked 1 is considered to be better than the team that is ranked 2, but the rankings provide no information as to *how much better* the first ranked basketball team is. In other words, the intervals between the ranks are not meaningful.

1.3.2.3 Interval Scale

This level of measurement involves being able to specify how far apart two stimuli are on a given dimension. On an ordinal scale, the difference between the basketball team ranked first and the team ranked second does not necessarily equal the distance between teams ranked third and fourth. On an interval scale however, differences of the same numerical size in scale values *are* equal. For example, on a standardized intelligence measure, a 10-point difference in IQ scores has the same meaning anywhere along the scale. Thus, the difference in IQ test scores between 80 and 90 is the same as the difference between 110 and 120. However, it would not be correct to say that a person with an IQ score of 100 is *twice* as intelligent as a person with a score of 50. The reason for this is because intelligence test scales (and other similar interval scales) do not have a true *zero* that represents a complete absence of intelligence.

1.3.2.4 Ratio Scale

This level of measurement replaces the arbitrary zero point of the interval scale with a true zero starting point that corresponds to the absence of the variable being measured. Thus, with a ratio scale, it is possible to state that a variable has, for example, twice, half, or three times as much of the variable measured than another. Take weight as an example. Weight has a true zero point (a weight of

zero means that the object is weightless) and the intervals between the units of measurement are equal. Thus, the difference between 10 and 15 g is equal to the difference between 45 and 50 g, and 80 g is *twice* as heavy as 40 g.

1.3.3 Choice of Test

Once the researcher has decided on the nature of the hypothesis to be tested (*test of difference* or *test of relationship*) and the levels of measurement of the variables to be included in the analysis (*nominal, ordinal, interval, ratio*), the next step is to choose an appropriate statistical test for analyzing the data. Table 1.1 presents a "Test Selection Grid" that will aid in the selection of an appropriate statistical procedure. From the Test Selection Grid, the researcher chooses a test by considering (1) the purpose of the statistical analysis, (2) the levels of measurement of the variables, (3) the number of sets of scores to be included in the analysis, and (4) whether the sets of scores are related or independent. It should be noted that the Test Selection Grid is restricted to only a number of the most commonly used procedures. There are far more statistical procedures available in SPSS for Windows.

TABLE 1.1

Test Selection Grid

	Relationship	One Set of Scores	Related Two Sets	Independent Two Sets	More than Two Sets
Nominal	Point biserial (r*pb*) (true dichotomy)	Single variable chi-square (χ^2) test	McNemar significance of change χ^2	Chi-square test of association	Chi-square goodness of fit
	Biserial (r*b*) (artificial dichotomy)				
	Phi (rΦ) (true/true)				
	Tetrachoric (r*t*) (artificial dichotomy /artificial dichotomy)				
Ordinal	Spearman's rho	Kolmogorov-Smirnov test for ranked data	Wilcoxon matched-pairs signed-ranks test	Mann-Whitney *U* test	Kruskal-Wallis test
Interval/ Ratio	Pearson's product-moment correlation	One-sample *t*-test	Related samples *t*-test	Independent samples *t*-test	One-way ANOVA (independent)
	Linear regression				Factorial ANOVA (independent)
	Multiple regression				Multivariate ANOVA (related)

2

Introduction to SPSS

2.1 Introduction

When SPSS, Inc., an IBM Company, was conceived in 1968, it stood for *Statistical Package for Social Sciences.* Since the company's purchase by IBM in 2009, IBM has decided to simply use the name SPSS to describe its core product of predictive analytics. IBM describes predictive analytics as tools that *help connect data to effective action by drawing reliable conclusions about current conditions and future events.*

SPSS is an integrated system of computer programs designed for the analysis of social sciences data. It is one of the most popular of the many statistical packages currently available for statistical analysis. Its popularity stems from the fact that the program:

- Allows for a great deal of flexibility in the format of data
- Provides the user with a comprehensive set of procedures for data transformation and file manipulation
- Offers the researcher a large number of statistical analyses commonly used in social sciences

For both the beginner and the advanced researchers, SPSS is an indispensable tool. It is not only an extremely powerful program, but also relatively easy to use once the researcher has been taught the rudiments. The Windows version of SPSS has introduced a *point-and-click* interface that allows the researcher to merely point-and-click through a series of windows and dialog boxes to specify the kind of analysis required, and the variables involved. This method eliminates the need to learn the very powerful syntax or command language used to execute SPSS analyses (in the older MS-DOS versions), and has proven to be highly popular for those researchers with little or no interest in learning the *syntax method*. Nevertheless, SPSS for Windows has retained the syntax method that permits the researcher to execute SPSS analyses by typing commands.

A question that is often asked by the beginner researcher is, "Which method of running SPSS is better?" Both the Windows and syntax method have their own advantages and disadvantages, and these are discussed in Section 2.3.

This chapter has been written with the beginner student and researcher in mind, and provides an overview on the two most basic functions of SPSS: (1) how to set up data files in SPSS for Windows, and (2) conducting SPSS analysis via the Windows method and the syntax method.

2.2 Setting Up a Data File

Suppose a survey has been conducted to investigate the extent to which Thais agree with increases in government spending in the three areas of defense, social security, and childcare services. Responses to these three issues were obtained from the questionnaire presented in Table 2.1.

TABLE 2.1

Survey Questionnaire

a. Gender 1. _____ Male 2. _____ Female

b. Age _____ (in years)

c. The following three statements relate to increases in government spending in the areas of defense, social security, and childcare services. Please consider these three statements carefully and then decide your level of agreement with the government's decision to increase spending. Please indicate your level of agreement by circling the number on each six-point scale.

i. Increased spending on defense

1_____	2_____	3_____	4_____	5_____	6_____
Strongly Disagree	Moderately Disagree	Barely Disagree	Barely Agree	Moderately Agree	Strongly Agree

ii. Increased spending on social security

1_____	2_____	3_____	4_____	5_____	6_____
Strongly Disagree	Moderately Disagree	Barely Disagree	Barely Agree	Moderately Agree	Strongly Agree

iii. Increased spending on childcare services

1_____	2_____	3_____	4_____	5_____	6_____
Strongly Disagree	Moderately Disagree	Barely Disagree	Barely Agree	Moderately Agree	Strongly Agree

2.2.1 Preparing a Codebook

Prior to data entry, it will be useful to prepare a codebook that contains the names of the variables in the questionnaire, their corresponding SPSS variable names, and their coding instructions. An important purpose of such a codebook is to allow the researcher to keep track of all the variables in the survey questionnaire and the way they are defined in the SPSS data file. Table 2.2 presents the codebook for the questionnaire above.

2.2.2 Data Set

Table 2.3 presents the responses obtained from a sample of 10 respondents to the questionnaire.

2.2.3 Creating an SPSS Data File

The following steps demonstrate how the data presented in Table 2.3 are entered into an SPSS data file.

TABLE 2.2

Codebook

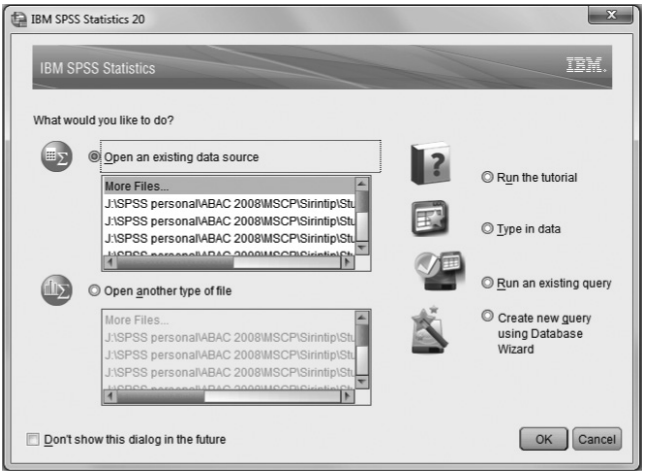
Variable	SPSS Variable Name	Code
Gender	Gender	1 = male 2 = female
Age	Age	Age in years
Defense	Defense	1 = Strongly Disagree 2 = Moderately Disagree 3 = Barely Disagree 4 = Barely Agree 5 = Moderately Agree 6 = Strongly Agree
Social Security	Social	1 = Strongly Disagree 2 = Moderately Disagree 3 = Barely Disagree 4 = Barely Agree 5 = Moderately Agree 6 = Strongly Agree
Childcare Services	Child	1 = Strongly Disagree 2 = Moderately Disagree 3 = Barely Disagree 4 = Barely Agree 5 = Moderately Agree 6 = Strongly Agree

TABLE 2.3

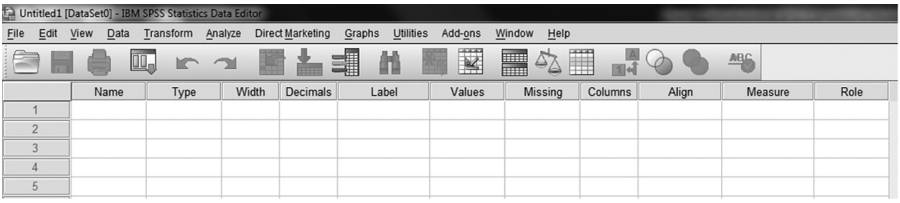
Raw Data

Gender	Age	Defense	Social	Child
1	24	4	2	1
1	18	5	1	4
2	33	2	5	6
1	29	5	3	4
2	26	3	5	5
2	19	2	5	2
1	36	4	4	3
2	34	3	6	6
1	20	3	5	1
2	21	2	5	3

1. When the SPSS program is launched, the following window will open.

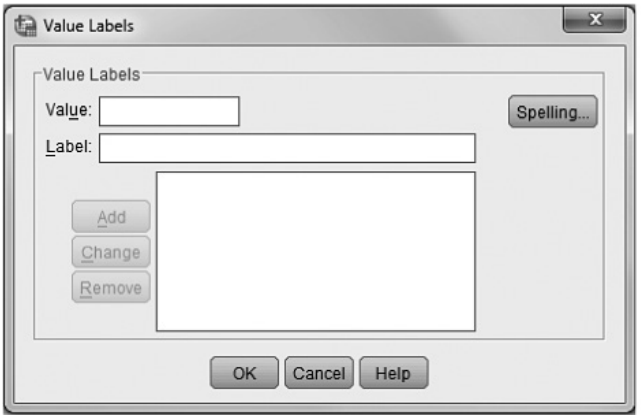


Since the purpose of the present exercise is to create a new data file, close this window by clicking Cancel. The following **Untitled1 [DataSet0]—IBM SPSS Statistics Data Editor** screen will then be displayed.

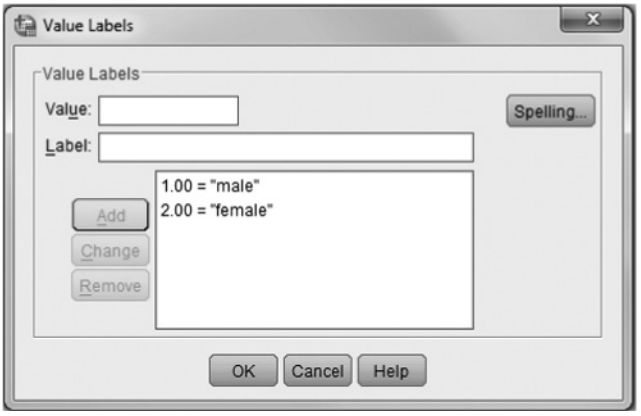


2. Prior to data entry, the variables in the data set must be named and defined. In the **Untitled1 [DataSet0]—IBM SPSS Statistics Data**

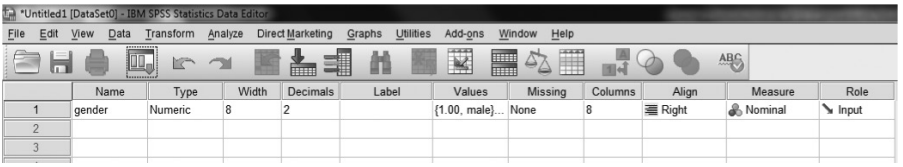
Editor screen, the names of the variables are listed down the side (under the **Name** column), with their characteristics listed along the top (Type, Width, Decimals, Label, Values, Missing, Columns, Align, Measure). The codebook presented in Table 2.2 will serve as a guide in naming and defining the variables. For example, the first variable is named GENDER and is coded 1 = male and 2 = female. Thus, in the first cell under **Name** in the **Data Editor** screen, type in the name GENDER. To assign the coded values (1 = male, 2 = female) to this variable, click the corresponding cell under **Values** in the **Data Editor** screen. Click the shaded area to open the following **Value Labels** window.



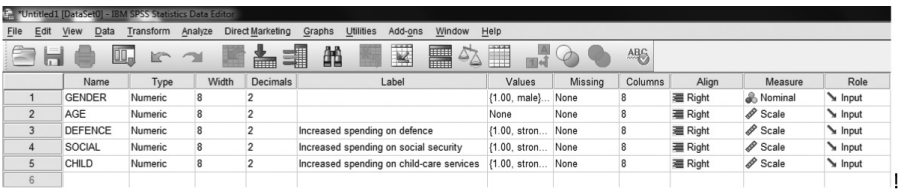
3. In order to define the code for male respondents, type **1** in the **Value:** cell, and in the **Label:** cell, type **Male**. Next, click Add to complete the coding for the male respondents. For female respondents, type **2** in the **Value:** cell, and in the **Label:** cell, type **Female**. Next, click Add to complete the coding for the female respondents. The completed **Value Labels** window is presented below.



Next, click OK to complete the coding for the GENDER variable and to return to the **Untitled1 [DataSet0]—IBM SPSS Statistics Data Editor** screen below.



4. Repeat the above coding procedure for the rest of variables in the codebook. Please note that the AGE variable is a *continuous* variable and therefore has no coded values.
5. If the researcher wishes to attach a label to a variable name (to provide a lengthy description for that variable), this can be done by typing a label in the corresponding cell in the **Label** column. For example, the researcher may wish to attach the label "**Increased spending on defense**" to the variable DEFENSE. This label will be printed in the analysis output generated by SPSS. The following **Untitled1 [DataSet0]—IBM SPSS Statistics Data Editor** screen displays the names of all the variables listed in the codebook, and where relevant, their **Label** and **Values** codes.

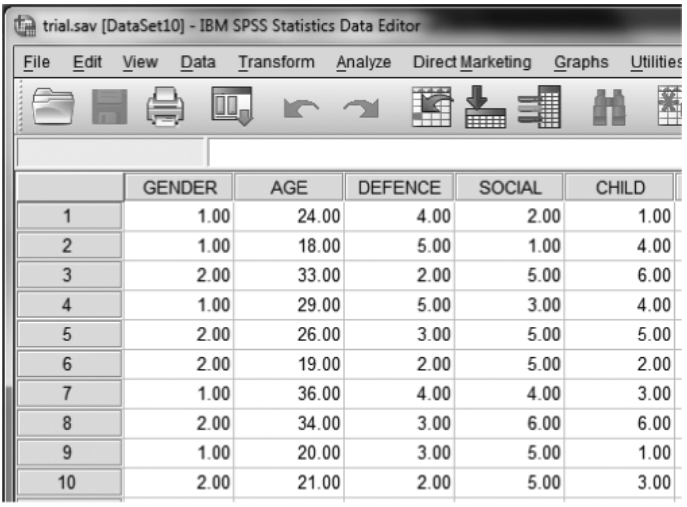


2.2.4 Data Entry

The data can only be entered via the **Data View** screen. Switch the present **Variable View** to **Data View** by clicking the **Data View** tab Data View Variable View at the bottom left-hand corner of the screen. In the **Data View** screen, the rows represent the respondents, and the columns represent the variables. Beginning with the first data cell (row 1, column 1), type in the data presented in Table 2.3. The following **Data View** screen shows the data obtained from the 10 respondents.

2.2.5 Saving and Editing a Data File

Once the data entry is completed, the data file can be saved. From the menu bar, click **File**, then **Save As**. Once it has been decided where the data file will



	GENDER	AGE	DEFENCE	SOCIAL	CHILD
1	1.00	24.00	4.00	2.00	1.00
2	1.00	18.00	5.00	1.00	4.00
3	2.00	33.00	2.00	5.00	6.00
4	1.00	29.00	5.00	3.00	4.00
5	2.00	26.00	3.00	5.00	5.00
6	2.00	19.00	2.00	5.00	2.00
7	1.00	36.00	4.00	4.00	3.00
8	2.00	34.00	3.00	6.00	6.00
9	1.00	20.00	3.00	5.00	1.00
10	2.00	21.00	2.00	5.00	3.00

be saved, type a name for the file. As this is a data file, SPSS will automatically append the suffix .**SAV** to the data file name (e.g., **TRIAL.SAV**).

To edit an existing file, click **File**, then **Open**, and then **Data** from the menu bar. Scroll through the names of the data files and double-click on the data file to open it.

2.3 SPSS Analysis: Windows Method versus Syntax Method

Once the SPSS data file has been created, the researcher can conduct the chosen analysis either through the Windows method (point-and-click) or the syntax method. The primary advantage of using the Windows method is clearly its ease of use. With this method, the researcher accesses the pull-down menu by clicking Analyze in either the **Data View** or **Variable View** mode, and then point-and-clicks through a series of windows and dialog boxes to specify the kind of analysis required and the variables involved. There is no need to type in any syntax or commands to execute the analysis. Although this procedure seems ideal at first, it is not always the method of choice for the more advanced and sophisticated users of the program. Rather, there is clearly a preference for the syntax method among these users. This preference stems from several good reasons from learning to use the syntax method.

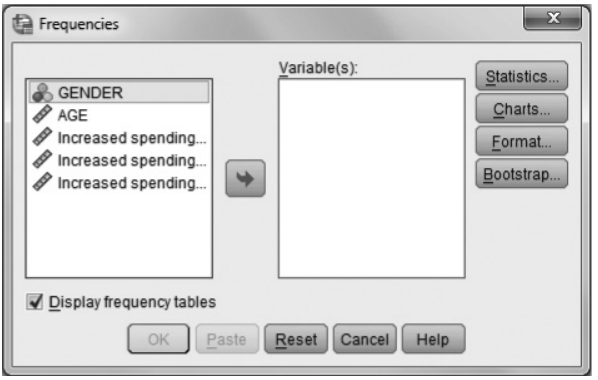
First, when conducting complex analysis, the ability to write and edit command syntax is advantageous. For example, if a researcher uses an incorrect a syntax command for a complex analysis and wants to go back and rerun it with minor changes, or if the researcher wishes to repeat an analysis multiple times with minor variations, it is often more efficient to write and

edit the syntax command directly than to repeat the Windows pull-down menu sequences. Second, from my teaching experience with SPSS, I believe that students have a better "feel" for statistics if they have to write syntax commands to generate the specific statistics they need, rather than merely relying on pull-down menus. In other words, it provides a better learning experience. Finally, and perhaps most important, several SPSS procedures are available only via the syntax method.

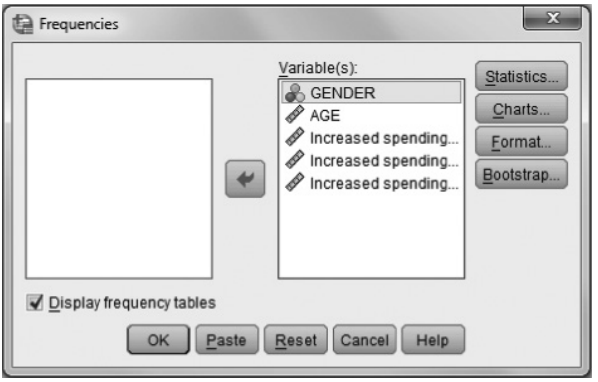
2.3.1 SPSS Analysis: Windows Method

Once the data have been entered, the researcher can begin the data analysis. Suppose the researcher is interested in obtaining general **descriptive statistics** for all of the variables entered in the data set TRIAL.SAV (see Section 2.2.5).

1. From the menu bar, click **Analyze**, then **Descriptive Statistics**, and then **Frequencies.** The following **Frequencies** window will open.

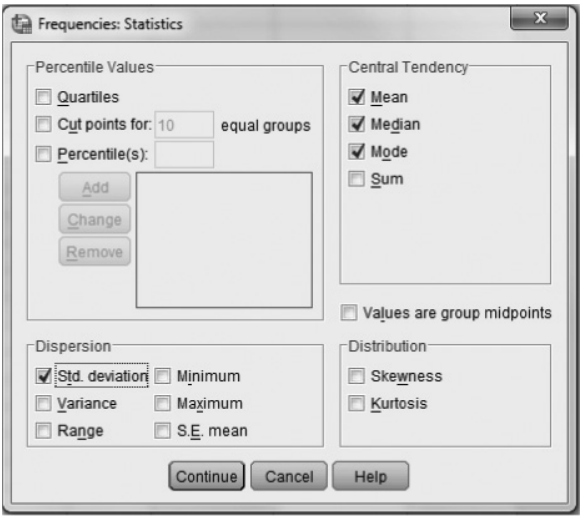


2. In the left-hand field containing the study's five variables, click (highlight) these variables, and then click ▸ to transfer the selected variables to the **Variable(s):** field.

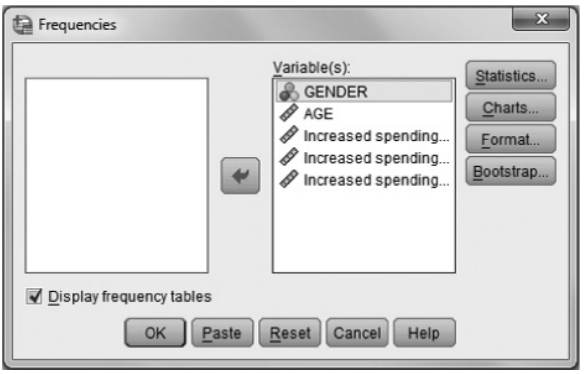


3. Click Statistics... to open the **Frequencies: Statistics** window below. Suppose the researcher is only interested in obtaining statistics for the **Mean**, **Median**, **Mode**, and **Standard Deviation** for the five variables. In the **Frequencies: Statistics** window, check the fields related to these statistics.

   Next click Continue.

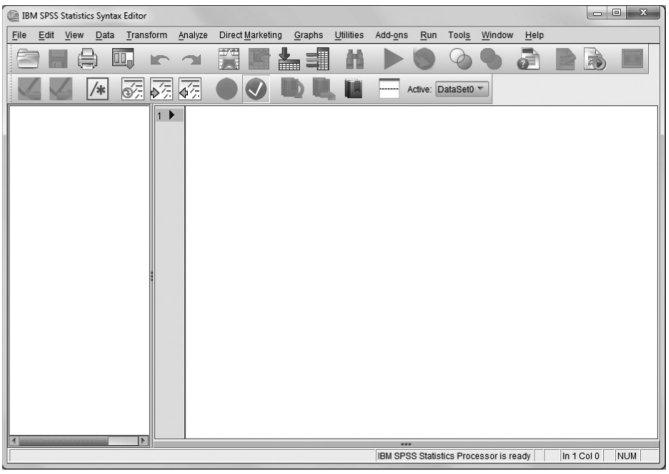


4. When the **Frequencies** window opens, run the analysis by clicking OK. See Table 2.4 for the results.



2.3.2 SPSS Analysis: Syntax Method

1. From the menu bar, click **File**, then **New**, and then **Syntax.** The following **IBM SPSS Statistics Syntax Editor** window will open.

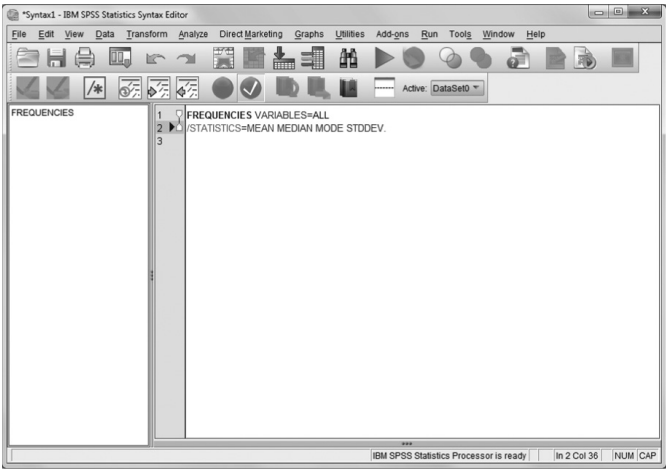


2. Type the **Frequencies** analysis syntax command in the **IBM SPSS Statistics Syntax Editor** window. If the researcher is interested in obtaining all descriptive statistics (and not just the mean, median, mode, and standard deviation), then replace the syntax

   ```
   /STATISTCS = MEAN MEDIAN MODE STDDEV.
   ```

   with

   ```
   /STATISTICS = ALL.
   ```



3. To run the Frequencies analysis, click ▶ or click Run and then **All**.

NOTE: The Appendix presents a summary of the SPSS syntax files employed for all the examples in this book.

2.3.3 SPSS Output

TABLE 2.4

Frequencies Output

Frequencies

Statistics

		Gender	Age	Increased Spending on Defence	Increased Spending on Social Security	Increased Spending on Childcare Services
N	Valid	10	10	10	10	10
	Missing	0	0	0	0	0
Mean		1.5000	26.0000	3.3000	4.1000	3.5000
Median		1.5000	25.0000	3.0000	5.0000	3.5000
Mode		1.00[a]	18.00[a]	2.00[a]	5.00	1.00[a]
Std. Deviation		.52705	6.66667	1.15950	1.59513	1.84089

[a] Multiple modes exist. The smallest value is shown.

Frequency Table

Gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	5	50.0	50.0	50.0
	Female	5	50.0	50.0	100.0
	Total	10	100.0	100.0	

Age

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	18.00	1	10.0	10.0	10.0
	19.00	1	10.0	10.0	20.0
	20.00	1	10.0	10.0	30.0
	21.00	1	10.0	10.0	40.0
	24.00	1	10.0	10.0	50.0
	26.00	1	10.0	10.0	60.0
	29.00	1	10.0	10.0	70.0
	33.00	1	10.0	10.0	80.0
	34.00	1	10.0	10.0	90.0
	36.00	1	10.0	10.0	100.0
	Total	10	100.0	100.0	

(*Continued*)

TABLE 2.4 (*Continued*)

Frequencies Output

Increased Spending on Defense

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	moderately disagree	3	30.0	30.0	30.0
	barely disagree	3	30.0	30.0	60.0
	barely agree	2	20.0	20.0	80.0
	moderately agree	2	20.0	20.0	100.0
	Total	10	100.0	100.0	

Increased Spending on Social Security

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	strongly disagree	1	10.0	10.0	10.0
	moderately disagree	1	10.0	10.0	20.0
	barely disagree	1	10.0	10.0	30.0
	barely agree	1	10.0	10.0	40.0
	moderately agree	5	50.0	50.0	90.0
	strongly agree	1	10.0	10.0	100.0
	Total	10	100.0	100.0	

Increased Spending on Childcare Services

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	strongly disagree	2	20.0	20.0	20.0
	moderately disagree	1	10.0	10.0	30.0
	barely disagree	2	20.0	20.0	50.0
	barely agree	2	20.0	20.0	70.0
	moderately agree	1	10.0	10.0	80.0
	strongly agree	2	20.0	20.0	100.0
	Total	10	100.0	100.0	

2.3.4 Results and Interpretation

The **Statistics** table presents the requested mean, median, mode, and standard deviation (SD) statistics for the five variables. The **Gender** variable is a nominal (categorical) variable and as such, its mean, median, and standard deviation statistics are not meaningful. The remaining four variables of **Age**, **Defense**, **Social**, and **Child** are measured at least at the ordinal level (i.e., they are continuous variables), and as such their mean, median, and standard deviation statistics can be interpreted.

The results presented in the **Statistics** table show that the 10 respondents in the survey have a mean age of 26 years and a median age of 25 years. Since

there is no one age that occurs more frequently than others, SPSS presents the lowest age value of 18 as the mode. For the three variables of "support for increased spending" on defense, social security, and childcare services, the results show that support for increased spending for social security is the highest (mean = 4.10; median = 5.00), followed by childcare services (mean = 3.50; median = 3.50), and defense (mean = 3.30; median = 3.00). The results also show that the variables of Defense and Child have multiple modes, and as such, SPSS has presented their lowest values (defense: mode = 2.00; child: mode = 1.00). The Social variable has a single mode of 5.00.

For the Age variable, the standard deviation shows that its average deviation (dispersion) from the mean is 6.66 years. For the Defense, Social, and Child variables, the results show that support for increased spending on childcare services has the largest average variation (SD = 1.84) from its mean score. The standard deviation scores for support for increased spending for defense (SD = 1.59) and social security (SD = 1.59) are similar.

The **frequency** table presents the breakdown of the frequency distributions for the five variables (Gender, Age, Defense, Social, Child). For each variable, the frequency table presents (1) the **frequency** of occurrence for each value within that variable, (2) the frequency for each value expressed as a **percentage** of the total sample, (3) the **valid percentage** for each value, controlling for missing cases, and (4) the **cumulative percentage** for each succeeding value within that variable. For example, the Frequency table for the Gender variable shows that there are five males and five females in the sample, and that these two groups represent 50% each of the total sample. Since there are no missing cases, the valid percentage values are identical to the percentage values. *If there are missing cases, then the valid percentage values should be interpreted.* The cumulative percentage presents the percentage of scores falling at or below each score. Thus, for the sample of 10 respondents, the five males in the sample represent 50% of the sample, and the additional five females represent a cumulative percentage of 100%.

The frequency tables for the Age, Defense, Social, and Child variables are interpreted in exactly the same way.

2.4 Missing Data

Encountering missing data in research is part and parcel of data analysis. The problem arises for a number of reasons; examples include subjects dropping out of an experimental condition because of boredom/fatigue; refusal to answer a particular question on a survey questionnaire; death; and refusal to participate in the post-test of a longitudinal study. Whatever the reasons, analysis of a "missing data set" can be problematic if the pattern of missing data is non-random. Any statistical results based on such data would

be biased to the extent that the variables included in the analysis are influenced by the pattern of non-randomness of the missing data. According to Tabachnick and Fidell (2001), if only a few data points (≤5%) are missing in a random pattern from a large data set, the problems are less serious and almost any procedure for handling missing values yields similar results. Thus, when screening data, it is important to identify the pattern of missing data.

2.4.1 Patterns of Missing Data

Suppose two variables (*X* and *Y*) are collected. *X* has no missing data, but *Y* does have some missing data.

Missing at random (MAR)—Missing data are termed MAR if the missing values of *Y* depend on *X*, but not on *Y*. Another way of stating this is that the missing values for *Y* can be explained by *X* in the data set. However, after accounting for *X*, the missing values of *Y* are random. An example can illustrate this. Assume that we know the marital status of respondents (the *X* variable), and we ask them about their income (the *Y* variable). Income is MAR if the probability of missing data on income depends on marital status, but within each category of marital status (single, married, divorced) the probability of missing values for income is unrelated to the value of income.

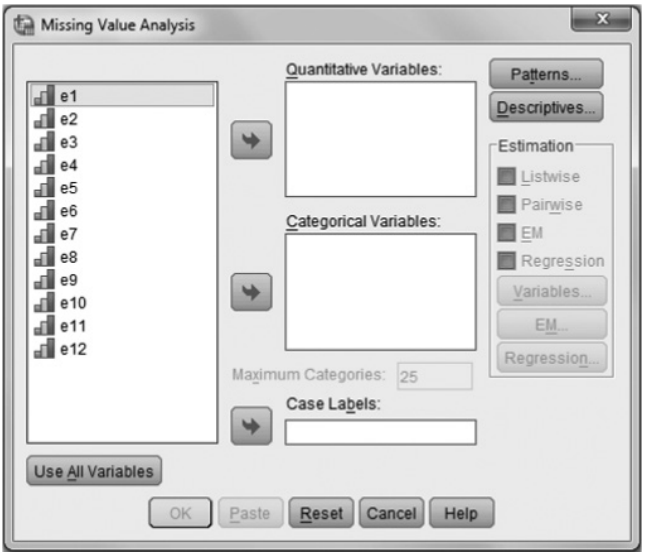
Missing completely at random (MCAR)—This is a higher level of randomness where the missing values of *Y* are unrelated to the value of *Y* itself or to any other variables in the data set. Using the income example, income is MCAR if respondents who do not report their income have the same average income as respondents who do report their income. In other words, there is no relationship at all (complete randomness) between missing values on the income variable and the values of other variables.

2.4.1.1 Test for Patterns in Missing Data

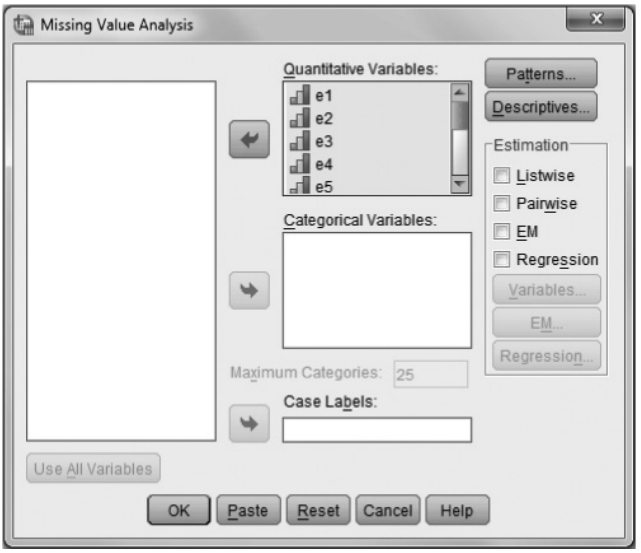
Consider the data set **MISSING.SAV**. This data set contains 12 items (**e1** to **e12**) written to measure people's attitude toward the issue of euthanasia. Each item was rated on a five-point Likert scale (1 = strongly disagree to 5 = strongly agree) with high scores indicating strong support for euthanasia. The sample consisted of 1928 respondents. As there are missing values in the data set, the SPSS **Missing Values Analysis** program can be used to highlight the pattern of missing values. The analysis includes requesting a *t* test ($\alpha = .05$) to test whether missingness is related any other variables. The *t* test will only be executed for variables with at least 5% of data missing.

2.4.1.2 Windows Method

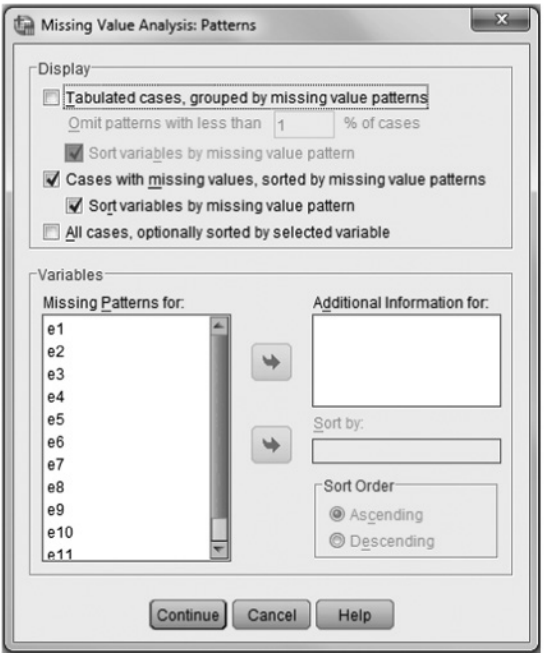
1. From the menu bar, click **Analyze**, then **Missing Value Analysis**. The following **Missing Value Analysis** window will open.



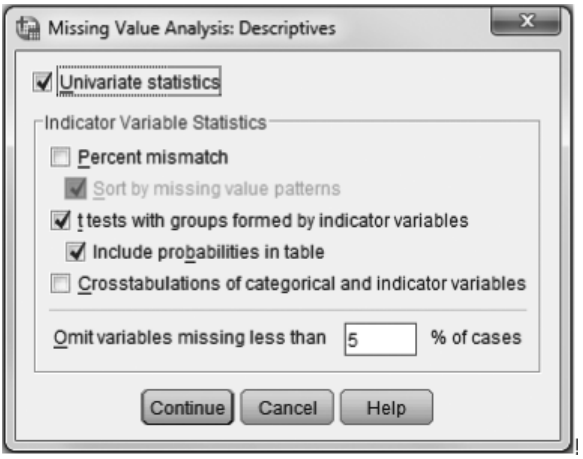
2. Transfer the variables **e1** to **e12** to the **Quantitative Variables:** field by clicking these variables (highlight) and then clicking .



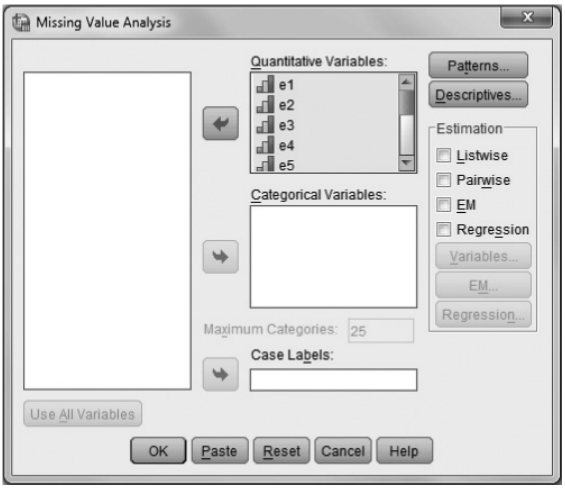
3. Click Patterns... to open the **Missing Value Analysis: Patterns** window. Check the cells for **Cases with missing values, sorted by missing value patterns** and **Sort variables by missing value pattern**. Click Continue to return to the **Missing Value Analysis** window.



4. Click Descriptives... to open the **Missing Value Analysis: Descriptives** window. Check the cells for **Univariate statistics**, ***t* test with groups formed by indicator variables**, and **Include probabilities in table**. In **Omit variables missing less than _% of cases,** type the number 5. Click Continue to return to the **Missing Value Analysis** window.



5. When the **Missing Value Analysis** window opens, click OK to complete the analysis. See Table 2.5 for the results.



2.4.1.3 SPSS Syntax Method

```
MVA E1 TO E12
/TTEST PROB PERCENT = 5.
```

2.4.1.4 SPSS Output

TABLE 2.5

Missing Value Analysis Output

MVA

Univariate Statistics

				Missing		No. of Extremes[a]	
	N	Mean	Std. Deviation	Count	Percent	Low	High
e1	1921	3.4446	1.24297	7	.4	0	0
e2	1921	3.4784	1.26230	7	.4	0	0
e3	1917	2.6922	1.29686	11	.6	0	0
e4	1920	3.3677	1.20153	8	.4	0	0
e5	1920	3.7188	1.21784	8	.4	0	0
e6	1807	3.4660	1.18491	121	6.3	124	0
e7	1919	3.3934	1.20348	9	.5	0	0
e8	1916	3.4034	1.11882	12	.6	125	0
e9	1914	3.0172	1.06221	14	.7	0	0
e10	1915	3.4198	1.14572	13	.7	0	0
e11	1918	2.9213	1.22636	10	.5	0	0
e12	1919	2.7384	1.26900	9	.5	0	0

[a] Number of cases outside the range (Q1 – 1.5*IQR, Q3 + 1.5*IQR).

(*Continued*)

TABLE 2.5 (*Continued*)

Missing Value Analysis Output

Separate Variance t Tests[a]

		e1	e2	e3	e4	e5	e6	e7	e8	e9	e10	e11	e12
e6	t	1.4	1.6	.4	1.1	.9	.	−.3	1.6	1.5	.5	−.8	−1.7
	df	127.8	125.8	126.8	128.9	127.6	.	127.1	126.5	125.2	127.9	126.4	128.1
	P(2-tail)	.171	.122	.723	.285	.389	.	.744	.116	.135	.620	.436	.095
	# Present	1807	1807	1803	1806	1806	1807	1805	1802	1800	1801	1805	1805
	# Missing	114	114	114	114	114	0	114	114	114	114	113	114
	Mean(Present)	3.4543	3.4903	2.6950	3.3749	3.7248	3.4660	3.3911	3.4140	3.0272	3.4231	2.9158	2.7263
	Mean(Missing)	3.2895	3.2895	2.6491	3.2544	3.6228	.	3.4298	3.2368	2.8596	3.3684	3.0088	2.9298

For each quantitative variable, pairs of groups are formed by indicator variables (present, missing).
[a] Indicator variables with less than 5% missing are not displayed.

Missing Patterns (Cases with Missing Values)

			Missing and Extreme Value Patterns[a]											
Case	# Missing	% Missing	e1	e2	e4	e5	e7	e12	e11	e3	e8	e10	e9	e6
11	1	8.3												S
17	1	8.3												S
22	1	8.3												S
25	1	8.3												S
30	1	8.3												S
45	1	8.3												S
61	1	8.3												S
64	1	8.3												S
81	1	8.3												S
89	1	8.3												S
1805	1	8.3									S			
252	1	8.3									S			
1918	1	8.3			S									-
1242	12	100.0	S	S	S	S	S	S	S		S	S	S	S
390	12	100.0	S	S	S	S	S	S	S	S	S	S	S	S
1847	12	100.0	S	S	S	S	S	S	S	S	S	S	S	S
1205	12	100.0	S	S	S	S	S	S	S	S	S	S	S	S
1025	12	100.0	S	S	S	S	S	S	S	S	S	S	S	S
1210	12	100.0	S	S	S	S	S	S	S	S	S	S	S	S
615	12	100.0	S	S	S	S	S	S	S	S	S	S	S	S

– Indicates an extreme low value, while + indicates an extreme high value. The range used is (Q1 – 1.5*IQR, Q3 + 1.5*IQR).
[a] Cases and variables are sorted on missing patterns.

2.4.1.5 Interpretation

The **Univariate Statistics** table shows that all variables (**e1** to **e12**) have missing values. However, only variable **e6** is tested because more than 5% of cases for this variable have missing values. The **Separate Variance t Tests** table shows no systematic relationship between missing values on **e6** and any of the other variables, $p > 0.05$. The **Missing Patterns** table (truncated to save space) shows that case number 11, among others, is missing **e6**, indicated by an **S** in the table. Case number 1242 and those below it in the table have missing values for all 12 variables (**e1** to **e12**).

2.4.2 Dealing with Missing Data

- Listwise deletion

  Cases with missing scores on any variable are excluded from all analysis. The effective sample size with listwise deletion includes only cases with complete records. An advantage of this method is that all analyses are conducted on the same number of cases. However, if missing observations are scattered across many cases, then deletion of cases can mean substantial loss of subjects.

- Pairwise deletion

  Cases are excluded only if they have missing data on variables involved in a particular computation. This means that the effective sample size can vary from analysis to analysis. This feature of pairwise deletion presents a potential problem for the statistical procedure of *structural equation modeling* as it can lead to the problem of a *nonpositive definite matrix* or *singularity*.

- Mean substitution

  This involves replacing a missing score with the overall sample average. This method is simple, but it tends to distort the underlying distribution of the data, reducing variability and making the distributions more peaked at the mean.

- Regression-based imputation

  A missing observation is replaced with a predicted score generated by using multiple regression based on non-missing scores on other variables.

- Expectation-maximization (EM) algorithm

  This procedure involves a two-step process. In the *E* (expectation) step, missing observations are imputed by predicted scores in a series of regressions where each missing variable is regressed on the remaining variables for a particular case. In the *M* (maximization) step, the whole imputed data set is submitted for the maximum likelihood estimation. These two steps are repeated until a stable

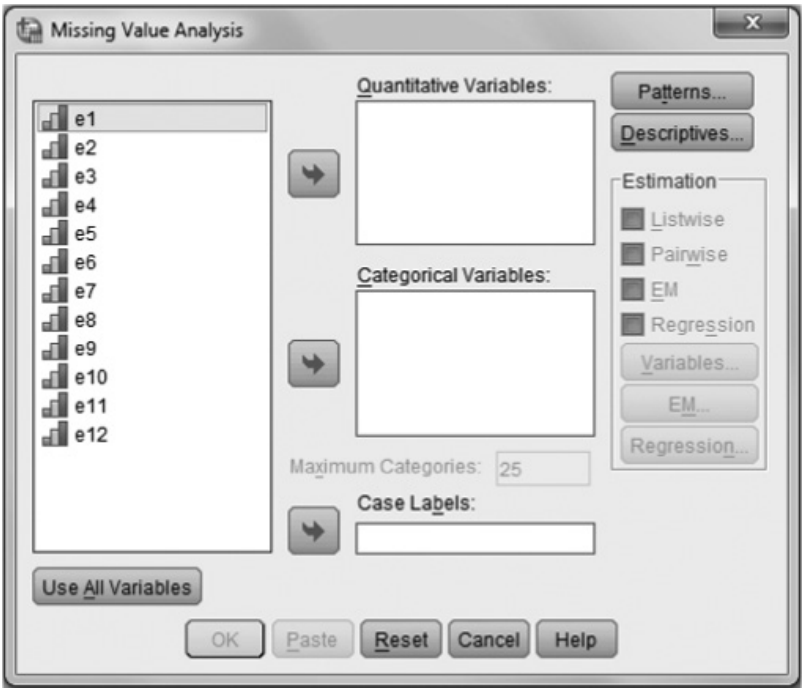
solution is reached across the *M* steps. According to Tabachnick and Fidell (2001), the EM method has the advantages of avoiding impossible matrices (e.g., nonpositive definite matrices), avoiding model overfitting (making the solution look overly good), and producing realistic estimates of variance.

2.4.3 Example of the Expectation-Maximization Method for Handling Missing Data

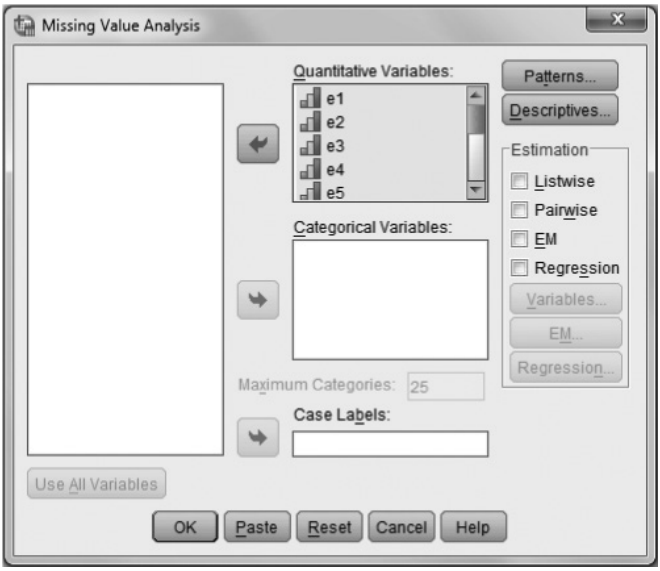
This example will demonstrate the use of the expectation-maximization method to impute scores to replace missing values in the data set **MISSING.SAV**. It can be seen from Table 2.5 that all 12 variables (**e1** to **e12**) have missing values.

2.4.3.1 Windows Method

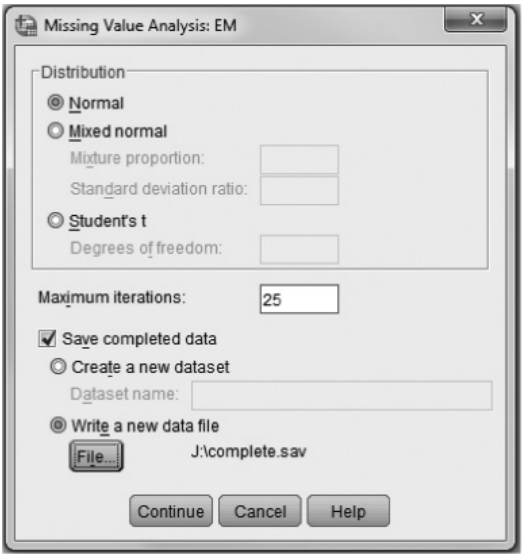
1. From the menu bar, click **Analyze**, then **Missing Value Analysis**. The following **Missing Value Analysis** window will open.



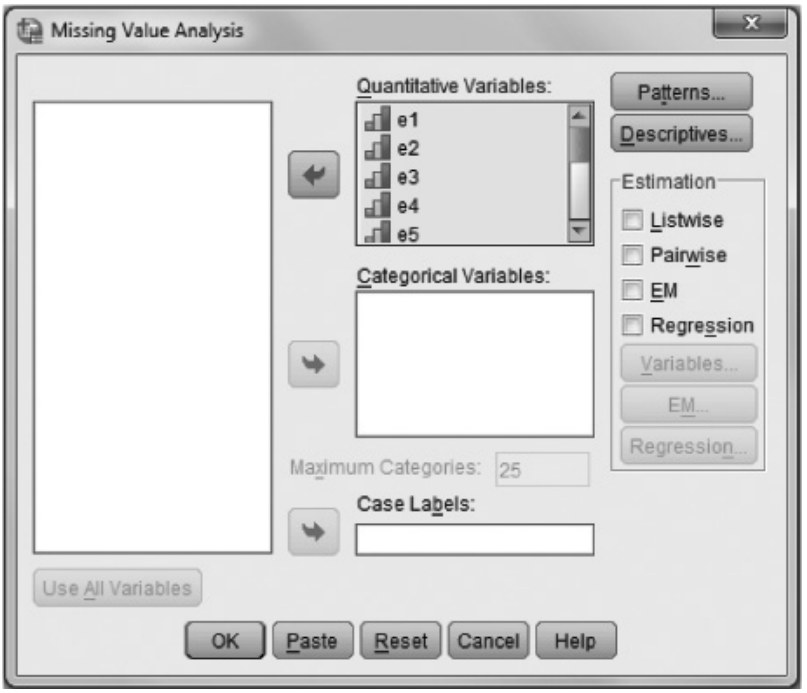
2. Transfer the variables **e1** to **e12** to the **Quantitative Variables:** field by clicking these variables (highlight) and then clicking ▸.



3. In the **Estimation** field, check the **EM** cell and then click EM... This will open the **Missing Values Analysis: EM** window. In the **Distribution** field, check **Normal**. Check the **Save completed data** cell. To save the imputed data set, say in the computer's J:\directory, click File... to look for the J:\directory and then save the imputed data set under the name **COMPLETE.SAV.** Click Continue to return to the **Missing Value Analysis** window.



4. When the **Missing Value Analysis** window opens, click OK to complete the analysis.



2.4.3.2 SPSS Syntax Method

```
MVA VARIABLES = E1 TO E12
/EM (OUTFILE = 'J:\COMPLETE.SAV').
```

2.4.3.3 Imputed Data Set

Retrieve the imputed data set **COMPLETE.SAV** from J:\directory. Run a simple Frequencies analysis on the variables (**e1** to **e12**) and the output will show that all missing values from the original data set (**MISSING.SAV**) have been replaced with scores imputed via the EM method. That is, the imputed data set contains no missing values.

Frequencies

Statistics

		e1	e2	e3	e4	e5	e6	e7	e8	e9	e10	e11	e12
N	Valid	1928	1928	1928	1928	1928	1928	1928	1928	1928	1928	1928	1928
	Missing	0	0	0	0	0	0	0	0	0	0	0	0

3

Multiple Response

3.1 Aim

MULT RESPONSE analysis allows the researcher to analyze research questions that can have multiple responses. For example, a research question may ask respondents to *name* all the newspapers read within last week, or to *circle* all newspapers read within last week from a list of newspapers. One way to generate descriptive statistics for each of the newspapers selected is to do a simple FREQUENCIES analysis. However, an ordinary FREQUENCIES analysis will only generate descriptive statistics for each nominated newspaper *separately* (e.g., the number and percentage of respondents who chose newspaper A). This procedure will not generate statistics on the basis of the entire "group" of newspapers nominated. For example, if the researcher is interested in the number of respondents who chose newspaper A as a *percentage of the total number of newspapers read*, then the MULT RESPONSE procedure should be used.

3.2 Methods of MULT RESPONSE Procedures

There are two ways to perform a frequency run with multiple response data. Whichever way the researcher chooses, the procedure will involve combining variables into groups. One way to organize multiple-response data is to create, for each possible response, a variable that can have one of two values, such as 1 for *yes* and 2 for *no*; this is the **multiple-dichotomy** method. Alternatively, on the basis of all responses collected from all respondents, the researcher can create variables to represent, for example, all the newspapers read. Each variable (newspaper) will have a value representing that newspaper, such a 1 for *Bangkok Post*, 2 for *The Nation*, and 3 for *Thai Rath*. This is the **multiple-response** method.

3.3 Example of the Multiple-Dichotomy Method

Suppose that in a survey of political party preference, the following question was asked.

"Why do you prefer that political party?" (you can choose more than one reason).

		1. Yes	2. No
1.	The party is honest.	____	____
2.	The party has integrity.	____	____
3.	The party is trustworthy.	____	____
4.	The party has always kept its promises.	____	____
5.	The party has strong leadership.	____	____

In this example, the respondent is asked to endorse the reasons for preferring a specific political party. Each reason is, therefore, a separate variable, with two possible values: 1 for *yes*, and 2 for *no*. The five reasons are named **HONEST**, **INTEG**, **TRUST**, **PROMISE**, and **LEADER** in the data set below. To do a **multiple-dichotomy** frequency analysis on the reasons given, choose either the Windows method (Section 3.3.2) or the syntax method (Section 3.3.3).

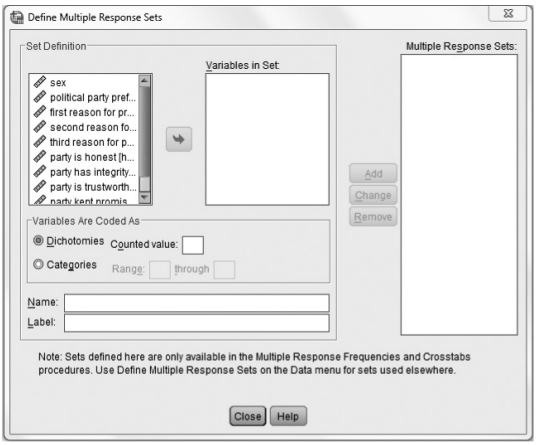
3.3.1 Data Entry Format

The data set has been saved under the name **EX3.SAV**

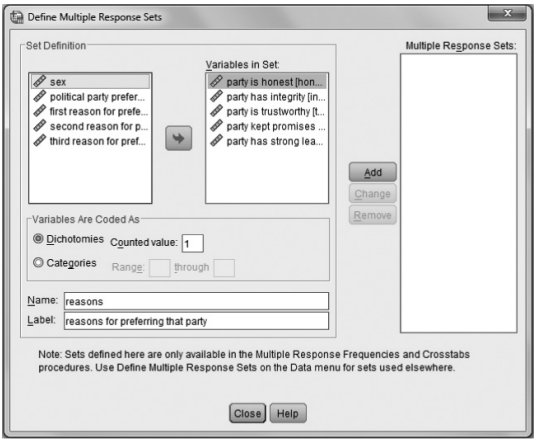
Variables	Column(s)	Code
Sex	1	1 = Male, 2 = Female
Party	2	1 = Labor, 2 = Liberal
		3 = National, 4 = Democrat
REASON 1	3	1 = The party is honest.
		2 = The party has integrity.
		3 = The party is trustworthy.
		4 = The party has always kept its promises.
		5 = The party has strong leadership.
REASON 2	4	As above
REASON 3	5	As above
HONEST	6	1 = yes, 2 = no
INTEG	7	1 = yes, 2 = no
TRUST	8	1 = yes, 2 = no
PROMISE	9	1 = yes, 2 = no
LEADER	10	1 = yes, 2 = no

3.3.2 Windows Method

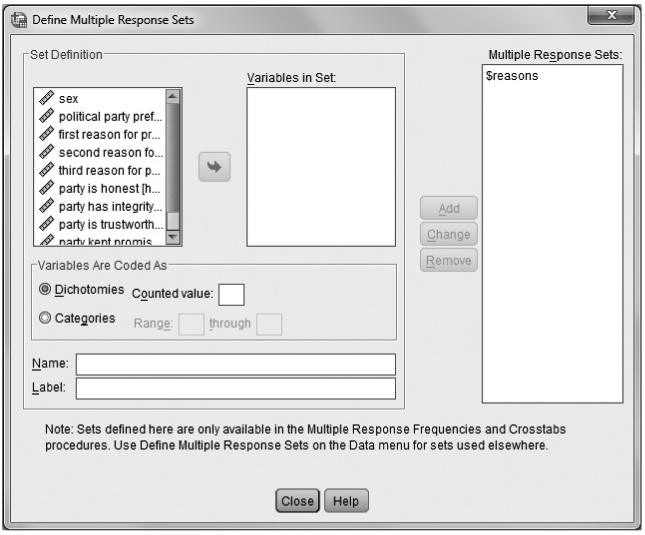
1. From the menu bar, click **Analyze**, then **Multiple Response**, and then **Define Variable Sets**. The following **Define Multiple Response Sets** window will open.



2. In the **Set Definition** field containing the study's variables, click (highlight) the variables (**HONEST**, **INTEG**, **TRUST**, **PROMISE**, **LEADER**) that will be grouped in the multiple response set. Click ▶ to transfer the selected variables to the **Variables in Set:** field. Since only those variables (reasons) that have been coded 1 (for *yes*) will be grouped for analysis, check the **Dichotomies** button, and in the **Counted value:** cell type **1**. Next, in the **Name:** field, type in a name for this multiple response set (e.g., *reasons*), and in the **Label:** field, type in a label (e.g., "*reasons for preferring that party*").

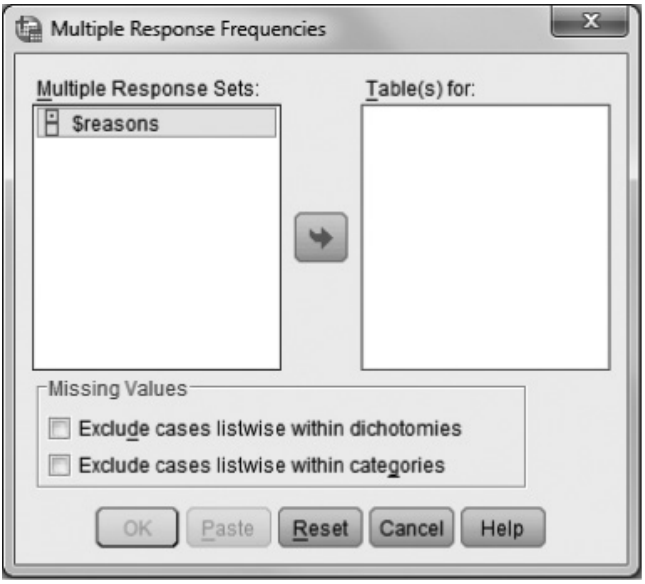


3. Click Add to transfer this response set to the **Multiple Response Sets:** field. The grouped response set is given the name *$reasons*.

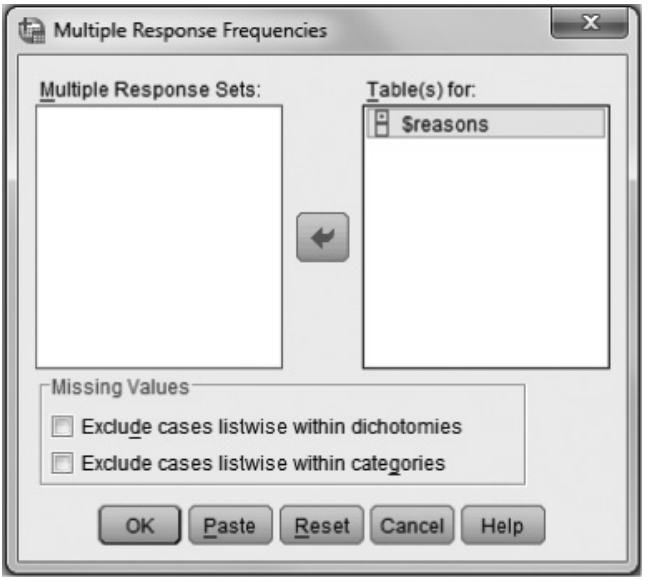


Click Close to close this window.

4. From the menu bar, click **Analyze**, then **Multiple Response**, and then **Frequencies.** The following **Multiple Response Frequencies** window will open.



5. Transfer the grouped response set *$reasons* (*reasons for preferring that party*) in the **Multiple Response Sets:** field to the **Table(s) for:** field by clicking (highlight) the response set, and then clicking ➔.



6. Click OK to run the multiple response frequencies analysis for the variables (HONEST, INTEG, TRUST, PROMISE, LEADER) in the grouped response set. See Table 3.1 for the results.

3.3.3 SPSS Syntax Method

Multiple-Dichotomy Frequency Analysis

```
MULT RESPONSE GROUPS = REASONS 'REASONS FOR PREFERRING
THAT PARTY' (HONEST TO LEADER(1))
/FREQUENCIES = REASONS.
```

3.3.4 SPSS Output

3.3.5 Results and Interpretation

In Table 3.1, the **N** column (beneath the **Responses** heading) presents the number of respondents who answered *yes* to each of the five reasons. Thus, of the 20 respondents included in the analysis, 13 endorsed *"party is honest"* as a reason for preferring that political party, 8 endorsed *"party has integrity"* as a reason, 10 endorsed *"party is trustworthy"* as a reason, 12 endorsed *"party kept promises"* as a reason, and 13 endorsed *"party has strong leader"* as a reason. Thus, a total of 56 *yes* responses were generated from the sample of 20 respondents.

TABLE 3.1

Multiple Response (Multiple-Dichotomy) Output

Multiple Response

Case Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
$reasons[a]	20	100.0%	0	.0%	20	100.0%

[a] Dichotomy group tabulated at value 1.

$reasons Frequencies

		Responses		Percent of Cases
		N	Percent	
Reasons for Preferring That Party[a]	party is honest	13	23.2%	65.0%
	party has integrity	8	14.3%	40.0%
	party is trustworthy	10	17.9%	50.0%
	party kept promises	12	21.4%	60.0%
	party has strong leader	13	23.2%	65.0%
Total		56	100.0%	280.0%

[a] Dichotomy group tabulated at value 1.

The **Percent** column presents the number of respondents who answered *yes* to each of the five reasons (in the **N** column) as a percentage of the total number of *yes* responses generated. For example, the 13 respondents who endorsed *"party is honest"* as a reason for preferring that political party represent 23.2% of the total number of *yes* responses (56) generated.

The **Percent of Cases** column presents the number of respondents who answered *yes* to each of the five reasons (in the **N** column) as a percentage of the total valid sample. For example, the 13 respondents who endorsed *"party is honest"* as a reason represent 65% of the total valid sample ($N = 20$ cases).

3.4 Example of the Multiple-Response Method

Using the same example as above, the following question was asked.

Why do you prefer that political party?

With the multiple-response method, the grouping of the multiple responses is different from the multiple-dichotomy method. A predetermined list of

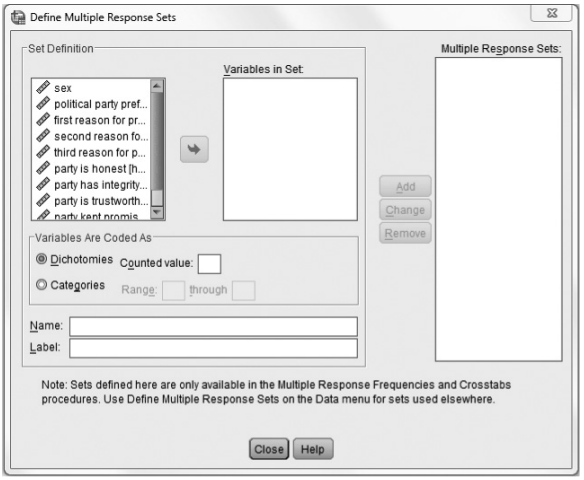
reasons will be used by the researcher in the multiple-response method to match the reasons chosen by the respondents. For example, a researcher may have the following list of reasons, each of which has been assigned a numerical value:

1 = The party is honest.
2 = The party has integrity.
3 = The party is trustworthy.
4 = The party has always kept its promises.
5 = The party has strong leadership.

Suppose that for this particular survey, each respondent is allowed to nominate a maximum of three reasons in response to the above question. These responses are labeled **REASON1**, **REASON2**, and **REASON3** in the data set **EX3.SAV**

3.4.1 Windows Method

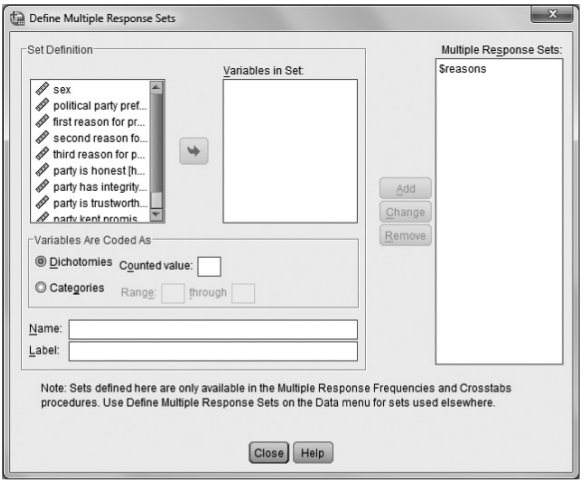
1. From the menu bar, click **Analyze**, then **Multiple Response**, and then **Define Variable Sets**. The following **Define Multiple Response Sets** window will open.



2. In the **Set Definition** field containing the study variables, click (highlight) the variables (**REASON1**, **REASON2**, **REASON3**) that will be grouped in the multiple response set. Click ▶ to transfer the selected variables to the **Variables in Set:** field. Since these three variables (reasons) have been coded from a predetermined list of

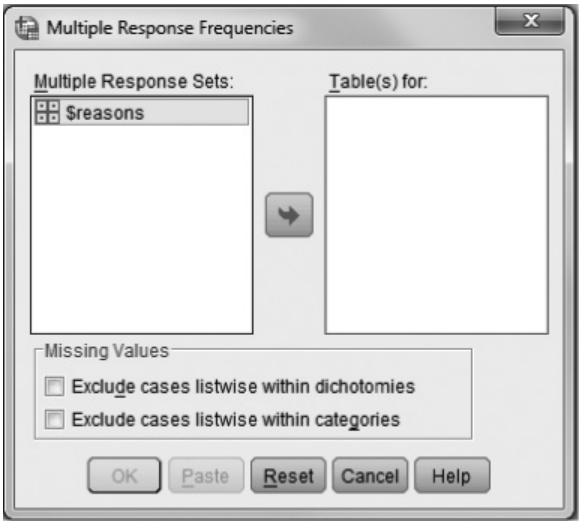
five reasons (see Section 3.4), check the **Categories** field, and in the **Range** fields type **1** through **5**. Next, in the **Name:** field, type in a name for this multiple response set (e.g., *reasons*), and in the **Label:** field, type in a label (e.g., *"reasons for preferring that party"*).

Click Add to transfer this response set to the **Multiple Response Sets:** field. The grouped response set is given the name *$reasons*.

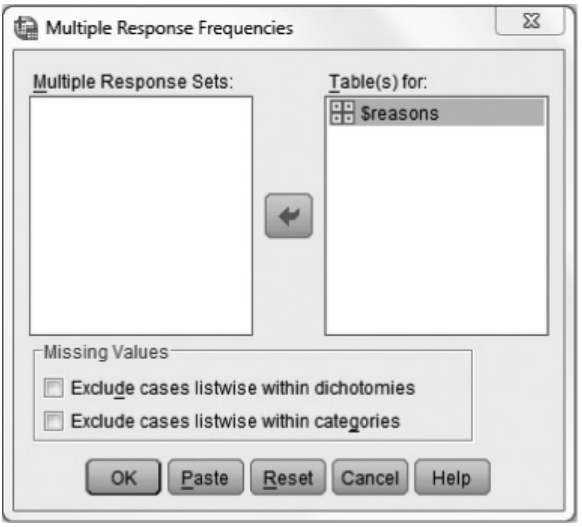


Click Close to close this window.

3. From the menu bar, click **Analyze**, then **Multiple Response**, and then **Frequencies.** The following **Multiple Response Frequencies** window will open.



4. Transfer the grouped response set *$reasons* (*reasons for preferring that party*) in the **Multiple Response Sets:** field to the **Table(s) for:** field by clicking the response set, and then clicking .



5. Click OK to run a multiple response frequencies analysis for the variables (REASON1, REASON2, REASON3) in the grouped response set. See Table 3.2 for the results.

3.4.2 SPSS Syntax Method

```
MULT RESPONSE GROUPS = REASONS 'REASONS FOR PREFERRING
THAT PARTY' (REASON1 TO REASON3 (1,5))
/FREQUENCIES = REASONS.
```

3.4.3 SPSS Output

3.4.4 Results and Interpretation

In Table 3.2, the **N** column (beneath the **Responses** heading) presents the number of respondents who chose each of the five reasons. Thus, of the 20 respondents included in the analysis, 11 chose *"party is honest"* as a reason for preferring that political party, 12 chose *"party has integrity"* as a reason, 12 chose *"party is trustworthy"* as a reason, 11 chose *"party kept promises"* as a reason, and 14 chose *"party has strong leader"* as a reason. Thus, a total of 60 responses were generated from the sample of 20 respondents.

TABLE 3.2

Multiple Response (Multiple-Response) Output

Multiple Response

Case Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
$reasons[a]	20	100.0%	0	.0%	20	100.0%

[a] Group.

$reasons Frequencies

		Responses		Percent of Cases
		N	Percent	
Reasons for Preferring That Party[a]	party is honest	11	18.3%	55.0%
	party has integrity	12	20.0%	60.0%
	party is trustworthy	12	20.0%	60.0%
	party kept promises	11	18.3%	55.0%
	party has strong leadership	14	23.3%	70.0%
Total		60	100.0%	300.0%

[a] Group.

The **Percent** column presents the number of respondents who chose each of the five reasons (in the **N** column) as a percentage of the total number of responses generated. For example, the 11 respondents who chose *"party is honest"* as a reason represent 18.3% of the total number of responses (60) generated.

The **Percent of Cases** column presents the number of respondents who selected each of the five reasons (in the **N** column) as a percentage of the total valid sample. For example, the 11 respondents who chose *"party is honest"* as a reason represent 55% of the total valid sample ($N = 20$ cases)

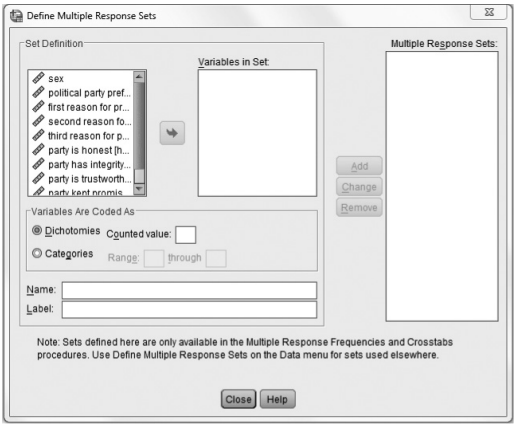
3.5 Cross-Tabulations

Cross-tabulations can be produced by MULT RESPONSE. Both individual and group variables can be tabulated together. Using the above example, suppose the researcher wants to cross-tabulate the *"reasons for preferring that party"* variable with the between-groups variable **SEX** (coded 1 = male,

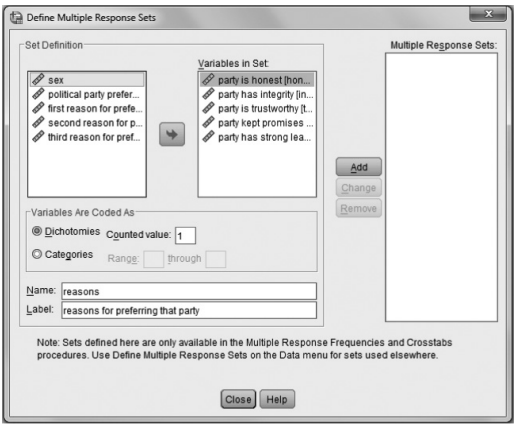
2 = female). The **multiple-dichotomy** method will be used to demonstrate this example. The same data set will be used for this example: **EX3.SAV**

3.5.1 Windows Method

1. From the menu bar, click **Analyze**, then **Multiple Response**, and then **Define Sets**. The following **Define Multiple Response Sets** window will open.

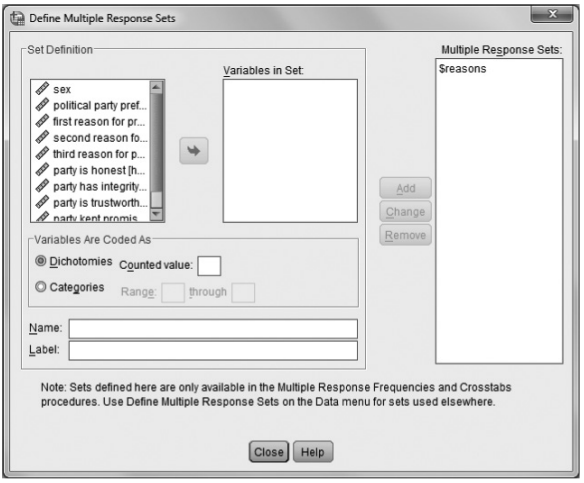


2. In the **Set Definition** field containing the study's variables, click (highlight) the variables (**HONEST**, **INTEG**, **TRUST**, **PROMISE**, **LEADER**) that will be grouped in the multiple response set. Click ▶ to transfer the selected variables to the **Variables in Set:** field. Since only those variables (reasons) that have been coded 1 (for *yes*) will be grouped for analysis, check the **Dichotomies** button, and in



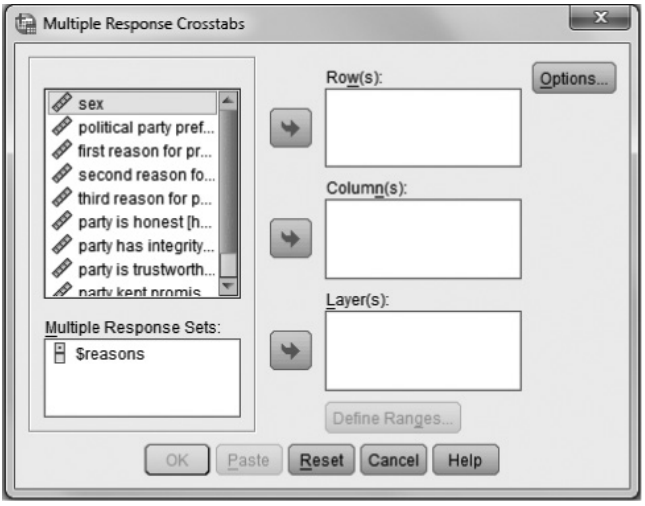
the **Counted value:** cell type **1**. Next, in the **Name:** field, type in a name for this multiple response set (e.g., *reasons*), and in the **Label:** field, type in a label (e.g., *"reasons for preferring that party"*).

Click Add to transfer this response set to the **Multiple Response Sets:** field. The grouped response set is given the name *$reasons*.

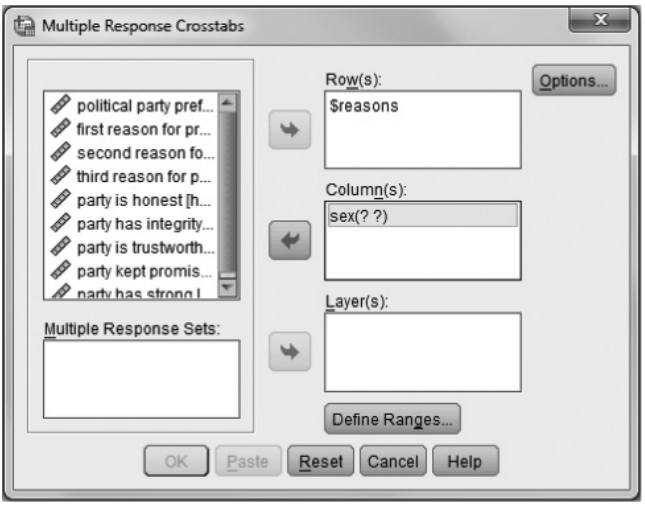


Click Close to close this window.

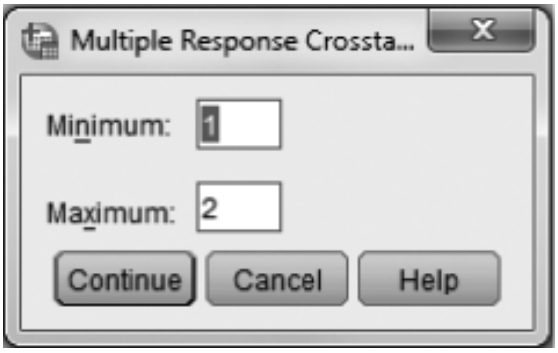
3. From the menu bar, click **Analyze**, then **Multiple Response**, and then **Crosstabs**. The following **Multiple Response Crosstabs** window will open.



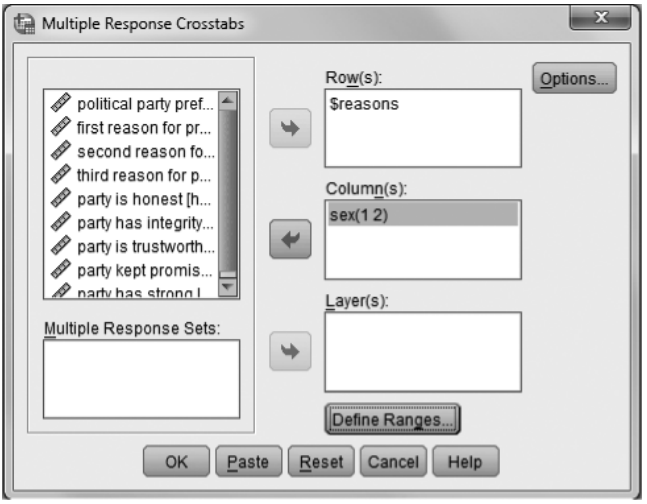
4. Transfer the grouped response set *$reasons* (*reasons for preferring that party*) in the **Multiple Response Sets:** field to the **Row(s):** field by clicking (highlight) the response set, and then clicking ▶. Next, transfer the between-groups variable **SEX** in the **Multiple Response Crosstabs** field to the **Column(s):** field by clicking (highlight) the **SEX** variable, and then clicking ▶.



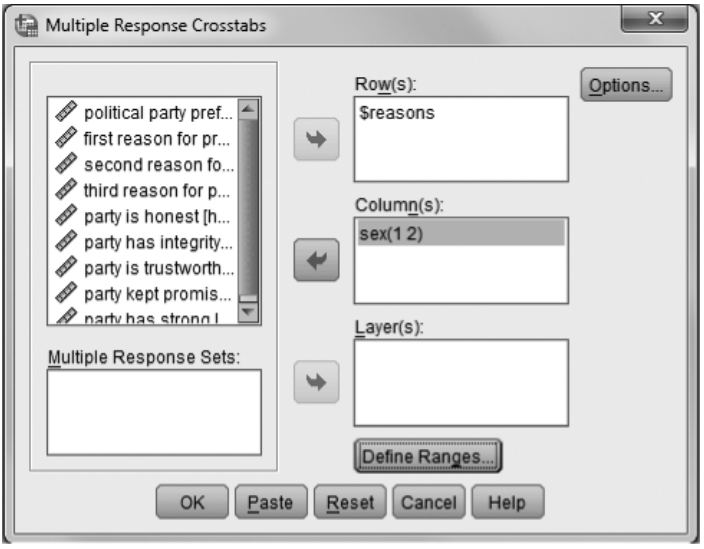
5. Click Define Ranges... to define the range for the between-groups variable **SEX** (coded 1 = male, 2 = female). When the following window opens, type the number **1** in the **Minimum:** field and the number **2** in the **Maximum:** field, and then click Continue.



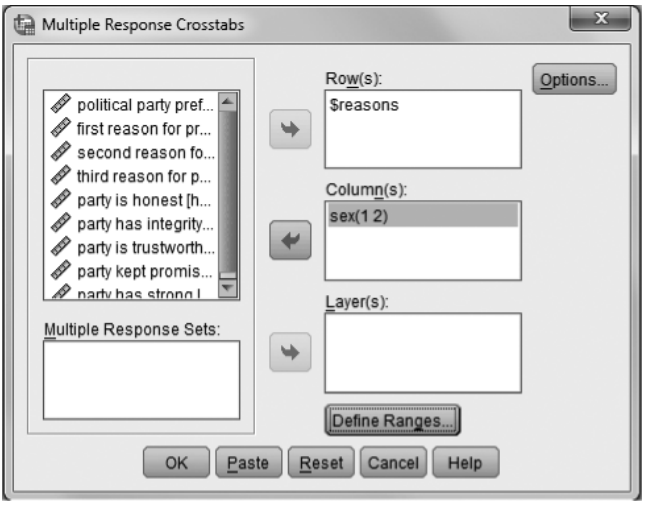
The following **Multiple Response Crosstabs** window will open.



6. Click Options... to obtain row, column, and total percentages in the cross-tabulation table.
7. When the following **Multiple Response Crosstabs: Options** window opens, check the **Row**, **Column**, and **Total** fields, and then click Continue.



8. When the **Multiple Response Crosstabs** window opens, click OK to run the cross-tabulation analysis. See Table 3.3 for the results.



3.5.2 SPSS Syntax Method

```
MULT RESPONSE GROUPS = REASONS 'REASONS FOR PREFERRING
THAT PARTY' (HONEST TO LEADER(1))
/VARIABLES SEX(1,2)
/TABLES = REASONS BY SEX
/CELLS = ALL.
```

This syntax file produces the **REASONS*SEX** cross-tabulation table. The **CELLS** syntax produces cell counts, row percentages, column percentages, and two-way table total percentages.

3.5.3 SPSS Output

3.5.4 Results and Interpretation

In Table 3.3, **Count** presents the *frequencies* for all the cells. Thus, seven males and six females chose *"party is honest"* as a reason for preferring that political party, three males and five females chose *"party has integrity"* as a reason, six males and four females chose *"party is trustworthy"* as a reason, seven males and five females chose *"party kept promises"* as a reason, and seven males and six females chose *"party has strong leader"* as a reason.

% within $reasons presents the number of male and female respondents who chose each of the five reasons (in the **Count** column) as a *percentage* of the number of respondents in each *reason* category. For example, a total of 13 respondents (7 males, 6 females) chose *"party is honest"* as a reason for preferring that political party. Thus, the seven males and six females represent 53.8% and 46.2% of respondents who chose that reason, respectively. Similarly, a total of eight respondents (three males, five females) chose

TABLE 3.3

Multiple Response Cross-Tabulation Output

Reasons by Sex

Multiple Response

Case Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
$reasons*sex	20	100.0%	0	.0%	20	100.0%

$reasons*sex Cross-Tabulation

			Sex		
			Male	Female	Total
Reasons for Preferring That Party[a]	party is honest	Count	7	6	13
		% within $reasons	53.8%	46.2%	
		% within sex	63.6%	66.7%	
		% of Total	35.0%	30.0%	65.0%
	party has integrity	Count	3	5	8
		% within $reasons	37.5%	62.5%	
		% within sex	27.3%	55.6%	
		% of Total	15.0%	25.0%	40.0%
	party is trustworthy	Count	6	4	10
		% within $reasons	60.0%	40.0%	
		% within sex	54.5%	44.4%	
		% of Total	30.0%	20.0%	50.0%
	party kept promises	Count	7	5	12
		% within $reasons	58.3%	41.7%	
		% within sex	63.6%	55.6%	
		% of Total	35.0%	25.0%	60.0%
		Count	7	6	13
		% within $reasons	53.8%	46.2%	
		% within sex	63.6%	66.7%	
		% of Total	35.0%	30.0%	65.0%
Total		Count	11	9	20
		% of Total	55.0%	45.0%	100.0%

Percentages and totals are based on respondents.

[a] Dichotomy group tabulated at value 1.

"party has integrity" as a reason for preferring that political party. Thus, the three males and five females represent 37.5% and 62.5% of respondents who selected that reason, respectively.

% within sex presents the number of male and female respondents who chose each of the five reasons (in the **Count** column) as a *percentage* of the total number of male and female respondents, respectively. For example, the total sample consists of 20 respondents (11 males, 9 females). Of the 11 male respondents, seven chose *"party is honest"* as a reason for preferring that political party. Similarly, of the nine female respondents, six chose *"party is honest"* as a reason for preferring that political party. Thus, the seven males and six females represent 63.6% and 66.7% of the total male and female respondents, respectively, who selected that reason. Similarly, of the 11 male respondents, three chose *"party has integrity"* as a reason for preferring that political party. Of the nine female respondents, five chose *"party has integrity"* as a reason for preferring that political party. Thus, the three males and five females represent 27.3% and 55.6% of the total male and female respondents, respectively, who selected that reason.

% of Total presents the two-way table total percentages. Thus, for the total sample of 20 respondents, the seven males who chose *"party is honest"* represent 35% of the total sample, while the six females who chose the same reason represent 30% of the total sample. Similarly, the three males who chose *"party has integrity"* represent 15% of the total sample, while the five females who chose the same reason represent 25% of the total sample.

4

t *Test for Independent Groups*

4.1 Aim

The independent *t* test is used for testing the difference between the means of two independent groups. It is particularly useful when the research question requires the comparison of variables (measured at least at the *ordinal* level) obtained from two independent samples.

For example:

"Do males and females differ in performance on a standardized achievement test?"

"What is the effect of drug versus no drug on rats' maze learning behavior?"

"Does the recidivism rate of juvenile offenders who are provided with father figures differ from those without father figures?"

4.2 Checklist of Requirements

In any one analysis, there must be:

- Only one independent (grouping) variable (IV) (e.g., subject's gender)
- Only two levels for that IV (e.g., male, female)
- Only one dependent variable (DV)

4.3 Assumptions

- **Independence**—The two groups are independent of one another.
- **Normality**—The dependent variable is normally distributed.
- **Homogeneity of variance**—That is, the distribution of the dependent variable for one of the groups being compared has the same variance as the distribution for the other group being compared.

4.4 Example

A researcher wants to investigate whether first-year male and female students at a university differ in their memory abilities. Ten male students and 10 female students were randomly selected from the first-year enrolment roll to serve as subjects. All 20 subjects were read 30 unrelated words and were then asked to recall as many of the words as possible. The numbers of words correctly recalled by each subject were recorded.

Males	Females
s1 16	s1 24
s2 14	s2 23
s3 18	s3 26
s4 25	s4 17
s5 17	s5 18
s6 14	s6 20
s7 19	s7 23
s8 21	s8 26
s9 16	s9 24
s10 17	s10 20

4.4.1 Data Entry Format

The data set has been saved under the name **EX4.SAV**

Variables	Column	Code
Gender	1	1 = male, 2 = female
Words	2	Number of words correctly recalled

4.4.2 Testing Assumptions

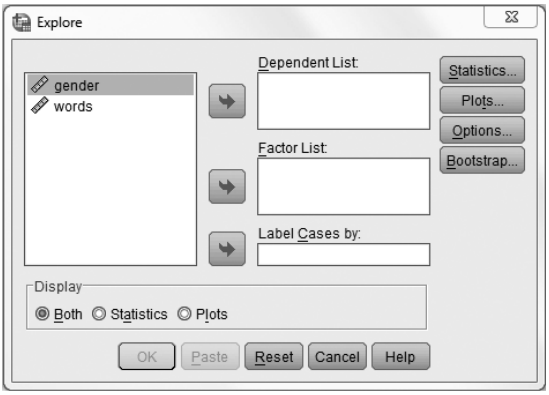
4.4.2.1 Independence

During data collection, ensure that the observations in one group are independent of the observations of the other group.

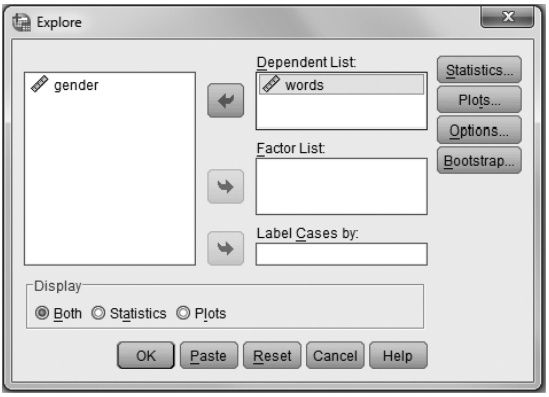
4.4.2.2 Normality

4.4.2.2.1 Windows Method

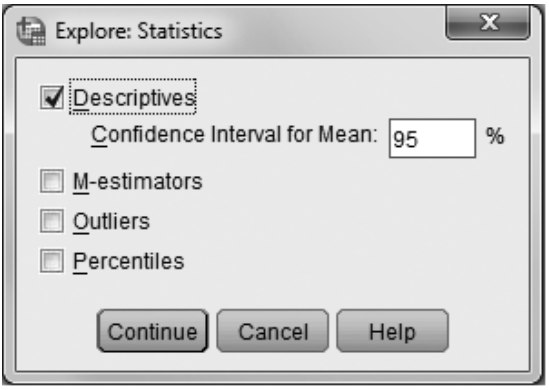
1. From the menu bar, click **Analyze**, then **Descriptive Statistics**, and then **Explore…** . The following **Explore** window will open.



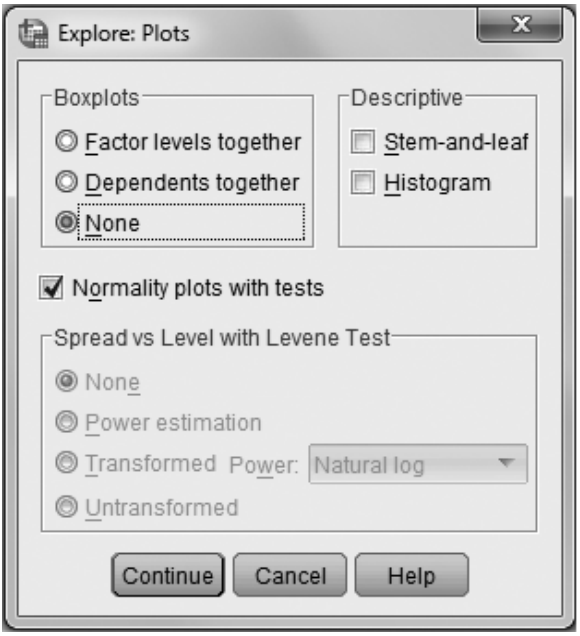
2. Transfer the **WORDS** variable to the **Dependent List:** field by clicking this variable (highlight) and then clicking .



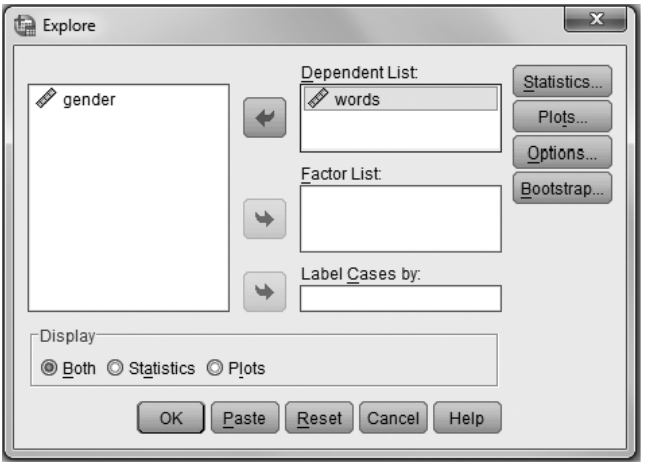
3. Click Statistics... to open the **Explore: Statistics** window. Check the **Descriptives** field and click Continue to return to the **Explore** window.



4. In the **Explore** window click Plots... to open the **Explore: Plots**



window. Check the **Normality plots with tests** field. Click Continue to return to the **Explore** window.



5. When the **Explore** window opens, click OK to complete the analysis. See Table 4.1 for the results.

TABLE 4.1

Explore Analysis (Selected) Output

Descriptives

			Statistic	Std. Error
Words	Mean		19.9000	0.87027
	95% Confidence Interval for Mean	Lower Bound	18.0785	
		Upper Bound	21.7215	
	5% Trimmed Mean		19.8889	
	Median		19.5000	
	Variance		15.147	
	Std. Deviation		3.89196	
	Minimum		14.00	
	Maximum		26.00	
	Range		12.00	
	Interquartile Range		6.75	
	Skewness		0.167	0.512
	Kurtosis		−1.234	0.992

Tests of Normality

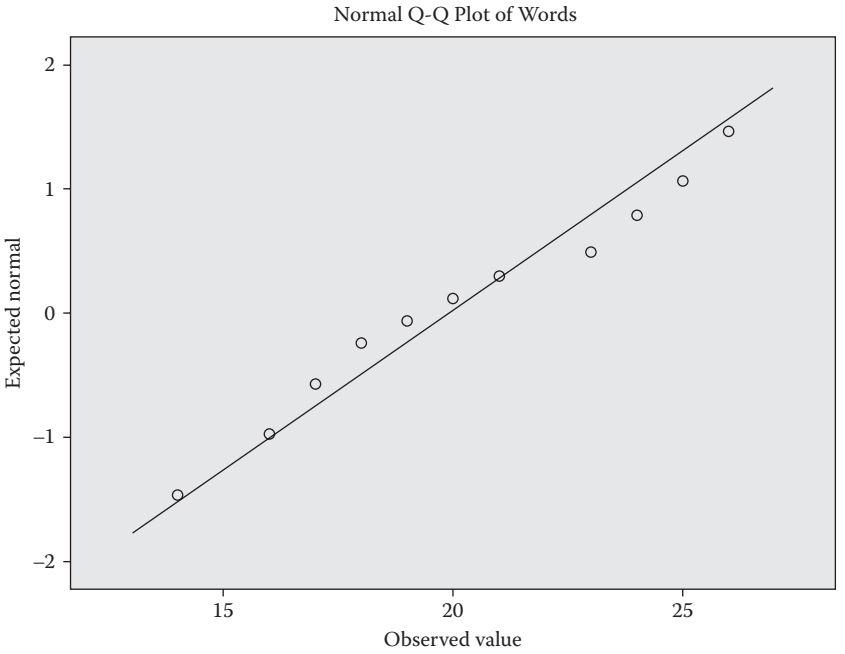
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	d*f*	Sig.	Statistic	d*f*	Sig.
Words	0.137	20	0.200*	0.936	20	0.201

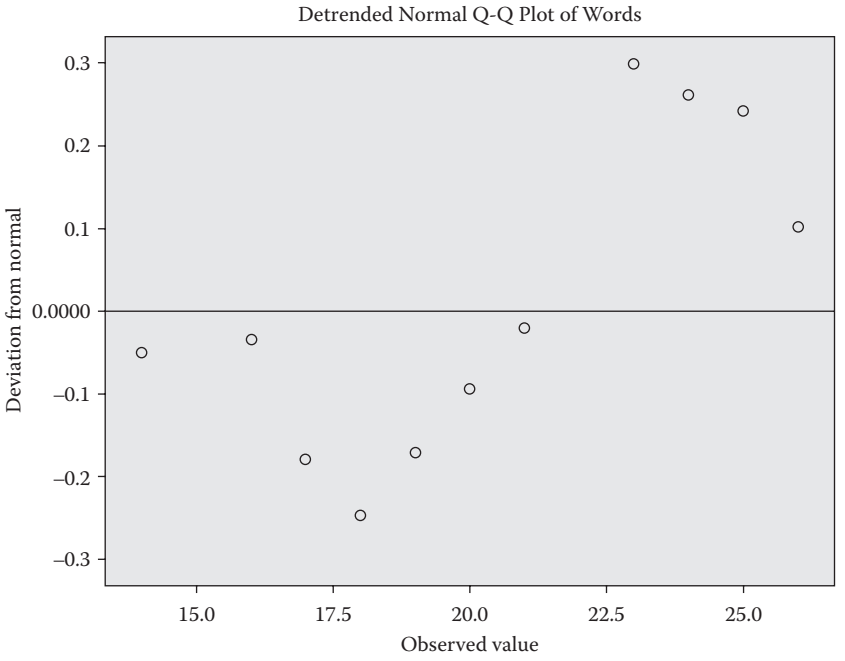
[a] Lilliefors significance correction.
* This is a lower bound of the true significance.

4.4.2.2.2 SPSS Syntax Method

```
EXAMINE VARIABLES = WORDS
/PLOT NPPLOT
/COMPARE GROUPS
/STATISTICS DESCRIPTIVES
/CINTERVAL 95
/MISSING LISTWISE
/NOTOTAL.
```

4.4.2.2.3 SPSS Output





4.4.2.2.4 Interpretation

The Kolmogorov-Smirnov statistic and the Shapiro-Wilk statistic are tests for normality, and if their significance levels are greater than 0.05, then normality is assumed. The Shapiro-Wilk statistic is calculated when the sample size is small (<50). For both the Kolmogorov-Smirnov and the Shapiro-Wilk tests, the computed significance levels are >0.05 (0.200 and 0.201, respectively). Therefore, normality can be assumed.

Another simple diagnostic test for normality is based on the skewness and kurtosis values. The statistical *z* value for the skewness value is calculated as:

$$Z_{\text{skewness}} = \frac{\text{skewness}}{\sqrt{\text{s.e. skewness}}}$$

The statistical *z* value for the kurtosis value is calculated as:

$$Z_{\text{kurtosis}} = \frac{\text{kurtosis}}{\sqrt{\text{s.e. kurtosis}}}$$

If the calculated *z* value exceeds the specified critical probability value, then the distribution is nonnormal. For example, a calculated *z* value exceeding ±2.58 will result in a rejection of the assumption of normality at the 0.01 critical probability (alpha) level. A calculated *z* value exceeding ±1.96 will result in a rejection of the assumption of normality at the 0.05 alpha level. Based on the obtained skewness statistics, the *z* value for the **WORDS** variable is $0.167/\sqrt{0.512} = \mathbf{0.23}$, which is less than ±1.96. Thus, it can be concluded that the distribution of this variable does not depart significantly from normality.

Another diagnostic test for normality is a visual check of the **Normal Q–Q Plot** that compares the cumulative distribution of the observed values with the expected values derived from the normal distribution. The normal distribution forms a straight diagonal line, and if a variable's distribution is normal, the data distribution will fall more or less on the diagonal. Inspection of the normal Q–Q plot shows very little departure from normality for the **WORDS** variable.

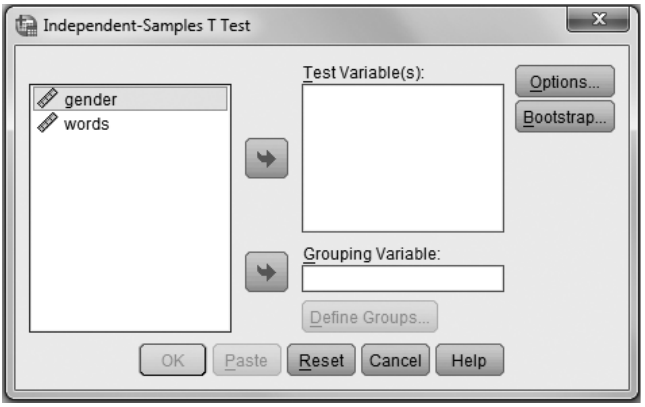
The **Detrended Normal Q–Q Plot** offers another visual check for normality. It shows the differences between the observed and expected values of a normal distribution, and plots the deviations of the scores from a straight line. If the distribution is normal, the scores should cluster around a horizontal line through zero with no pattern. The figure shows little deviation from normality.

4.4.2.3 Homogeneity of Variance

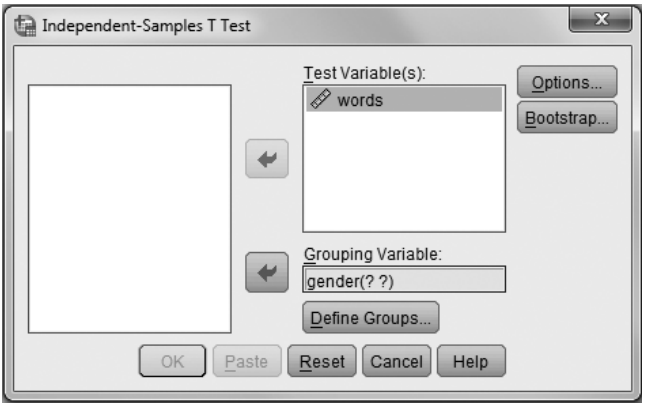
The homogeneity assumption is checked in SPSS by Levene's test.

4.4.3 Windows Method: Independent-Samples *t* Test

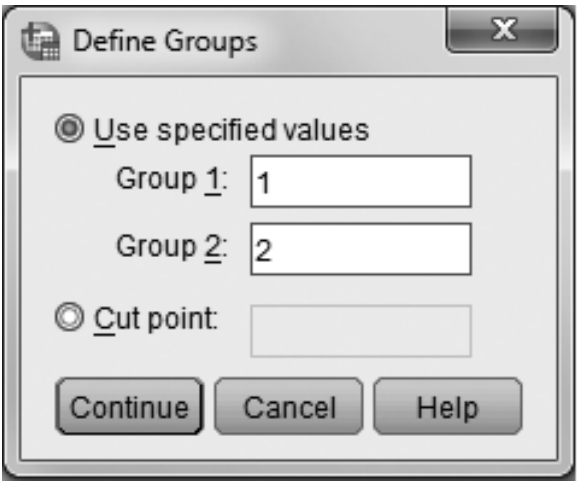
1. From the menu bar, click **Analyze**, then **Compare Means**, and then **Independent-Samples T Test**. The following window will open.



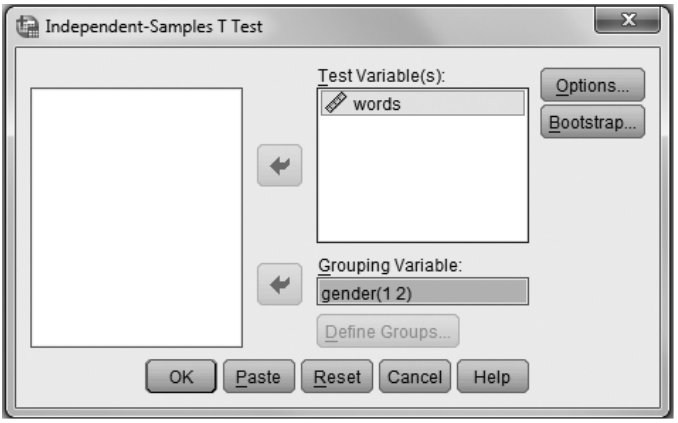
2. Since **GENDER** is the grouping (independent) variable, transfer it to the **Grouping Variable:** field by clicking (highlight) the variable and then clicking ▶. As **WORDS** is the test (dependent) variable, transfer it to the **Test Variable(s):** field by clicking (highlight) the variable and then clicking ▶.



3. Click Define Groups... to define the range for the grouping variable **GENDER** (coded 1 = male, 2 = female). When the following **Define Groups** window opens, type **1** in the **Group 1:** field and **2** in the **Group 2:** field, and then click Continue.



4. When the following **Independent-Samples** *t* **Test** window opens, run the *t* test analysis by clicking OK. See Table 4.2 for the results.



4.4.4 SPSS Syntax Method

```
t-TEST GROUPS = GENDER(1 2)
/MISSING = ANALYSIS
/VARIABLES = WORDS
/CRITERIA = CI(0.95).
```

4.4.5 SPSS Output

TABLE 4.2

Independent *t* Test Output

Group Statistics

	Gender	*N*	Mean	Std. Deviation	Std. Error Mean
WORDS	MALE	10	17.7000	3.3350	1.0546
	FEMALE	10	22.1000	3.1780	1.0050

Independent Samples Test

		Levene's Test for Equality of Variances		*t* test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	*t*	d*f*	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
WORDS	Equal variances assumed	.087	.772	−3.020	18	.007	−4.4000	1.4568	−7.4606	−1.3394
	Equal variances not assumed			−3.020	17.958	.007	−4.4000	1.4568	−7.4611	−1.3389

4.4.6 Results and Interpretation

The assumption of **homogeneity of variance** is tested by **Levene's test for equality of variances**, which tests the hypothesis that the two population variances are equal. In this example, the Levene statistic is $F = 0.087$ and the corresponding level of significance is large (i.e., $p > 0.05$) (see Table 4.2). Thus, the assumption of homogeneity of variance has not been violated, and the **equal variances assumed** t test statistic can be used for evaluating the null hypothesis of equality of means. If the significance level of the Levene statistic is small (i.e., $p < 0.05$), the assumption that the population variances are equal is rejected and the **equal variances not assumed** t test statistic should be used.

The results from the t test analysis indicate the following:

- There is a significant difference between the male and female samples in the number of words correctly recalled, $t(df = 18) = -3.02$, $p < 0.01$. The mean values indicate that females correctly recalled significantly more words ($M = 22.10$) than males ($M = 17.70$).
- The confidence interval information shows that the null hypothesis value (i.e., zero) does not fall within this interval (Lower = –7.4606, Upper = –1.3394). Therefore, the null hypothesis of equality of means can be rejected.
- Eta-square (η^2) is 0.3363. This number can be interpreted in exactly the same way as R^2 in correlation and regression. That is, approximately 33.63% of the variability in the number of words correctly recalled was explained by the gender manipulation. Eta-square is calculated by:

$$\eta^2 = \frac{t^2}{t^2 + df}$$

$$0.3363 = -3.02^2 \div -3.02^2 + 18$$

5

Paired-Samples t *Test*

5.1 Aim

The paired-samples *t* test is used in **repeated measures** or **correlated groups** design, in which each subject is tested twice on the same variable. A common experiment of this type involves the *before and after* design. The test can also be used for the **matched group** design in which pairs of subjects that are matched on one or more characteristics (e.g., IQ, grades, and so forth) serve in the two conditions. As the subjects in the groups are matched and not independently assigned, this design is also referred to as a **correlated groups** design.

5.2 Checklist of Requirements

- In any one analysis, there must be only two sets of data.
- The two sets of data must be obtained from (1) the same subjects, or (2) from two matched groups of subjects.

5.3 Assumption

- The sampling distribution of the means should be normal.

5.4 Example

A researcher designed an experiment to test the effect of drug X on eating behavior. The amount of food eaten by a group of rats in a one-week period, prior to ingesting drug X, was recorded. The rats were then given drug X, and the amount of food eaten in a one-week period was again recorded. The following amounts of food in grams were eaten during the "before" and "after" conditions.

	Food Eaten	
	Before Ingesting Drug X	**After Ingesting Drug X**
s1	100	60
s2	180	80
s3	160	110
s4	220	140
s5	140	100
s6	250	200
s7	170	100
s8	220	180
s9	120	140
s10	210	130

5.4.1 Data Entry Format

The data set has been saved under the name **EX5.SAV**

Variables	Column	Code
• BEFORE	• 1	• Food eaten in grams
• AFTER	• 2	• Food eaten in grams

5.4.2 Testing Assumption

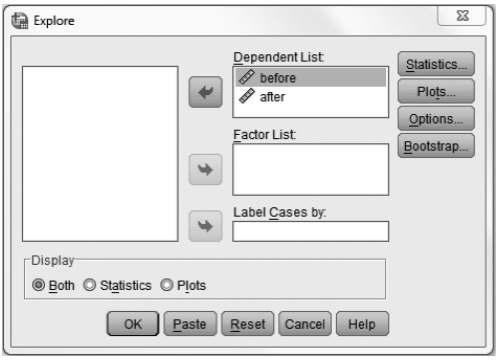
5.4.2.1 Normality

For the present example, normality will be tested using the normal Q–Q plot, the detrended normal Q–Q plot, and the *z* test for skewness.

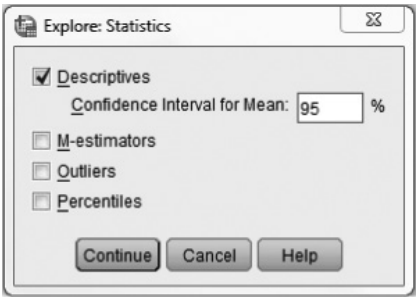
5.4.2.2 Windows Method

1. From the menu bar, click **Analyze**, then **Descriptive Statistics**, and then **Explore….** The following **Explore** window will open.

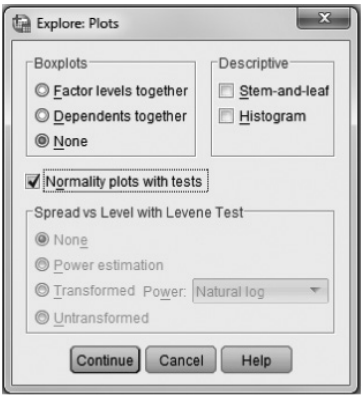
2. Transfer the **BEFORE** and **AFTER** variables to the **Dependent List:** field by clicking these variables (highlight) and then clicking .



3. Click Statistics... to open the **Explore: Statistics** window. Check the **Descriptives** field and click Continue to return to the **Explore** window.



4. In the **Explore** window click Plots... to open the **Explore: Plots** window. Check the **Normality plots with tests** field.



Click Plots... to return to the **Explore** window.

TABLE 5.1

Explore Analysis (Selected) Output

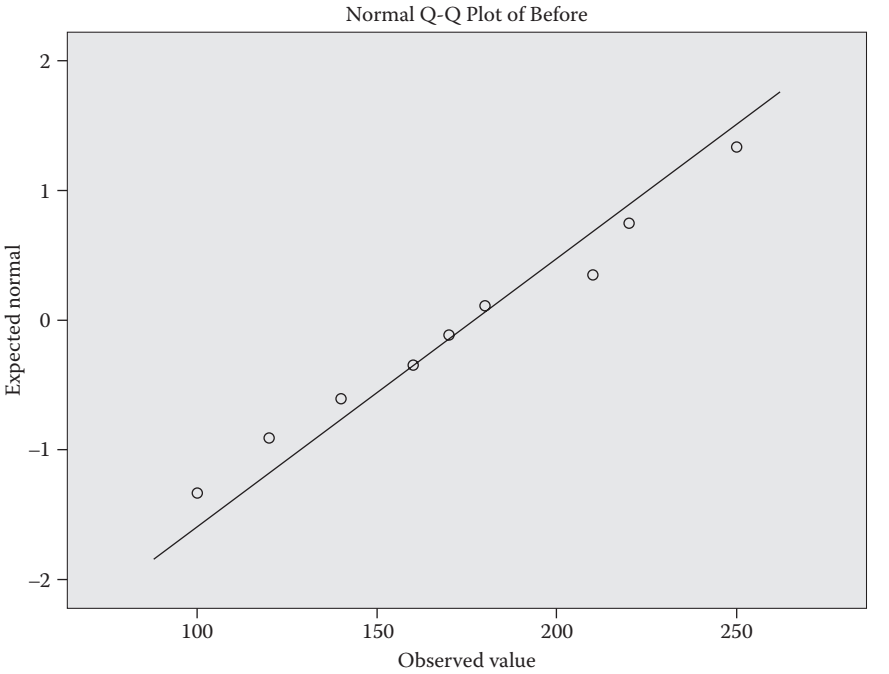
Descriptives				
			Statistic	**Std. Error**
Before	Mean		177.0000	15.27889
	95% confidence interval for mean	Lower bound	142.4368	
		Upper bound	211.5632	
	5% trimmed mean		177.2222	
	Median		175.0000	
	Variance		2334.444	
	Std. deviation		48.31609	
	Minimum		100.00	
	Maximum		250.00	
	Range		150.00	
	Interquartile range		85.00	
	Skewness		–0.140	0.687
	Kurtosis		–0.972	1.334
After	Mean		124.0000	13.67886
	95% confidence interval for mean	Lower bound	93.0563	
		Upper bound	154.9437	
	5% trimmed mean		123.3333	
	Median		120.0000	
	Variance		1871.111	
	Std. deviation		43.25634	
	Minimum		60.00	
	Maximum		200.00	
	Range		140.00	
	Interquartile range		55.00	
	Skewness		0.421	0.687
	Kurtosis		–0.315	1.334

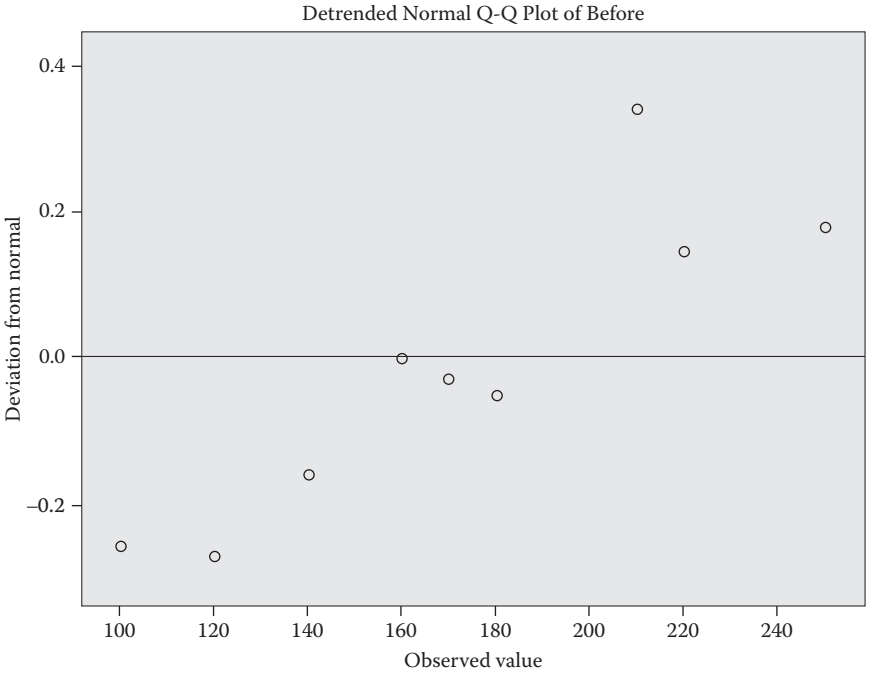
5. When the **Explore** window opens, click Continue to complete the analysis. See Table 5.1 for the results.

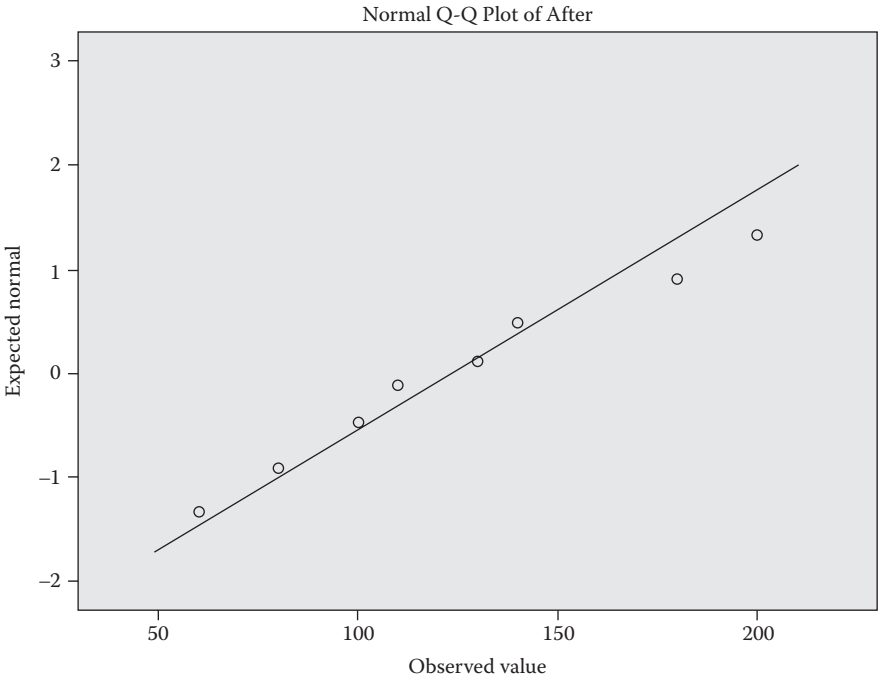
5.4.2.3 SPSS Syntax Method

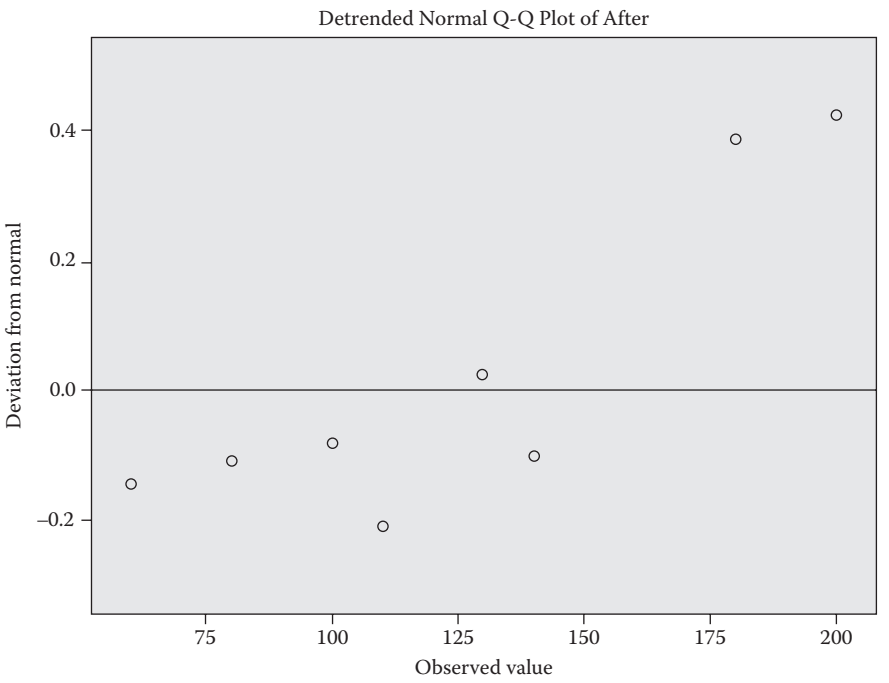
```
EXAMINE VARIABLES = BEFORE AFTER
/PLOT NPPLOT
/STATISTICS DESCRIPTIVES
/CINTERVAL 95
/MISSING LISTWISE
/NOTOTAL.
```

5.4.2.4 SPSS Output






Normal Q-Q Plot of After
Expected normal
3
2
1
0
–1
–2
50
100
150
200
Observed value


Detrended Normal Q-Q Plot of After
Deviation from normal
0.4
0.2
0.0
–0.2
75
100
125
150
175
200
Observed value

5.4.2.5 Interpretation

Similar to the independent *t* test, a simple diagnostic test for normality is based on the skewness and kurtosis values. The statistical *z* value for the skewness value is calculated as:

$$Z_{\text{skewness}} = \frac{\text{skewness}}{\sqrt{\text{s.e. skewness}}}$$

The statistical *z* value for the kurtosis value is calculated as:

$$Z_{\text{kurtosis}} = \frac{\text{kurtosis}}{\sqrt{\text{s.e. kurtosis}}}$$

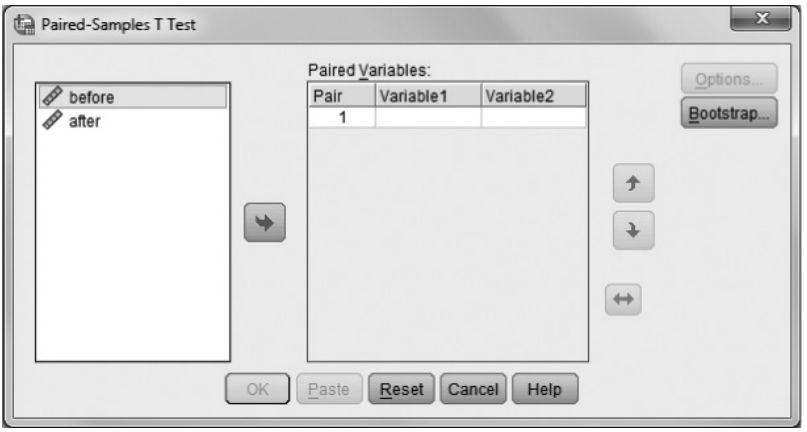
If the calculated *z* value exceeds the specified critical probability value, then the distribution is abnormal. For example, a calculated *z* value exceeding ±2.58 will result in a rejection of the assumption of normality at the 0.01 critical probability (alpha) level. A calculated *z* value exceeding ±1.96 will result in a rejection of the assumption of normality at the 0.05 alpha level. Based on the obtained skewness statistics, the *z* values for the **BEFORE** and **AFTER** variables are 0.17 and 0.51, respectively, and are less than ±1.96. Thus, it can be concluded that the distributions of these variables do not depart significantly from normality.

Another diagnostic test for normality is a visual check of the **Normal Q–Q Plot** that compares the cumulative distribution of the observed values with the expected values derived from the normal distribution. The normal distribution forms a straight diagonal line, and if a variable's distribution is normal, the data distribution will fall more or less on the diagonal. Inspection of the normal Q–Q plots shows very little departure from normality for both the **BEFORE** and **AFTER** variables.

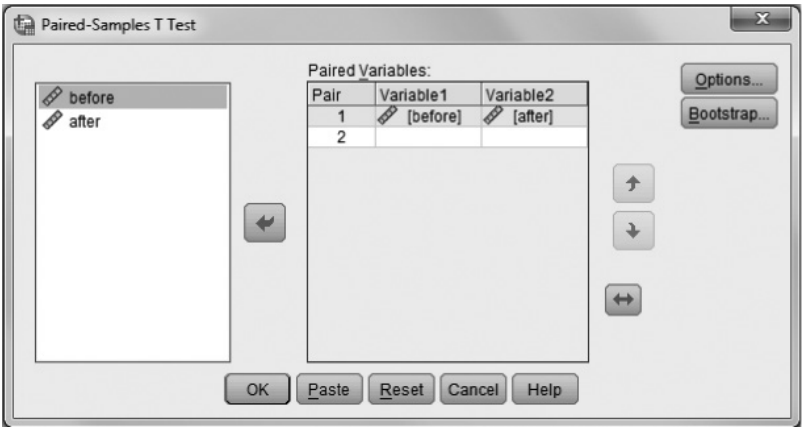
Similarly, a visual check of the **Detrended Normal Q–Q Plot**, which plots the deviations of the scores from a straight line, shows little deviation from normality for both the **BEFORE** and **AFTER** variables.

5.4.3 Windows Method: Paired-Samples *t* Test

1. From the menu bar, click **Analyze**, then **Compare Means**, and then **Paired-Samples *t* Test.** The following window will open.



2. Transfer both the **BEFORE** and **AFTER** variables to the **Paired Variables:** field by clicking (highlight) these two variables, and then clicking ▶. Click OK to run the *t* Test analysis. See Table 5.2 for the results.



5.4.4 SPSS Syntax Method

```
T-TEST PAIRS = BEFORE WITH AFTER (PAIRED)
/CRITERIA = CI(.9500)
/MISSING = ANALYSIS.
```

5.4.5 SPSS Output

TABLE 5.2

Paired-Samples *t* Test Output

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Before	177.0000	10	48.31609	15.27889
	After	124.0000	10	43.25634	13.67886

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Before and After	10	0.745	0.013

Paired Samples Test

		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	*t*	d*f*	Sig. (2-tailed)
Pair 1	before - after	53.00000	33.01515	10.44031	29.38239	76.61761	5.076	9	.001

5.4.6 Results and Interpretation

The result from the analysis indicates that there is a significant difference in the amount of food eaten before and after drug X was ingested, $t(df = 9) = 5.08$, $p < .01$ (see **Paired Samples Test** table). The mean values indicate that significantly less food was consumed after ingestion of drug X ($M = 124.00$) than before ($M = 177.00$).

6

One-Way Analysis of Variance, with Post Hoc Comparisons

6.1 Aim

The one-way analysis of variance (ANOVA) is an extension of the independent *t* test. It is used when the researcher is interested in whether the means from several (>2) independent groups differ. For example, if a researcher is interested in investigating whether four ethnic groups differ in their IQ scores, the one-way ANOVA can be used.

6.2 Checklist of Requirements

- In any analysis, there must be only one independent variable (e.g., ethnicity).
- There should be more than two levels for that independent variable (e.g., Australian, American, Chinese, African).
- There must be only one dependent variable.

6.3 Assumptions

- **Normality**—The dependent variable is normally distributed.
- **Homogeneity of variance**—The groups have approximately equal variance on the dependent variable.

6.4 Example

A researcher is interested in finding out whether intensity of electric shock will affect the time required to solve a set of difficult problems. Eighteen subjects are randomly assigned to the three experimental conditions of "Low

Shock," "Medium Shock," and "High Shock." The total time (in minutes) required to solve all the problems is the measure recorded for each subject.

Shock Intensity					
Low		**Medium**		**High**	
s1	15	s7	30	s13	40
s2	10	s8	15	s14	35
s3	25	s9	20	s15	50
s4	15	s10	25	s16	43
s5	20	s11	23	s17	45
s6	18	s12	20	s18	40

6.4.1 Data Entry Format

The data set has been saved under the name **EX6.SAV**

Variables	Column(s)	Code
• SHOCK	• 1	• 1 = low, 2 = medium, 3 = high
• TIME	• 2	• Time in minutes

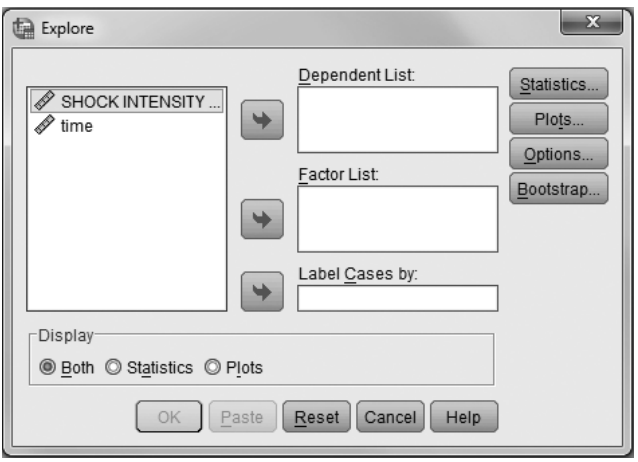
6.4.2 Testing Assumptions

6.4.2.1 Normality

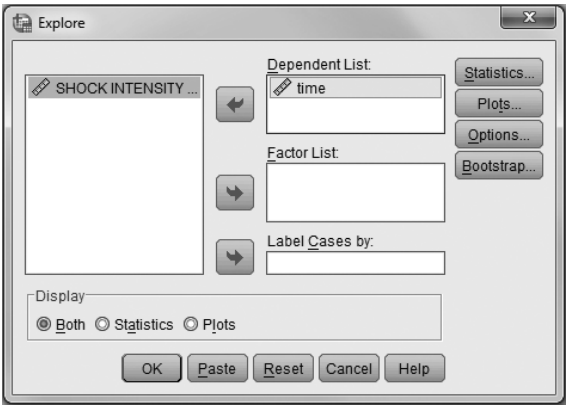
For the present example, normality will be tested using the normal Q–Q plot, the detrended Q–Q plot, and the z test for skewness.

6.4.2.1.1 Windows Method

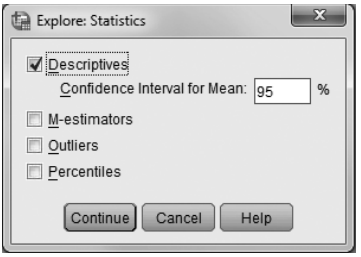
1. From the menu bar, click **Analyze**, then **Descriptive Statistics**, and then **Explore…** . The following **Explore** window will open.



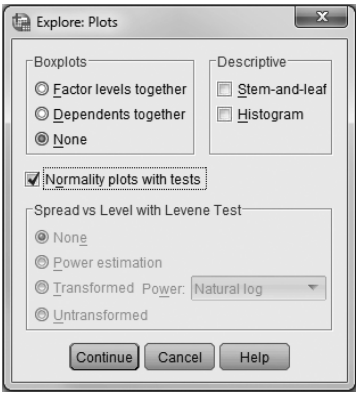
2. Transfer the **TIME** variable to the **Dependent List:** field by clicking this variable (highlight) and then clicking .



3. Click Statistics... to open the **Explore: Statistics** window. Check the **Descriptives** field and click Continue to return to the **Explore** window.



4. In the **Explore** window click Plots... to open the **Explore: Plots** window. Check the **Normality plots with tests** field.



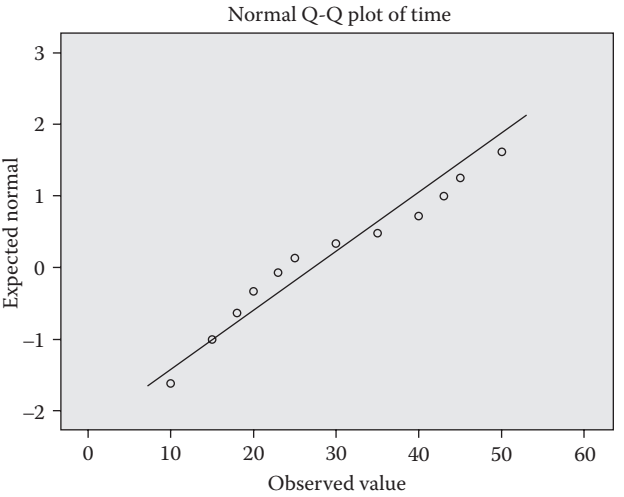
Click Continue to return to the **Explore** window.

5. When the **Explore** window opens, click OK to complete the analysis. See Table 6.1 for the results.

6.4.2.1.2 SPSS Syntax Method

```
EXAMINE VARIABLES = TIME
/PLOT NPPLOT
/STATISTICS DESCRIPTIVES
/CINTERVAL 95
/MISSING LISTWISE
/NOTOTAL.
```

6.4.2.1.3 SPSS Output



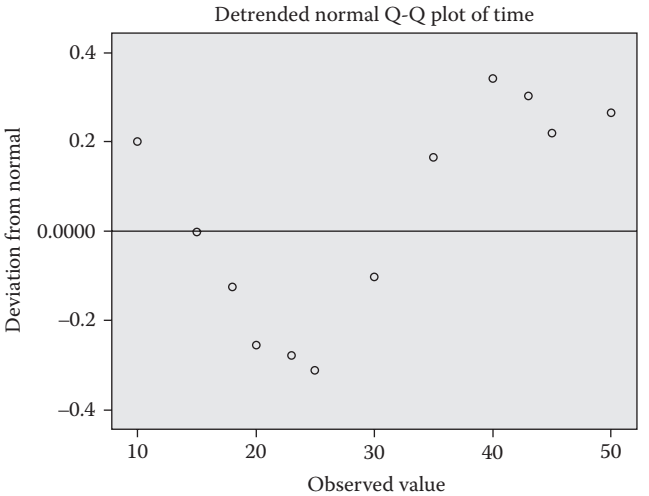


TABLE 6.1

Explore Analysis (Selected) Output

Descriptives

			Statistic	Std. Error
Time	Mean		27.1667	2.85402
	95% Confidence Interval for Mean	Lower Bound	21.1452	
		Upper Bound	33.1881	
	5% Trimmed Mean		26.8519	
	Median		24.0000	
	Variance		146.618	
	Std. Deviation		12.10858	
	Minimum		10.00	
	Maximum		50.00	
	Range		40.00	
	Interquartile Range		22.75	
	Skewness		.516	.536
	Kurtosis		−1.032	1.038

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	d*f*	Sig.	Statistic	d*f*	Sig.
Time	.182	18	.118	.921	18	.136

[a] Lilliefors significance correction.

6.4.2.1.4 Interpretation

For both the Kolmogorov-Smirnov and the Shapiro-Wilk tests, the computed significance levels are greater than 0.05 (0.118 and 0.136, respectively). Therefore, normality can be assumed.

A test for normality is based on the skewness and kurtosis values. The statistical *z* value for the skewness value is calculated as:

$$Z_{\text{skewness}} = \frac{\text{skewness}}{\sqrt{\text{s.e. skewness}}}$$

The statistical *z* value for the kurtosis value is calculated as:

$$Z_{\text{kurtosis}} = \frac{\text{kurtosis}}{\sqrt{\text{s.e. kurtosis}}}$$

A calculated *z* value exceeding ±1.96 will result in a rejection of the assumption of normality at the 0.05 alpha level. Based on the obtained skewness statistics, the *z* value for the **TIME** variable is 0.70, which is less than ±1.96. Thus, it can be concluded that the distribution of this variable does not depart significantly from normality.

Another diagnostic test for normality is a visual check of the **Normal Q–Q Plot**, which compares the cumulative distribution of the observed values with the expected values derived from the normal distribution. The normal distribution forms a straight diagonal line, and if a variable's distribution is normal, the data distribution will fall more or less on the diagonal. Inspection of the normal Q-Q plot shows very little departure from normality for the **TIME** variable.

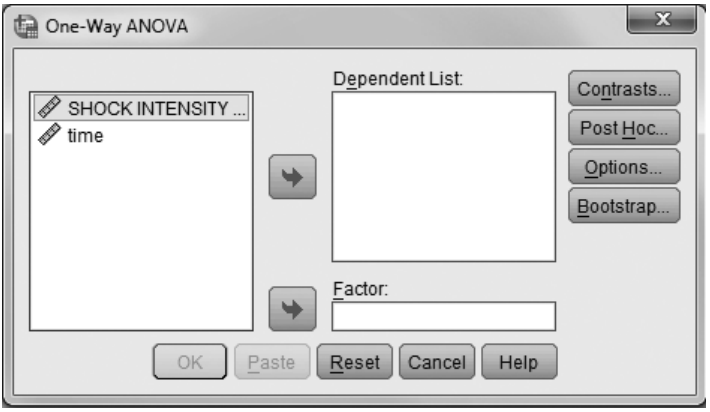
The **Detrended Normal Q-Q Plot** offers another visual check for normality. It shows the differences between the observed and expected values of a normal distribution, and plots the deviations of the scores from a straight line. If the distribution is normal, the scores should cluster around a horizontal line through zero with no pattern. The figure shows little deviation from normality.

6.4.2.2 Homogeneity of Variance

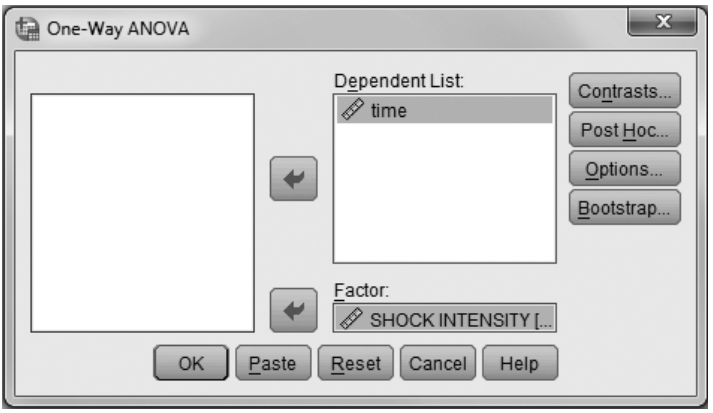
The homogeneity assumption is checked in SPSS by Levene's test.

6.4.3 Windows Method: One-Way ANOVA

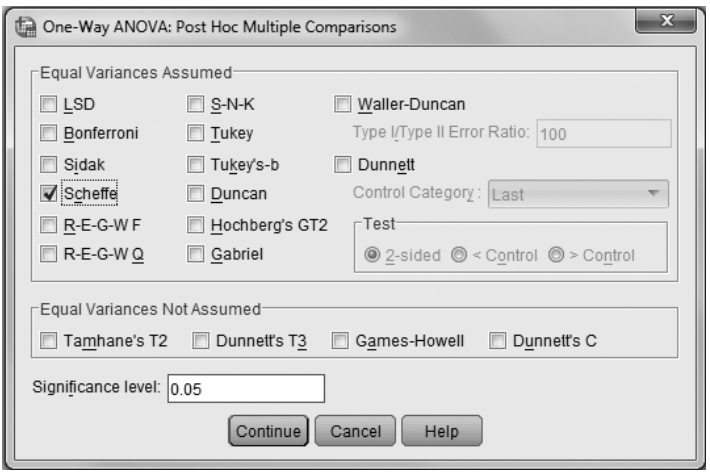
1. From the menu bar, click **Analyze**, then **Compare Means**, and then **One-Way ANOVA.** The following **One-Way ANOVA** window will open.



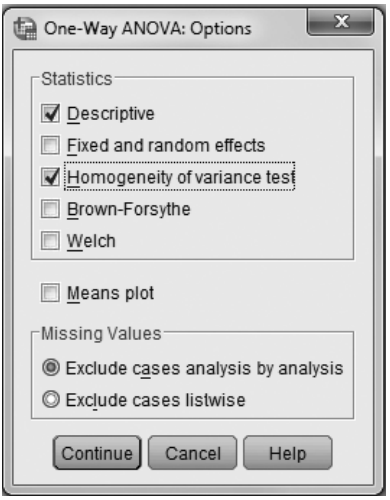
2. Transfer the dependent variable **TIME** to the **Dependent List:** field by clicking (highlight) the variable and then clicking ▶. Transfer the independent variable **SHOCK** to the **Factor:** field by clicking (highlight) the variable and then clicking ▶.



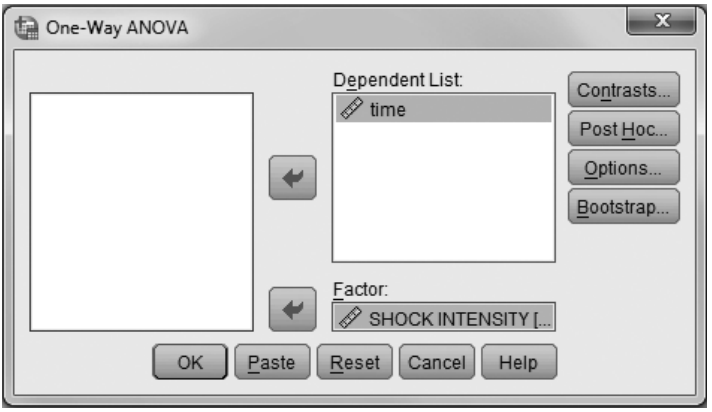
3. Since the one-way ANOVA will only perform an omnibus analysis of the *overall* differences between the three levels (low, medium, high) of the independent variable **SHOCK**, it will not analyze the differences between the *specific* shock levels. To obtain multiple comparisons between the three shock levels (low shock versus medium shock, low shock versus high shock, medium shock versus high shock), the researcher needs to perform a **post hoc** comparison test. Click Post Hoc... to achieve this. When the following **One-Way ANOVA: Post Hoc Multiple Comparisons** window opens, check the **Scheffe** field to run the Scheffé post hoc test. Next, click Continue.



4. When the **One-Way ANOVA** window opens, click Options... to open the **One-Way ANOVA: Options** window. Check the **Descriptive** box and the **Homogeneity of variance test** box and then click Continue.



5. When the following **One-Way ANOVA** window opens, run the analysis by clicking OK. See Table 6.2 for the results.



6.4.4 SPSS Syntax Method

```
ONEWAY TIME BY SHOCK
/STATISTICS DESCRIPTIVES HOMOGENEITY
/MISSING ANALYSIS
/POSTHOC = SCHEFFE ALPHA(0.05).
```

6.4.5 SPSS Output

TABLE 6.2

One-Way ANOVA Output

Descriptives

Time

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean: Lower Bound	95% Confidence Interval for Mean: Upper Bound	Minimum	Maximum
LOW SHOCK	6	17.1667	5.11534	2.08833	11.7985	22.5349	10.00	25.00
MEDIUM SHOCK	6	22.1667	5.11534	2.08833	16.7985	27.5349	15.00	30.00
HIGH SHOCK	6	42.1667	5.11534	2.08833	36.7985	47.5349	35.00	50.00
Total	18	27.1667	12.10858	2.85402	21.1452	33.1881	10.00	50.00

Test of Homogeneity of Variances

Time

Levene Statistic	d*f*1	d*f*2	Sig.
.000	2	15	1.000

Anova

Time

	Sum of Squares	d*f*	Mean Square	*F*	Sig.
Between Groups	2100.000	2	1050.000	40.127	.000
Within Groups	392.500	15	26.167		
Total	2492.500	17			

Post Hoc Tests

Multiple Comparisons

Dependent Variable: TIME

Scheffe

(I) SHOCK INTENSITY	(J) SHOCK INTENSITY	Mean Difference (I–J)	Std. Error	Sig.	95% Confidence Interval: Lower Bound	95% Confidence Interval: Upper Bound
LOW SHOCK	MEDIUM SHOCK	−5.0000	2.95334	.269	−13.0147	3.0147
	HIGH SHOCK	−25.0000*	2.95334	.000	−33.0147	−16.9853
MEDIUM SHOCK	LOW SHOCK	5.0000	2.95334	.269	−3.0147	13.0147
	HIGH SHOCK	−20.0000*	2.95334	.000	−28.0147	−11.9853
HIGH SHOCK	LOW SHOCK	25.0000*	2.95334	.000	16.9853	33.0147
	MEDIUM SHOCK	20.0000*	2.95334	.000	11.9853	28.0147

* The mean difference significant at the .05 level.

6.4.6 Results and Interpretation

The assumption of **homogeneity of variance** is tested by the **Levene statistic**, which tests the hypothesis that the population variances are equal. In this example, the Levene statistic is $F = 0.000$ and the corresponding level of significance is large (i.e., $p > .05$) (see Table 6.1). Thus, the assumption of homogeneity of variance has not been violated.

The results from the analysis (Table 6.1) indicate that the intensity of the electric shock has a significant effect on the time taken to solve the problems, $F(2,15) = 40.13$, $p < .001$. The mean values for the three shock levels indicate that, as the shock level increased (from low to medium to high), so did the time taken to solve the problems (low: $M = 17.17$; medium: $M = 22.17$; high: $M = 42.17$).

6.4.7 Post Hoc Comparisons

While the highly significant F-ratio ($p < .001$) indicates that the means of the three shock levels differ significantly, it does not indicate the *location* of this difference. For example, the researcher may want to know whether the overall difference is due primarily to the difference between "Low Shock" and "High Shock" levels, or between "Low Shock" and "Medium Shock" levels, or between "Medium Shock" and "High Shock" levels. To test for differences between specific shock levels, a number of post hoc comparison techniques can be used. For this example, the more conservative **Scheffé** test was used.

In the **Multiple Comparisons** table, in the column labeled **Mean Difference (I–J)**, the mean difference values accompanied by asterisks indicate which shock levels differ significantly from each other at the 0.05 level of significance. The results indicate that the high shock level is significantly different from both the low shock and medium shock levels. The low shock level and the medium shock level do not differ significantly. These results show that the overall difference in time taken to solve complex problems between the three shock-intensity levels is due to the significantly greater amount of time taken by the subjects in the high shock condition.

7

Factorial Analysis of Variance

7.1 Aim

The factorial univariate ANOVA is an extension of the one-way ANOVA in that it involves the analysis of two or more independent variables. It is used in experimental designs in which every level of every factor is paired with every level of every other factor. It allows the researcher to assess the effects of each independent variable separately, as well as the joint effect or interaction of variables. Factorial designs are labeled either by the number of factors involved, or in terms of the number of levels of each factor. Thus, a factorial design with two independent variables (e.g., gender, ethnicity) and with two levels for each independent variable (male/female; Australian/Chinese) are called either a **2-way factorial** or a **2 × 2 factorial**.

7.2 Checklist of Requirements

- In any one analysis, there must be two or more independent variables (due to the complexity in interpreting higher-order interactions, most factorial designs are limited to three or four independent variables or factors).
- There can be two or more levels for each independent variable.
- There must be only one dependent variable.

7.3 Assumptions

- **Independence**—The samples are independently drawn from the source population(s).
- **Normality**—The dependent variable is normally distributed.

- **Homogeneity of variance**—The distribution of the dependent variable for one of the groups being compared has the same variance as the distribution for the other group being compared.

7.4 Example 1: Two-Way Factorial (2 × 2 Factorial)

A researcher is interested in determining the effects of two learning strategies (A and B) on the memorization of a hard versus an easy list of syllables. The factorial combination of these two independent variables (2 × 2) yields four experimental conditions: Strategy A-Easy List, Strategy A-Hard List, Strategy B-Easy List, and Strategy B-Hard List. A total of 24 subjects is randomly assigned to the four experimental conditions. The researcher recorded the total number of mistakes made by each subject.

	Strategy A	Strategy B
Easy list	s1 6	s13 20
	s2 13	s14 18
	s3 11	s15 14
	s4 8	s16 14
	s5 9	s17 12
	s6 5	s18 16
Hard list	s7 15	s19 16
	s8 17	s20 13
	s9 23	s21 15
	s10 21	s22 20
	s11 22	s23 11
	s12 20	s24 12

7.4.1 Data Entry Format

The data set has been saved under the name **EX7a.SAV**

Variables	Column(s)	Code
• STRATEGY	1	1 = Strategy A, 2 = Strategy B
• LIST	2	1 = Easy, 2 = Hard
• ERRORS	3–4	Number of errors made

7.4.2 Testing Assumptions

7.4.2.1 Normality

For the present example, normality will be tested using the z test for skewness, normal Q-Q plot, and the detrended normal Q-Q plot.

7.4.2.1.1 Windows Method

With the data set opened, follow step 1 through step 5 in Section 6.4.2.1.1 in Chapter 6. Treat **ERRORS** as the dependent variable. See Table 7.1 for the results.

7.4.2.1.2 SPSS Syntax Method

```
EXAMINE VARIABLES = ERRORS
/PLOT NPPLOT
/STATISTICS DESCRIPTIVES
/CINTERVAL 95
/MISSING LISTWISE
/NOTOTAL.
```

7.4.2.1.3 SPSS Output

TABLE 7.1

Explore Analysis (Selected) Output

Descriptives

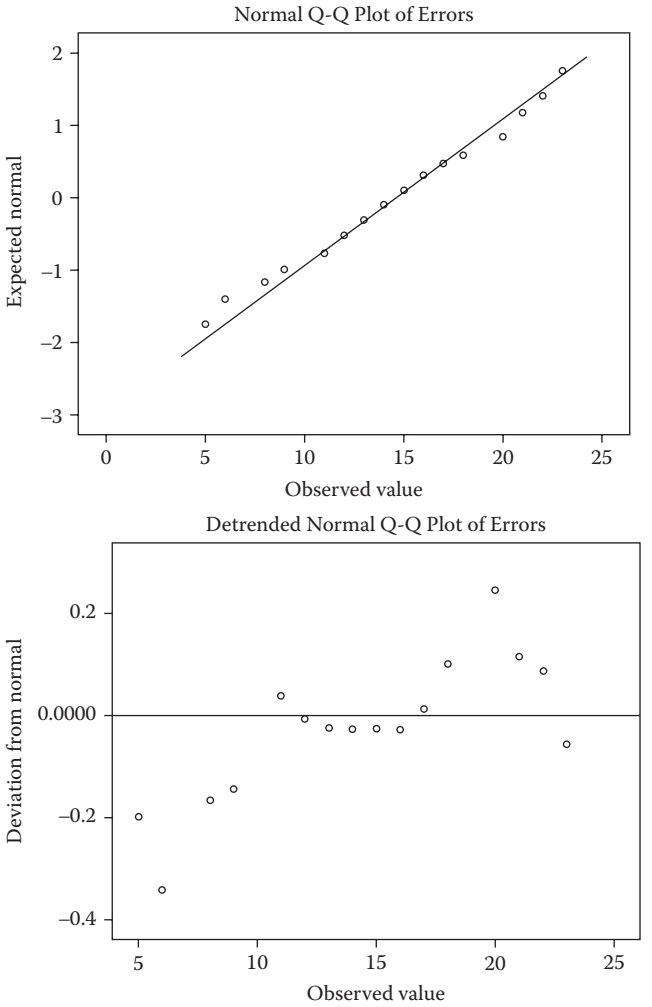
			Statistic	Standard Error
Errors	Mean		14.6250	1.00868
	95% Confidence Interval for mean	Lower Bound	12.5384	
		Upper Bound	16.7116	
	5% Trimmed mean		14.6944	
	Median		14.5000	
	Variance		24.418	
	Standard deviation		4.94151	
	Minimum		5.00	
	Maximum		23.00	
	Range		18.00	
	Interquartile range		8.25	
	Skewness		–.146	.472
	Kurtosis		–.622	.918

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	d*f*	Significance	Statistic	d*f*	Significance
Errors	.112	24	.200*	.975	24	.790

* This is a lower bound of the true significance.

[a] Lilliefors significance correction.



7.4.2.1.4 Interpretation

For both the Kolmogorov-Smirnov and the Shapiro-Wilk tests, the computed significance levels are greater than .05 (.200 and .790, respectively). Therefore, normality can be assumed.

Based on the obtained skewness statistics, the *Z*-skewness value for the **ERRORS** variable of **ERRORS** is .21 ($Z\text{skewness} = \text{skewness}/\sqrt{\text{s.e. skewness}}$) which is less than ±1.96. Therefore, it can be concluded that the distribution of this variable does not deviate significantly from normality.

Inspection of the normal Q-Q plot and the detrended normal Q-Q plot also shows very little departure from normality for the ERRORS variable.

7.4.2.2 Homogeneity of Variance

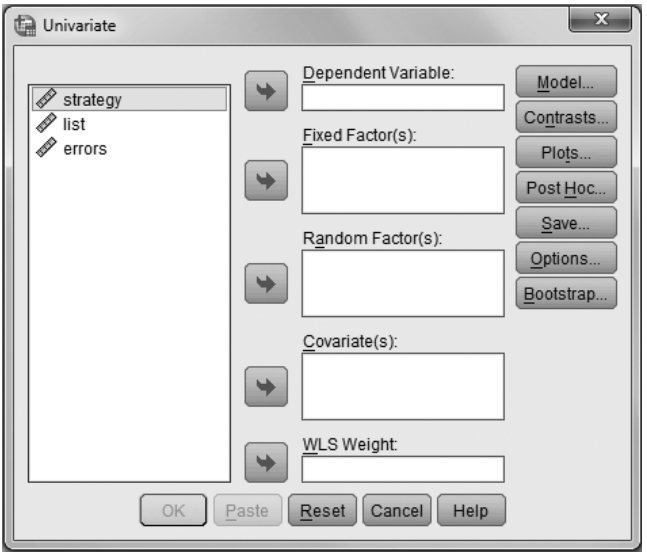
The homogeneity assumption is checked in SPSS by Levene's test.

7.4.2.3 Independence

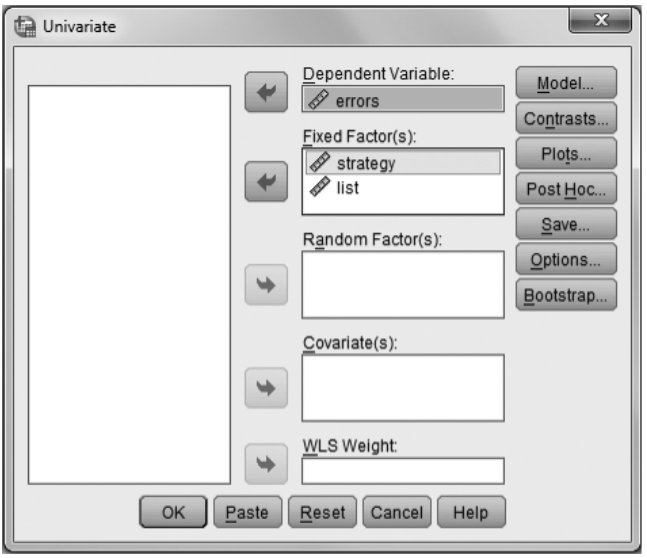
During data collection, ensure that the observations in one group are independent of the observations in the other group.

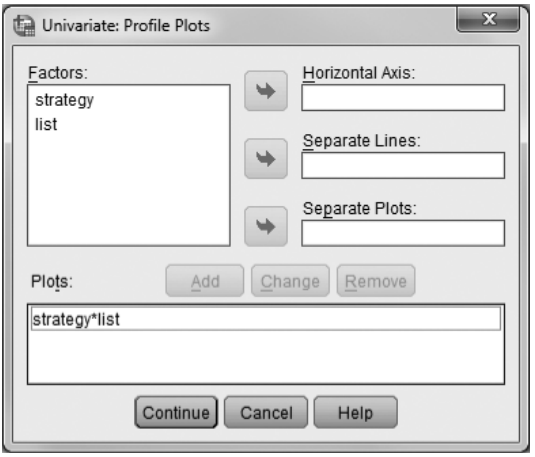
7.4.3 Windows Method: Factorial ANOVA

1. From the menu bar, click **Analyze**, then **General Linear Model**, and then **Univariate**. The following **Univariate** window will open.

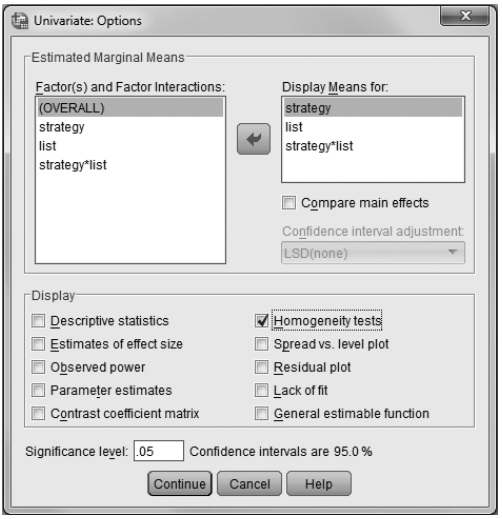


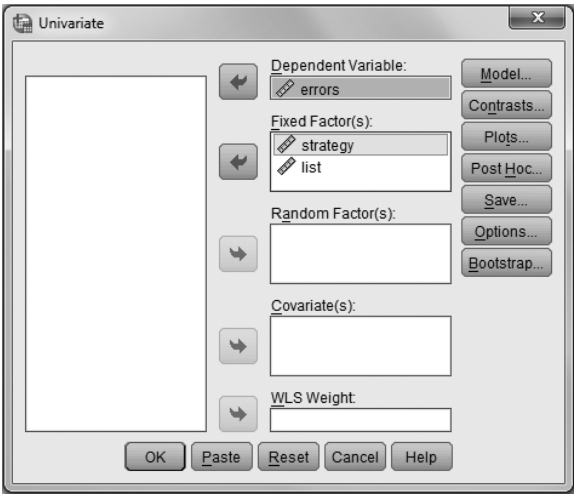
2. Transfer the **ERRORS** dependent variable of **ERRORS** to the **Dependent Variable:** field by clicking (highlight) the variable and then clicking ▶. Transfer the **STRATEGY** and **LIST** independent variables of **STRATEGY** and **LIST** to the **Fixed Factor(s):** field by clicking (highlight) the variables and then clicking ▶.
3. Click Plots... to plot a graph of the **STRATEGY*LIST** interaction. The following **Univariate: Profile Plots** window will open. Transfer the **STRATEGY** variable to the **Horizontal Axis:** field by clicking (highlight) the variable and then clicking ▶. Transfer the **LIST** variable to the **Separate Lines:** field by clicking (highlight) the variable and then clicking ▶. Next, click Add to transfer the **STRATEGY*LIST** interaction to the **Plots:** field. When this is done, click Continue.





4. Click Options... in the **Univariate** window to obtain descriptive statistics (estimated marginal means) for the full 2 x 2 **STRATEGY*LIST** interaction. When the **Univariate: Options** window opens, click (highlight) **STRATEGY**, **LIST**, and **STRATEGY*LIST** in the **Factor(s) and Factor Interactions:** field, and then click ▶ to transfer these factors and factor interaction to the **Display Means for:** field. Check the **Homogeneity tests** box. Click Continue to return to the **Univariate** window.
5. When the **Univariate** window opens, click OK to run the analysis. See Table 7.2 for the results.





7.4.4 SPSS Syntax Method

```
UNIANOVA ERRORS BY STRATEGY LIST
/METHOD = SSTYPE(3)
/INTERCEPT = INCLUDE
/PLOT = PROFILE(STRATEGY*LIST)
/EMMEANS = TABLES(STRATEGY)
/EMMEANS = TABLES(LIST)
/EMMEANS = TABLES(STRATEGY*LIST)
/PRINT = HOMOGENEITY
/CRITERIA = ALPHA(.05)
/DESIGN = STRATEGY LIST STRATEGY*LIST.
```

7.4.5 SPSS Output

TABLE 7.2

2 × 2 ANOVA Output

Univariate Analysis of Variance

Between-Subjects Factor

		Value Label	*N*
Strategy	1.00	STRATEGY A	12
	2.00	STRATEGY B	12
List	1.00	EASY	12
	2.00	HARD	12

Levene's Test of Equality of Error Variances[a]

Dependent Variable: Errors

F	d*f*1	d*f*2	Significance
.016	3	20	.997

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

[a] Design: Intercept + strategy + list + strategy × list.

Tests of Between-Subjects Effects

Dependent Variable: Errors

Source	Type III Sum of Squares	d*f*	Mean Square	*F*	Sig.
Corrected model	372.125[a]	3	124.042	13.091	.000
Intercept	5133.375	1	5133.375	541.781	.000
Strategy	5.042	1	5.042	.532	.474
List	145.042	1	145.042	15.308	.001
Strategy * list	222.042	1	222.042	23.434	.000
Error	189.500	20	9.475		
Total	5695.000	24			
Corrected total	561.625	23			

[a] R Squared = .663 (Adjusted R Squared = .612)

Estimated Marginal Means

1. Strategy

Dependent Variable: Errors

			95% Confidence Interval	
Strategy	Mean	Std Error	Lower Bound	Upper Bound
STRATEGY A	14.167	.889	12.313	16.020
STRATEGY B	15.083	.889	13.230	16.937

TABLE 7.2 (*Continued*)

2 × 2 ANOVA Output

Estimated Marginal Means

2. List

Dependent Variable: Errors

			95% Confidence Interval	
List	**Mean**	**Std Error**	**Lower Bound**	**Upper Bound**
EASY	12.167	.889	10.313	14.020
HARD	17.083	.889	15.230	18.937

3. Strategy * List

Dependent Variable: Errors

				95% Confidence Interval	
Strategy	**List**	**Mean**	**Std. Error**	**Lower Bound**	**Upper Bound**
STRATEGY A	EASY	8.667	1.257	6.045	11.288
	HARD	19.667	1.257	17.045	22.288
STRATEGY B	EASY	15.667	1.257	13.045	18.288
	HARD	14.500	1.257	11.879	17.121

7.4.6 Results and Interpretation

The assumption of **homogeneity of variance** is tested by **Levene's test of equality of error variances**, which tests the hypothesis that the population error variances are equal. In this example, the Levene statistic is $F = 0.016$ and the corresponding level of significance is large (i.e., $p > .05$) (see Table 7.2). Thus, the assumption of homogeneity of variance has not been violated.

7.4.6.1 Main Effect

The main effect of **STRATEGY** is not significant, $F(1,20) = 0.53$, $p > .05$ (see Table 7.2). From the estimated marginal means, the number of errors made by the **Strategy A group** ($M = 14.167$) is not significantly different from the number of errors produced by the **Strategy B group** ($M = 15.083$) (collapsing across the two LIST levels).

The main effect of **LIST** is significant, $F(1,20) = 15.31$, $p < .05$. From the estimated marginal means, it can be determined that the subjects produced significantly more errors in the hard list (M = 17.08) than in the easy list (M = 12.16) (collapsing across the two STRATEGY levels).

7.4.6.2 Interaction Effect

The **STRATEGY*LIST** interaction is significant, $F(1,20) = 23.43$, $p < .001$. To interpret the interaction, the task is made easier by graphing the

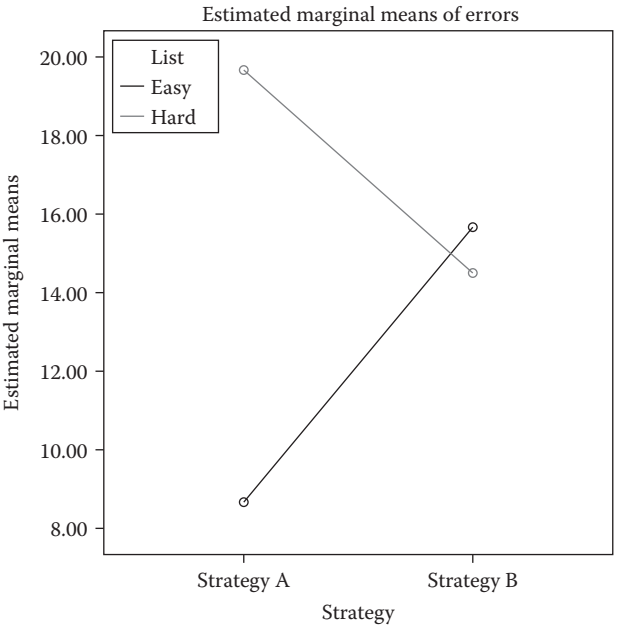


FIGURE 7.1
2 (STRATEGY) × 2 (LIST) interaction effect.

STRATEGY*LIST estimated marginal means from Table 7.2, as shown in Figure 7.1.

From Figure 7.1, it can be determined that the effect of learning strategy on the number of errors made is dependent on the difficulty of the list learned. Under strategy A, subjects made more errors on the hard list than on the easy list, but under strategy B, the effect is opposite, with subjects making more errors on the easy list than on the hard list.

7.4.7 Post Hoc Test for Simple Effects

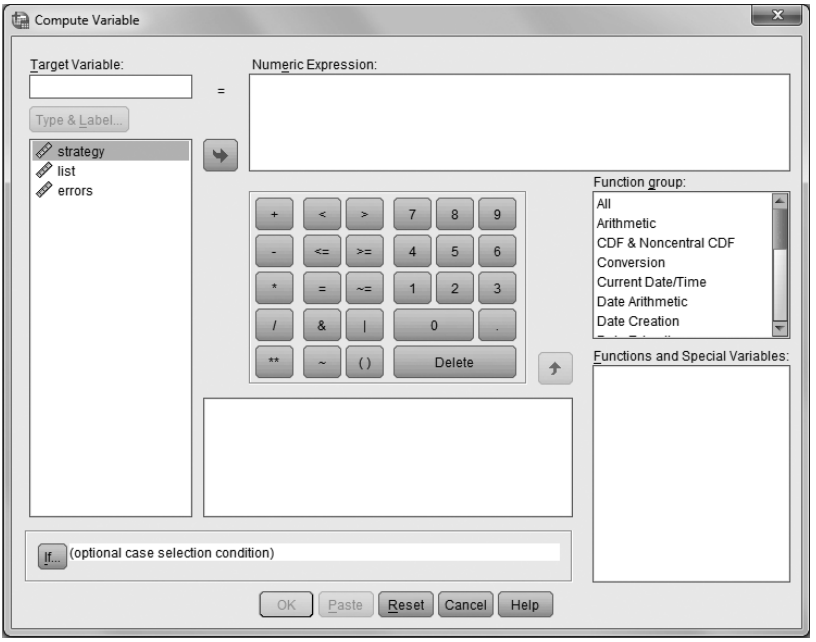
The significant interaction effect indicates that the outcome of one independent variable on the dependent variable is dependent on the second independent variable, i.e., the four experimental conditions (Strategy A-Easy List, Strategy A-Hard List, Strategy B-Easy List, Strategy B-Hard List) differ significantly in affecting the number of errors made. Nonetheless, the interaction effect does not show where the differences are, i.e., between which experimental conditions. To identify specific differences, **post hoc** comparisons can be used to "tease apart" the interaction. This is equivalent to the test for simple effects, i.e., the effect of one factor (IV1) at one level of the other factor (IV2).

Unfortunately, in a factorial design such as this, post hoc comparisons between the four experimental conditions (Strategy A-Easy List, Strategy A-Hard List, Strategy B-Easy List, Strategy B-Hard List) cannot be directly conducted within the SPSS GLM Univariate program. Rather, post hoc comparisons can only be executed through the **ONEWAY ANOVA** analysis. This procedure requires some data manipulation to convert the four experimental conditions into four levels of the same grouping variable. These four levels (Strategy A-Easy List, Strategy A-Hard List, Strategy B-Easy List, Strategy B-Hard List) can then be compared using the **Scheffé** post hoc test via the **ONEWAY ANOVA** analysis.

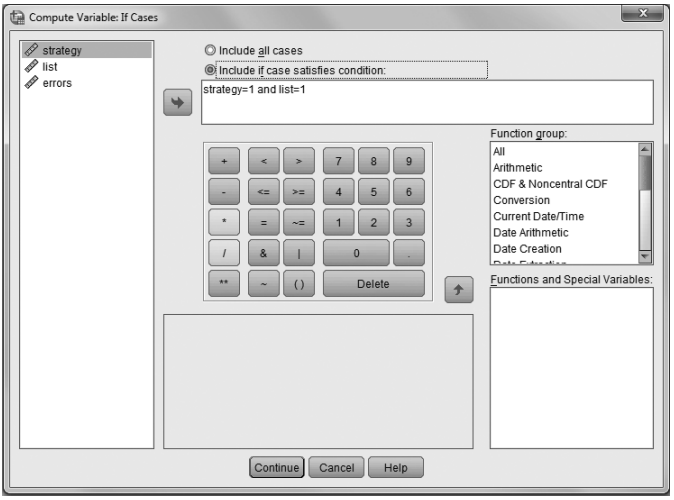
7.4.8 Data Transformation

7.4.8.1 Windows Method

1. The first step is to create a new grouping variable called **GROUP** that contains the four levels (Strategy A-Easy List, Strategy A-Hard List, Strategy B-Easy List, Strategy B-Hard List) generated by the **STRATEGY*LIST** interaction. From the menu bar, click **Transform** and then **Compute Variable.** The following **Compute Variable** window will open.



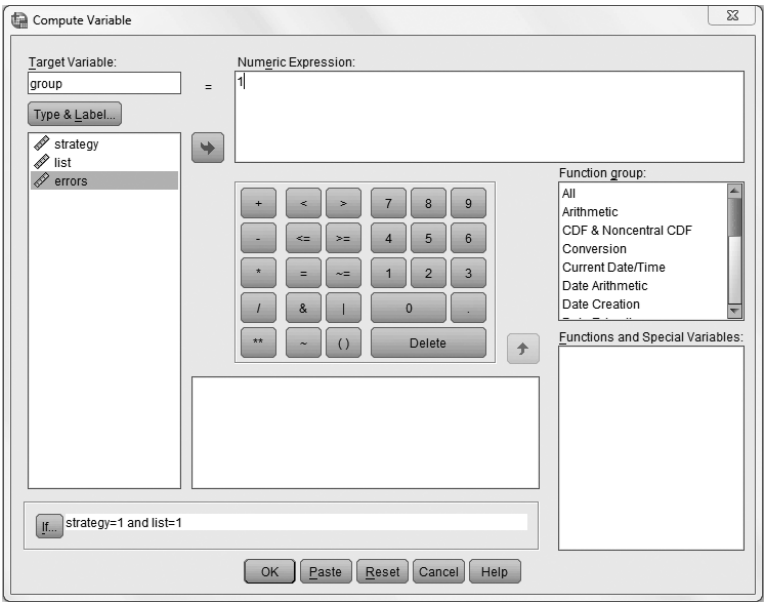
2. Click [If...] to open the **Compute Variable: If Cases** window below. Ensure that the **Include if case satisfies the condition:** field is checked. Using the data entry format codes presented in Section 7.4.1, create the first level (**Strategy A-Easy List**) by typing **strategy = 1 and list = 1** in the field. Click [Continue].

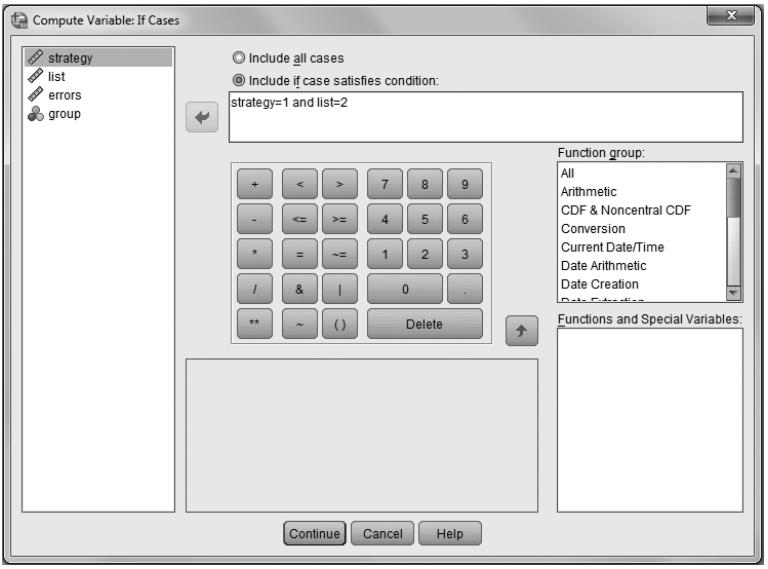


3. Since the **Strategy A-Easy List** level is the first of four levels within a new grouping variable called **GROUP**, this level will be coded **1** within the **GROUP** variable. When the following **Compute Variable** window opens, type **GROUP** in the **Target Variable:** field and **1** in the **Numeric Expression:** field. Click [OK] to create the first level of **Strategy A-Easy List** (coded **GROUP = 1**) within the new grouping variable of **GROUP**.
4. Repeat steps 1 to 3 to create the other three levels: **Strategy A-Hard List** (coded **GROUP = 2**), **Strategy B-Easy List** (coded **GROUP = 3**), and **Strategy B-Hard List** (coded **GROUP = 4**).

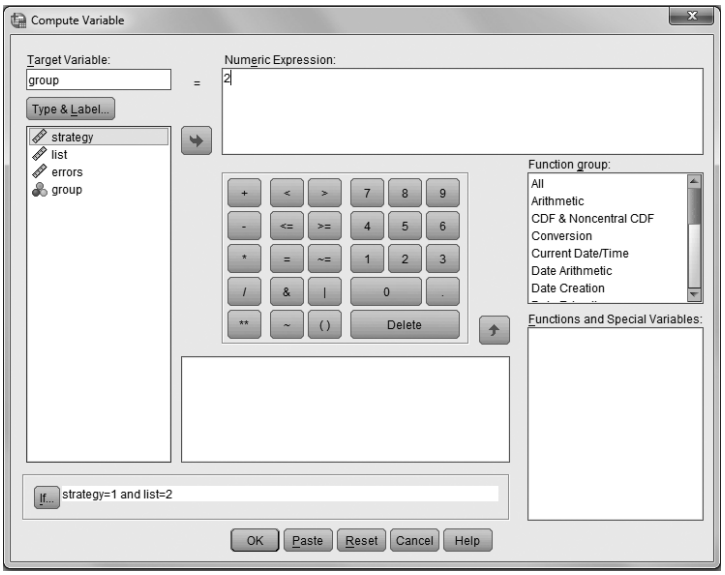
   For example, to create the second level of **Strategy A-Hard List**, open the **Compute Variable: If Cases** window. Type **strategy = 1 and list = 2** in the **Include if case satisfies condition:** field. Click [Continue].

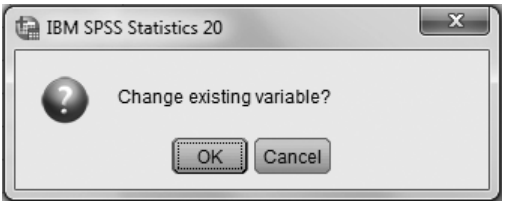
   When the following **Compute Variable** window opens, type **GROUP** in the **Target Variable:** field and **2** in the **Numeric Expression:** field. Click [OK] to create the second level of **Strategy A-Hard List** (coded **GROUP = 2**) within the new grouping variable of **GROUP**.
5. SPSS will then ask whether you want to "Change existing variable?" Click [OK] to complete the creation of this second level.

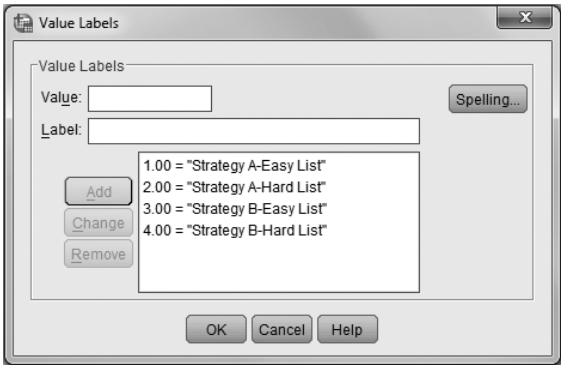




6. To aid interpretation of the obtained results, **Value Labels** in the data set should be activated and labels attached to the numerical codes for the four levels. To do this, open the data set, and under **Variable View**, click the **Values** field for the **GROUP** variable. Type in the value labels as indicated in the **Value Labels** window below.



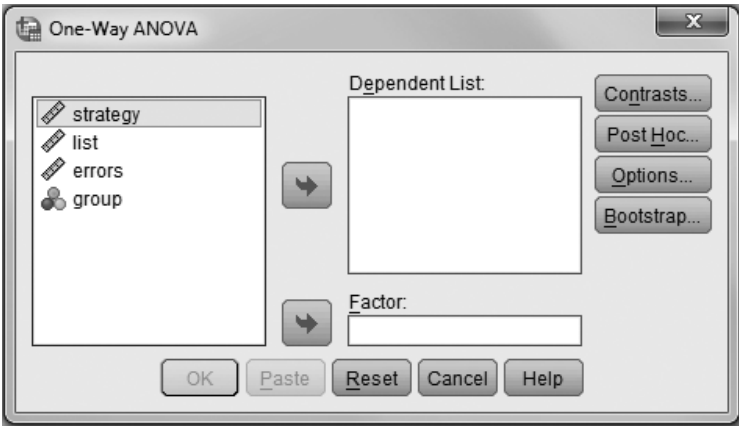




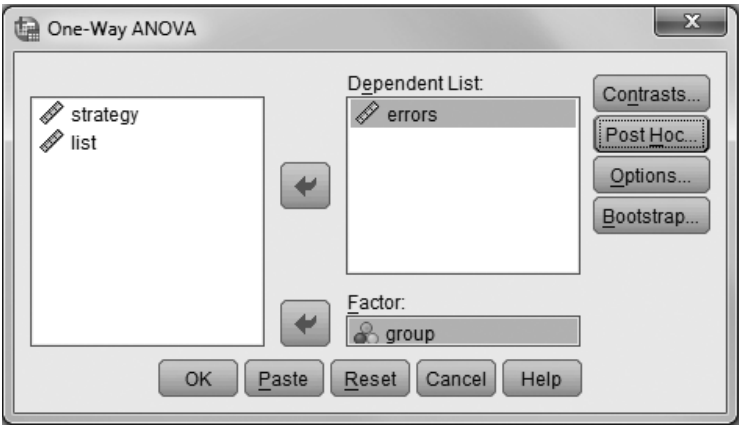
7.4.8.2 Post Hoc Comparisons: Windows Method

Once the four levels (Strategy A-Easy List, Strategy A-Hard List, Strategy B-Easy List, Strategy B-Hard List) have been created, **Scheffé** post hoc comparisons can be carried out to test for differences (simple effects) between these four levels.

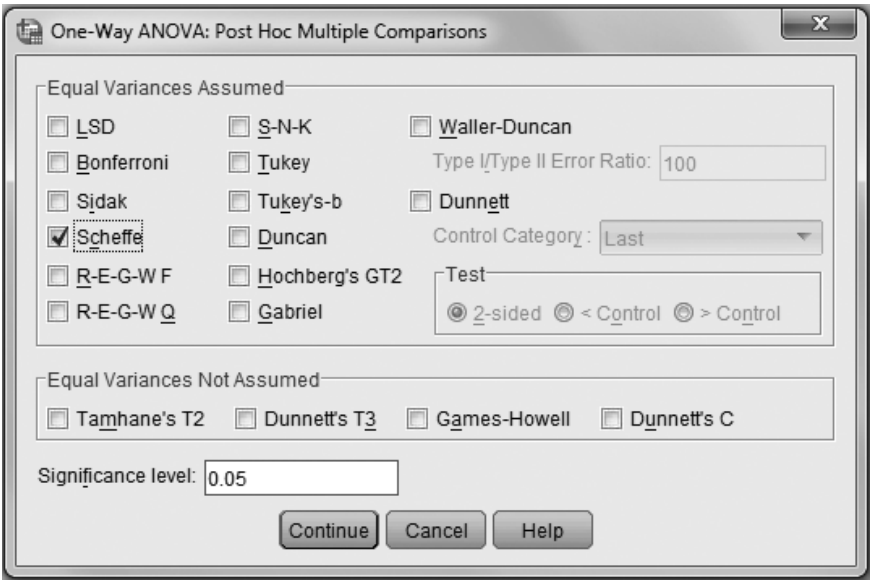
1. From the menu bar, click **Analyze**, then **Compare Means**, and then **One-Way ANOVA**. The following **One-Way ANOVA** window will open. Note that the listing of variables now includes the newly created grouping variable of **GROUP**, which contains the four levels **Strategy A-Easy List**, **Strategy A-Hard List**, **Strategy B-Easy List**, and **Strategy B-Hard List**.



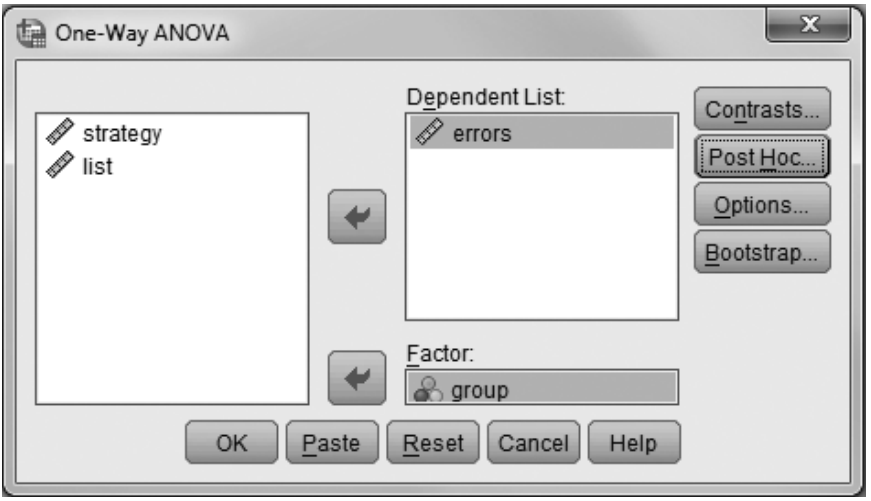
2. Transfer the dependent variable **ERRORS** to the **Dependent List:** field by clicking (highlight) the variable and then clicking ▸. Transfer the independent variable **GROUP** to the **Factor:** field by clicking (highlight) the variable and then clicking ▸. Click Post Hoc... to execute a post hoc comparison between the four levels (**Strategy A-Easy List**, **Strategy A-Hard List**, **Strategy B-Easy List**, and **Strategy B-Hard List**).



3. When the following **One-Way ANOVA: Post Hoc Multiple Comparisons** window opens, check the **Scheffe** field to run the Scheffé post hoc test. Next, click Continue.



4. When the **One-Way ANOVA** window opens, run the analysis by clicking OK. See Table 7.3 for the results.



7.4.8.3 Post Hoc Comparisons: SPSS Syntax Method

```
IF (STRATEGY EQ 1 AND LIST EQ 1) GROUP = 1.
IF (STRATEGY EQ 1 AND LIST EQ 2) GROUP = 2.
IF (STRATEGY EQ 2 AND LIST EQ 1) GROUP = 3.
IF (STRATEGY EQ 2 AND LIST EQ 2) GROUP = 4.

VALUE LABELS GROUP 1 'STRATEGY A-EASY LIST'
  2 'STRATEGY A-HARD LIST'
  3 'STRATEGY B-EASY LIST'
  4 'STRATEGY B-HARD LIST'.

ONEWAY ERRORS BY GROUP
/STATISTICS DESCRIPTIVES
/MISSING ANALYSIS
/POSTHOC = SCHEFFE ALPHA(0.05).
```

7.4.9 SPSS Output

7.4.10 Results and Interpretation

The results from the Scheffé comparisons (see Table 7.2) indicate that the significant **STRATEGY*LIST** interaction is due primarily to the subjects making significantly less errors in the Strategy A-Easy List condition (M = 8.66) than in the Strategy B-Easy List (M = 15.66), Strategy A-Hard List (M = 19.66), and Strategy B-Hard List (M = 14.50) conditions. No other conditions are significantly different from each other.

TABLE 7.3

Scheffé Post Hoc Comparisons Output

Descriptives

Errors

					95% Confidence Interval for Mean			
	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
STRATEGY A-EASY LIST	6	8.6667	3.01109	1.22927	5.5067	11.8266	5.00	13.00
STRATEGY B-HARD LIST	6	19.6667	3.07679	1.25610	16.4378	22.8956	15.00	23.00
STRATEGY B-EASY LIST	6	15.6667	2.94392	1.20185	12.5772	18.7561	12.00	20.00
STRATEGY B-HARD LIST	6	14.5000	3.27109	1.33542	11.0672	17.9328	11.00	20.00
Total	24	14.6250	4.94151	1.00868	12.5384	16.7116	5.00	23.00

(*Continued*)

TABLE 7.3 (*Continued*)

Scheffé Post Hoc Comparisons Output

Multiple Comparisons

Dependent Variable: Errors

Scheffe

(I) Group	(J) Group	Mean Difference (I–J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Strategy A-Easy List	Strategy A-Hard List	−11.00000*	1.77717	.000	−16.4182	-5.5818
	Strategy B-Easy List	−7.00000*	1.77717	.008	−12.4182	-1.5818
	Strategy B-Hard List	−5.83333*	1.77717	.032	−11.2516	-.4151
Strategy A-Hard List	Strategy A-Easy List	11.00000*	1.77717	.000	5.5818	16.4182
	Strategy B-Easy List	4.00000	1.77717	.201	−1.4182	9.4182
	Strategy B-Hard List	5.16667	1.77717	.065	−.2516	10.5849
Strategy B-Easy List	Strategy A-Easy List	7.00000*	1.77717	.008	1.5818	12.4182
	Strategy A-Hard List	−4.00000	1.77717	.201	−9.4182	1.4182
	Strategy B-Hard List	1.16667	1.77717	.933	−4.2516	6.5849
Strategy B-Hard List	Strategy A-Easy List	5.83333*	1.77717	.032	.4151	11.2516
	Strategy A-Hard List	−5.16667	1.77717	.065	−10.5849	.2516
	Strategy B-Easy List	−1.16667	1.77717	.933	−6.5849	4.2516

* The mean difference is significant at the .05 level.

7.5 Example 2: Three-Way Factorial (2 × 2 × 2 Factorial)

The previous example can be extended to a three-way factorial design. Assume that the researcher, in addition to determining the effects of the two types of learning strategy (A and B) on the memorization of easy and difficult lists, is also concerned with determining the effects of high- versus low-intensity shock on learning in the above conditions. The design includes shock level as an additional factor, making the entire experiment a three-factor design. The researcher records the total number of mistakes made by each subject. The three-way factorial combination of the three independent variables yields the following eight experimental groups, with six cases per group.

	Strategy A		Strategy B	
	Easy List	**Hard List**	**Easy List**	**Hard List**
Low Shock	s1 17	s7 25	s13 13	s19 27
	s2 10	s8 18	s14 18	s20 29
	s3 11	s9 19	s15 10	s21 31
	s4 9	s10 19	s16 16	s22 36
	s5 10	s11 16	s17 22	s23 40
	s6 6	s12 13	s18 18	s24 28
High Shock	s25 16	s31 26	s37 13	s43 41
	s26 8	s32 18	s38 19	s44 34
	s27 5	s33 12	s39 14	s45 40
	s28 7	s34 20	s40 16	s46 41
	s29 8	s35 15	s41 23	s47 35
	s30 9	s36 19	s42 20	s48 33

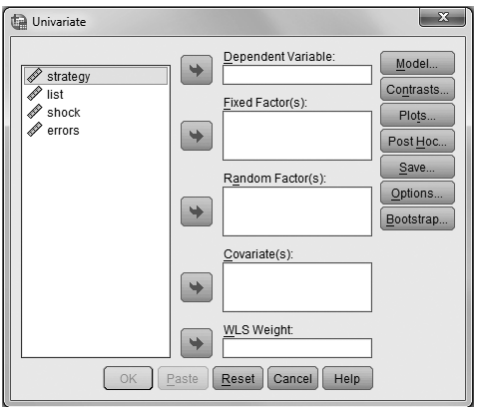
7.5.1 Data Entry Format

The data set has been saved under the name: **EX7b.SAV**

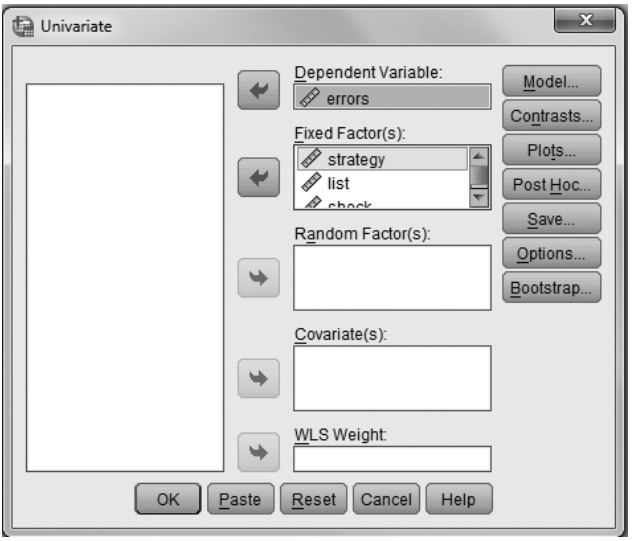
Variables	Column(s)	Code
• STRATEGY	1	1 = Strategy A, 2 = Strategy B
• LIST	2	1 = Easy, 2 = Hard
• SHOCK	3	1 = Low, 2 = High
• ERRORS	4	Number of errors made

7.5.2 Windows Method

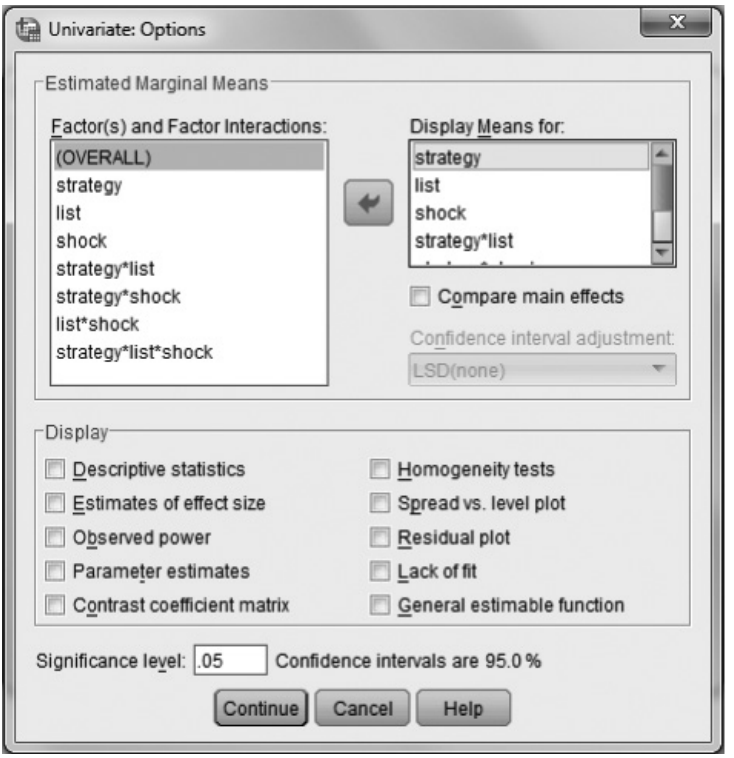
1. From the menu bar, click **Analyze**, then **General Linear Model**, and then **Univariate**. The following **Univariate** window will open.

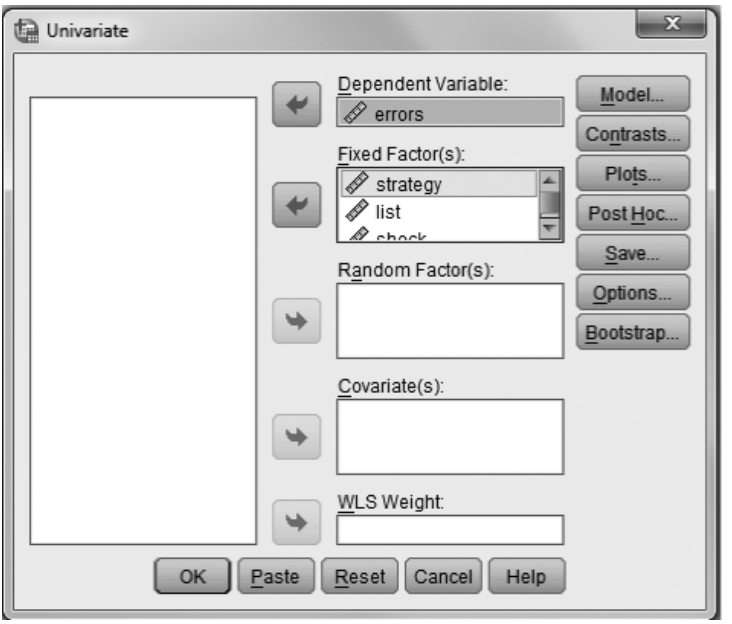


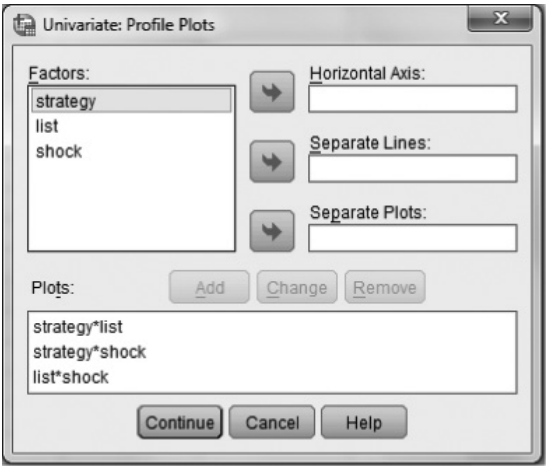
2. Transfer the dependent variable **ERRORS** to the **Dependent Variable:** field by clicking (highlight) the variable and then clicking ▶. Transfer the independent variables **STRATEGY, LIST** and **SHOCK** to the **Fixed Factor(s):** field by clicking (highlight) the variables and then clicking ▶.



3. Click Options... to obtain means for the **STRATEGY**, **LIST**, and **SHOCK** factors, and the **STRATEGY*LIST**, **STRATEGY*SHOCK**, **LIST*SHOCK**, and **STRATEGY*LIST*SHOCK** interactions. When the **Univariate: Options** window opens, click (highlight) **STRATEGY**, **LIST**, **SHOCK**, **STRATEGY*LIST**, **STRATEGY*SHOCK**, **LIST*SHOCK**, and **STRATEGY*LIST*SHOCK** in the **Factor(s) and Factor Interactions:** field, and then click ▶ to transfer these factors and factor interactions to the **Display Means for:** field. Click Continue to return to the **Univariate** window.
4. When the **Univariate** window opens, click Plots... to plot graphs of the three two-way interactions (**STRATEGY*LIST**, **STRATEGY*SHOCK**, **LIST*SHOCK**).
5. When the following **Univariate: Profile Plots** window opens, transfer the **STRATEGY** variable to the **Horizontal Axis:** field by clicking (highlight) the variable and then clicking ▶. Transfer the **LIST** variable to the **Separate Lines:** field by clicking (highlight) the variable and then clicking ▶. Next, click Add to transfer the **STRATEGY*LIST** interaction to the **Plots:** field. Repeat this procedure to add the **STRATEGY*SHOCK** and **LIST*SHOCK** interactions to the **Plots:** field. When this is done, click Continue.


Univariate: Options
Estimated Marginal Means
Factor(s) and Factor Interactions:
(OVERALL)
strategy
list
shock
strategy*list
strategy*shock
list*shock
strategy*list*shock
Display Means for:
strategy
list
shock
strategy*list
Compare main effects
Confidence interval adjustment:
LSD(none)
Display
Descriptive statistics
Estimates of effect size
Observed power
Parameter estimates
Contrast coefficient matrix
Homogeneity tests
Spread vs. level plot
Residual plot
Lack of fit
General estimable function
Significance level: .05
Confidence intervals are 95.0 %
Continue
Cancel
Help


Univariate
Dependent Variable:
errors
Fixed Factor(s):
strategy
list
Random Factor(s):
Covariate(s):
WLS Weight:
Model...
Contrasts...
Plots...
Post Hoc...
Save...
Options...
Bootstrap...
OK
Paste
Reset
Cancel
Help



6. When the Univariate window opens, run the analysis by clicking OK. See Table 7.4 for the results.

7.5.3 SPSS Syntax Method

```
GLM ERRORS BY STRATEGY LIST SHOCK
/PLOT = PROFILE(STRATEGY*LIST)
/PLOT = PROFILE(STRATEGY*SHOCK)
/PLOT = PROFILE(LIST*SHOCK)
/EMMEANS = TABLES(STRATEGY)
/EMMEANS = TABLES(LIST)
/EMMEANS = TABLES(SHOCK)
/EMMEANS = TABLES(STRATEGY*LIST)
/EMMEANS = TABLES(STRATEGY*SHOCK)
/EMMEANS = TABLES(LIST*SHOCK)
/EMMEANS = TABLES(STRATEGY*LIST*SHOCK).
```

7.5.4 SPSS Output

TABLE 7.4

2 × 2 × 2 ANOVA Output

General Linear Model			
Between-Subject Factors			
		Value Label	***N***
Strategy	1.00	STRATEGY A	24
	2.00	STRATEGY B	24
List	1.00	EASY	24
	2.00	HARD	24
Shock	1.00	LOW	24
	2.00	HIGH	24

TABLE 7.4 (*Continued*)

2 × 2 × 2 ANOVA Output

Tests of Between-Subjects Effects

Dependent Variable: Errors

Source	Type III Sum of Squares	d*f*	Mean Square	*F*	Sig.
Corrected model	4090.479[a]	7	584.354	33.802	.000
Intercept	18921.021	1	18921.021	1094.491	.000
Strategy	1645.021	1	1645.021	95.157	.000
List	2093.521	1	2093.521	121.100	.000
Shock	20.021	1	20.021	1.158	.288
Strategy * list	247.521	1	247.521	14.318	.001
Strategy * shock	54.187	1	54.187	3.134	.084
List * shock	25.521	1	25.521	1.476	.231
Strategy * list * shock	4.688	1	4.688	.271	.605
Error	691.500	40	17.288		
Total	23703.000	48			
Corrected total	4781.979	47			

[a] R Squared = .855 (Adjusted R Squared = .830)

Estimated Marginal Means

1. Strategy

Dependent Variable: Errors

			95% Confidence Interval	
Strategy	Mean	Std. Error	Lower Bound	Upper Bound
STRATEGY A	14.000	.849	12.285	15.715
STRATEGY B	25.708	.849	23.993	27.424

2. List

Dependent Variable: Errors

			95% Confidence Interval	
List	Mean	Std. Error	Lower Bound	Upper Bound
EASY	13.250	.849	11.535	14.965
HARD	26.458	.849	24.743	28.174

3. Shock

Dependent Variable: Errors

			95% Confidence Interval	
Shock	Mean	Std. Error	Lower Bound	Upper Bound
LOW	19.208	.849	17.493	20.924
HIGH	20.500	.849	18.785	22.215

(*Continued*)

TABLE 7.4 (*Continued*)

2 × 2 × 2 ANOVA Output

4. Strategy * List

Dependent Variable: Errors

				95% Confidence Interval	
Strategy	**List**	**Mean**	**Std. Error**	**Lower Bound**	**Upper Bound**
STRATEGY A	EASY	9.667	1.200	7.241	12.092
	HARD	18.333	1.200	15.908	20.759
STRATEGY B	EASY	16.833	1.200	14.408	19.259
	HARD	34.583	1.200	32.158	37.009

5. Strategy * Shock

Dependent Variable: Errors

				95% Confidence Interval	
Strategy	**Shock**	**Mean**	**Std. Error**	**Lower Bound**	**Upper Bound**
STRATEGY A	LOW	14.417	1.200	11.991	16.842
	HIGH	13.583	1.200	11.158	16.009
STRATEGY B	LOW	24.000	1.200	21.574	26.426
	HIGH	27.417	1.200	24.991	29.842

6. List * Shock

Dependent Variable: Errors

				95% Confidence Interval	
List	**Shock**	**Mean**	**Std. Error**	**Lower Bound**	**Upper Bound**
EASY	LOW	13.333	1.200	10.908	15.759
	HIGH	13.167	1.200	10.741	15.592
HARD	LOW	25.083	1.200	22.658	27.509
	HIGH	27.833	1.200	25.408	30.259

7. Strategy * List * Shock

Dependent Variable: Errors

					95% Confidence Interval	
Strategy	**List**	**Shock**	**Mean**	**Std. Error**	**Lower Bound**	**Upper Bound**
STRATEGY A	EASY	LOW	10.500	1.697	7.069	13.931
		HIGH	8.833	1.679	5.403	12.264
	HARD	LOW	18.333	1.697	14.903	21.764
		HIGH	18.333	1.697	14.903	21.764
STRATEGY B	EASY	LOW	16.167	1.697	12.736	19.597
		HIGH	17.500	1.679	14.069	20.931
	HARD	LOW	31.833	1.697	28.403	35.264
		HIGH	37.333	1.679	33.903	40.764

7.5.5 Results and Interpretation

7.5.5.1 Main Effects

The main effect of **STRATEGY** is significant, $F(1,40) = 95.16$, $p < .001$ (see Table 7.4). From the estimated marginal means, subjects made significantly more errors under **Strategy B** ($M = 25.71$) than under **Strategy A** ($M = 14.00$) (collapsing across the **LIST** and **SHOCK** factors).

The main effect of **LIST** is significant, $F(1,40) = 121.10$, $p < .001$. Subjects made significantly more errors in the **hard list** condition ($M = 26.45$) than in the **easy list** condition ($M = 13.25$) (collapsing across the **STRATEGY** and **SHOCK** factors).

The main effect for **SHOCK** is not significant, $F(1,40) = 1.16$, $p > .05$. The number of errors made under the **low shock** condition ($M = 19.21$) is not significantly different from the number of errors made under the **high shock** condition ($M = 20.50$) (collapsing across the **LIST** and **STRATEGY** factors).

7.5.5.2 Two-Way Interactions

*7.5.5.2.1 Strategy*List Interaction*

The **STRATEGY*LIST** interaction is significant, $F(1,40) = 14.32$, $p < .01$. Interpretation of this interaction can be facilitated by graphing the **STRATEGY*LIST** estimated marginal means from Table 7.4, as shown in Figure 7.2.

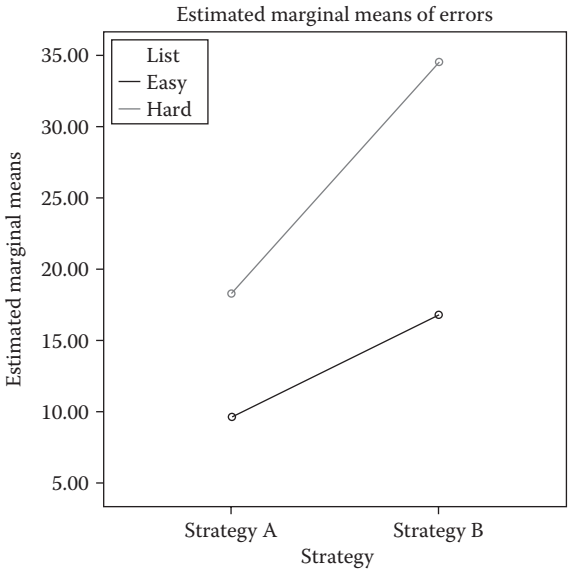


FIGURE 7.2
2 (STRATEGY) × 2 (LIST) interaction effect.

The significant interaction effect indicates that the effect of learning strategy on the number of errors made is dependent on the difficulty of the list learned. While the number of errors made increased from Strategy A to Strategy B when learning either the hard or easy list, the increase is more pronounced when learning the hard list than the easy list.

7.5.5.2.2 Post Hoc Comparisons

Post hoc comparisons can now be carried out to clarify the above interaction, i.e., to locate differences between the four experimental conditions. The **Windows** method for carrying out this post hoc comparison is identical to that presented in **Section 7.4.8.1, Data Transformation**, from step 1 through to step 6, and **Section 7.4.8.2, Post Hoc Comparisons: Windows Method**, from step 1 through to step 4. The **SPSS syntax method** used for carrying out these comparisons is identical to that presented in **Section 7.4.8.3, Post Hoc Comparisons: SPSS Syntax Method**. See Table 7.5 for the results.

7.5.2.2.3 SPSS Output

TABLE 7.5

Scheffé Post Hoc Comparisons

Multiple Comparisons						
Dependent Variable: Errors						
Scheffe						
					95% Confidence Interval	
(I) GROUP	**(J) GROUP**	**Mean Difference (I–J)**	**Std. Error**	**Sig.**	**Lower Bound**	**Upper Bound**
STRATEGY A-EASY LIST	STRATEGY A-HARD LIST	−8.66667*	1.73633	.000	−13.7138	−3.6195
	STRATEGY B-EASY LIST	−7.16667*	1.73633	.002	−12.2138	−2.1195
	STRATEGY B-HARD LIST	−24.91667*	1.73633	.000	−29.9638	−19.8695
STRATEGY A-HARD LIST	STRATEGY A-EASY LIST	8.66667*	1.73633	.000	3.6195	13.7138
	STRATEGY B-EASY LIST	1.50000	1.73633	.862	−3.5471	6.5471
	STRATEGY B-HARD LIST	−16.25000*	1.73633	.000	−21.2971	−11.2029

TABLE 7.5 (*Continued*)

Scheffé Post Hoc Comparisons

Multiple Comparisons						
Dependent Variable: Errors						
Scheffe						
					95% Confidence Interval	
(I) GROUP	**(J) GROUP**	**Mean Difference (I–J)**	**Std. Error**	**Sig.**	**Lower Bound**	**Upper Bound**
STRATEGY B-EASY LIST	STRATEGY A-EASY LIST	7.16667*	1.73633	.002	2.1195	12.2138
	STRATEGY A-HARD LIST	−1.50000	1.73633	.862	−6.5471	3.5471
	STRATEGY B-HARD LIST	−17.75000*	1.73633	.000	−22.7971	−12.7029
STRATEGY B-HARD LIST	STRATEGY A-EASY LIST	24.91667*	1.73633	.000	19.8695	29.9638
	STRATEGY A-HARD LIST	16.25000*	1.73633	.000	11.2029	21.2971
	STRATEGY B-HARD LIST	17.75000*	1.73633	.000	12.7029	22.7971

* The mean difference is significant at the .05 level.

7.5.2.2.4 Results and Interpretation

Results of the post hoc comparisons (Table 7.5) indicate that the significant **STRATEGY*LIST** interaction is due primarily to subjects making significantly less errors in the Strategy A-Easy List condition than in the other three experimental conditions (A-Hard, B-Easy, B-Hard), and to subjects making significantly more errors in the Strategy B-Hard List condition than in the other three experimental conditions (B-Easy, A-Easy, A-Hard).

*7.5.2.2.5 Strategy*Shock Interaction*

The **STRATEGY*SHOCK** interaction is not significant, $F(1,40) = 3.13$, $p > .05$. Interpretation of this interaction can be facilitated by graphing the **STRATEGY*SHOCK** estimated marginal means from Table 7.4, as shown in Figure 7.3.

As the interaction is not significant, the result can be interpreted in terms of the significant main effect for **STRATEGY**. That is, the effect of **STRATEGY** on the number of errors made is not dependent on the levels of

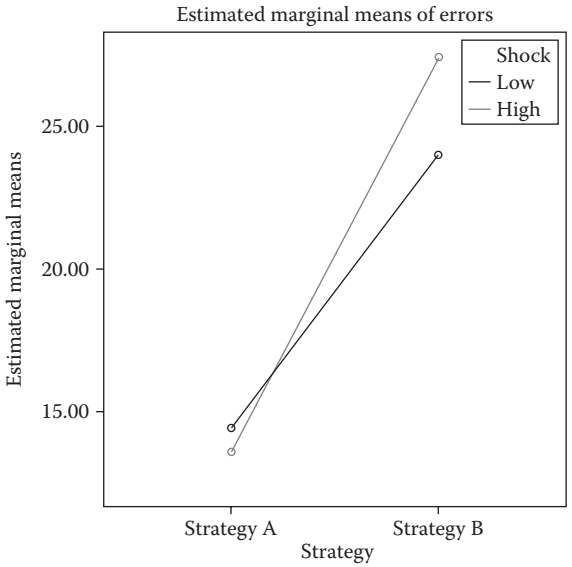


FIGURE 7.3
2 (STRATEGY) × 2 (SHOCK) interaction effect.

SHOCK, such that regardless of **SHOCK** level, subjects made significantly more errors under Strategy B ($M = 25.71$) than under Strategy A ($M = 14.00$).

*7.5.2.2.6 List*Shock Interaction*

The **LIST*SHOCK** interaction is not significant, $F(1,40) = 1.47$, $p > .05$. Interpretation of this interaction can be facilitated by graphing the **LIST*SHOCK** estimated marginal means from Table 7.4, as shown in Figure 7.4.

The effect of **LIST** on the number of errors made is not dependent on the levels of **SHOCK**, such that regardless of **SHOCK** level, subjects made more errors on the hard list ($M = 26.45$) than on the easy list ($M = 13.25$).

7.5.6 Strategy*List*Shock Interaction

In order to plot the three-way **STRATEGY*LIST*SHOCK** interaction, some data transformation must be carried out. Suppose the researcher decides to plot the variables **STRATEGY** and **LIST** (i.e., **STRATEGY*LIST** interaction) against the variable **SHOCK**. The first step is to create a new grouping variable called **GROUP** that contains the four levels generated by the **STRATEGY*LIST** interaction (Strategy A-Easy List, Strategy A-Hard List, Strategy B-Easy List, Strategy B-Hard List). To do this, follow step 1 through step 6 in Section 7.4.8.1.

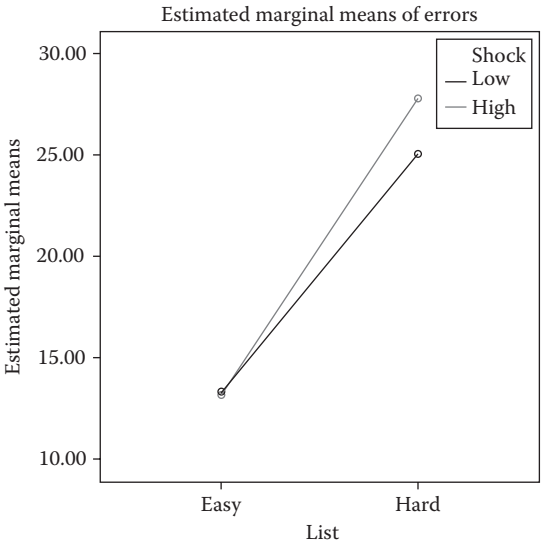
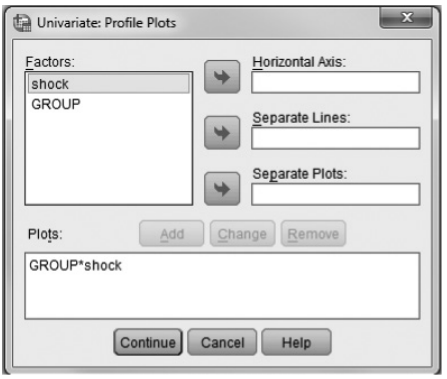


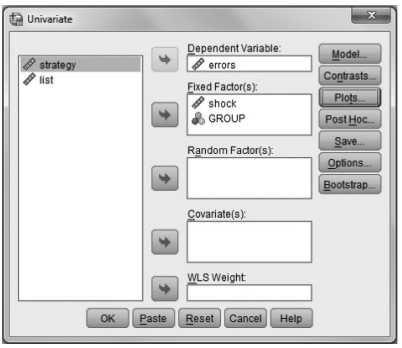
FIGURE 7.4
2 (LIST) × 2 (SHOCK) interaction effect.

7.5.6.1 Windows Method

1. When the data transformation procedure has been completed, the newly created grouping variable **GROUP** (containing the four levels Strategy A-Easy List, Strategy A-Hard List, Strategy B-Easy List, and Strategy B-Hard List) will be added to the data set. Click Plots... in the **Univariate** window to open the **Univariate: Profile Plots** window. Transfer the newly created **GROUP** variable to the **Horizontal Axis:** field by clicking (highlight) the variable and then clicking ▶. Transfer the **SHOCK** variable to the **Separate Lines:** field by clicking (highlight) the variable and then clicking ▶. Next, click Add to transfer the **SHOCK*GROUP** interaction to the **Plots** field. When this is done, click Continue.



2. When the **Univariate** window opens, run the analysis by clicking OK. See Table 7.6 for the results.



7.5.6.2 SPSS Syntax Method

```
IF (STRATEGY EQ 1 AND LIST EQ 1) GROUP = 1.
IF (STRATEGY EQ 1 AND LIST EQ 2) GROUP = 2.
IF (STRATEGY EQ 2 AND LIST EQ 1) GROUP = 3.
IF (STRATEGY EQ 2 AND LIST EQ 2) GROUP = 4.
VALUE LABELS GROUP 1 'STRATEGY A-EASY LIST'
                   2 'STRATEGY A-HARD LIST'
                   3 'STRATEGY B-EASY LIST'
                   4 'STRATEGY B-HARD LIST'.
UNIANOVA ERRORS BY SHOCK GROUP
/METHOD = SSTYPE(3)
/INTERCEPT = INCLUDE
/PLOT = PROFILE(GROUP*SHOCK)
/EMMEANS = TABLES(SHOCK)
/EMMEANS = TABLES(GROUP)
/EMMEANS = TABLES(SHOCK*GROUP)
/CRITERIA = ALPHA(.05)
/DESIGN = GROUP SHOCK SHOCK*GROUP.
```

7.5.7 SPSS Output

TABLE 7.6

2 × 2 × 2 ANOVA Output

Between-Subjects Factors			
		Value Label	*N*
SHOCK	1.00	LOW	24
	2.00	HIGH	24
GROUP	1.00	STRATEGY A-EASY LIST	12
	2.00	STRATEGY A-HARD LIST	12
	3.00	STRATEGY B-EASY LIST	12
	4.00	STRATEGY B-HARD LIST	12

TABLE 7.6 (*Continued*)

2 × 2 × 2 ANOVA Output

Tests of Between-Subjects Effects

Dependent Variable: Errors

Source	Type III Sum of Squares	d*f*	Mean Square	*F*	Significance
Corrected model	4090.479[a]	7	584.354	33.802	.000
Intercept	18921.021	1	18921.021	1094.491	.000
Group	3986.063	3	1328.688	76.858	.000
Shock	20.021	1	20.021	1.158	.288
Shock * Group	84.396	3	28.132	1.627	.198
Error	691.500	40	17.288		
Total	23703.000	48			
Corrected total	4781.979	47			

Estimated Marginal Means

1. Shock

Dependent Variable: Errors

Shock	Mean	Standard Error	95% Confidence Interval	
			Lower Bound	Upper Bound
LOW	19.208	.849	17.493	20.924
HIGH	20.500	.849	18.785	22.215

2. GROUP

Dependent Variable: Errors

GROUP	Mean	Standard Error	95% Confidence Interval	
			Lower Bound	Upper Bound
STRATEGY A-EASY LIST	9.667	1.200	7.241	12.092
STRATEGY A-HARD LIST	18.333	1.200	15.908	20.759
STRATEGY B-EASY LIST	16.833	1.200	14.408	19.259
STRATEGY B-HARD LIST	34.583	1.200	32.158	37.009

3. Shock * GROUP

Dependent Variable: Errors

Shock	GROUP	Mean	Standard Error	95% Confidence Interval	
				Lower Bound	Upper Bound
LOW	STRATEGY A-EASY LIST	10.500	1.697	7.069	13.931
	STRATEGY A-HARD LIST	18.333	1.697	14.903	21.764
	STRATEGY B-EASY LIST	16.167	1.697	12.736	19.597
	STRATEGY B-HARD LIST	31.833	1.697	28.403	35.264
HIGH	STRATEGY A-EASY LIST	8.833	1.697	5.403	12.264
	STRATEGY A-HARD LIST	18.333	1.697	14.903	21.764
	STRATEGY B-EASY LIST	17.500	1.697	14.069	20.931
	STRATEGY B-HARD LIST	37.333	1.697	33.903	40.764

[a] R Squared = .855 (adjusted R Squared = .830)

7.5.7.1 Shock × Group Interaction

The three-way interaction between **GROUP** and **SHOCK** is not significant, $F(3,40) = 1.63$, $p > .05$. Interpretation of this interaction can be facilitated by graphing the **SHOCK*GROUP** estimated marginal means from Table 7.6, as shown in Figure 7.5.

As the three-way interaction is not significant, it is legitimate to interpret the significant main effects of **LIST** and **STRATEGY**. For instance, the results from Table 7.3 indicate that more errors were made on the hard list ($M = 26.45$) than on the easy list ($M = 13.25$), and under Strategy B ($M = 25.71$) than under Strategy A ($M = 14.00$), regardless of shock level.

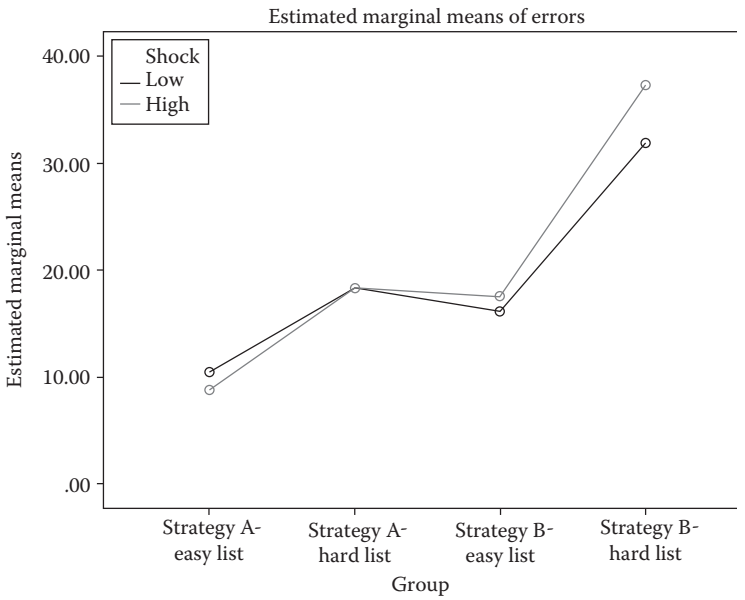


FIGURE 7.5
4 (GROUP) × 2 (SHOCK) interaction effect.

8

General Linear Model (GLM) Multivariate Analysis

8.1 Aim

In experiments involving multiple independent variables and one dependent variable, the *General Linear Model (GLM) univariate analysis of variance* is usually used to answer questions about the effects of the independent variables on the dependent variable. The final example in the previous chapter examined the effects of three independent variables (type of learning strategy, shock level, and difficulty of material) on the dependent variable of the number of errors made by each subject in learning one list of material. As the experiment involved only one dependent variable, the GLM univariate ($2 \times 2 \times 2$) analysis of variance was the appropriate test to use. However, if the experiment had required each subject to learn four different lists of material (instead of one list), the GLM univariate analysis would no longer be appropriate. This is because the dependent variable is no longer a single measure but four different scores obtained for each subject. While the GLM univariate analysis of variance can be conducted on an individual basis for each of the four dependent variables, the GLM multivariate analysis of variance is more appropriate. Unlike univariate tests, GLM multivariate analysis takes into account the interrelation among dependent variables, and analyzes the variables simultaneously (reducing the probability of committing Type I error).

8.2 Checklist of Requirements

- Depending on the research question and/or hypotheses to be tested, the experiment can include or exclude "classification" on independent variables.

- When independent variables are included, there can be two or more levels for each independent variable.
- There should be two or more dependent variables.

8.3 Assumptions

- **Independence of observations**—The observations must be autonomous (i.e., responses between groups of respondents should not be correlated). This can be tested with an intraclass correlation coefficient if lack of independence of observations is suspected.
- **Linearity**—Linear relationships among all pairs of dependent variables must be assumed.
- **Homogeneity of covariance matrices**—The variance of all DVs must be equal for all treatment groups defined by the IVs.
- **Multivariate normality**—The set of dependent variables must follow a multivariate normal distribution (i.e., any linear combination of the dependent variables must follow a Gaussian distribution). However, multivariate normality is difficult to test, although the likelihood of multivariate normality is increased if the variables are all normally distributed.

8.4 Example 1: GLM Multivariate Analysis: One-Sample Test

The one-sample GLM model incorporating multiple dependent variables is an extension of the one-sample *t* test, and is the simplest example of GLM. Where the one-sample *t* test is used to examine the hypothesis that the sample does not differ from a population with a known mean, the one-sample GLM tests the hypothesis that several observed means do not differ from a set of constants. That is, it tests the hypothesis that a set of means is equal to 0.

Suppose that a sports psychologist has recorded the following running times (in seconds) from five men in four different events: 50-yard dash, 100-yard dash, 200-yard dash, and 300-yard dash. The hypothesis to be tested is that the observed sample comes from a population with specified values for the means of the running events. For illustrative purposes, the standard values are taken to be 6 seconds for the 50-yard dash, 12 seconds for the 100-yard dash, 25 seconds for the 200-yard dash, and 40 seconds for the 300-yard dash. Since GLM automatically tests the hypothesis that a set of means is equal to 0, the normative values must be subtracted from the observed scores

and the hypothesis that the differences are 0 is tested. The running times for the four events are shown below.

	Running Events			
	50 (Yard)	**100 (Yard)**	**200 (Yard)**	**300 (Yard)**
s1	9	20	42	67
s2	9	19	45	66
s3	8	17	38	62
s4	9	18	46	64
s5	8	22	39	59

8.4.1 Data Entry Format

The data set has been saved under the name **EX8a.SAV**.

Variables	Column(s)	Code
T1	1	Running speed in seconds
T2	2	Running speed in seconds
T3	3	Running speed in seconds
T4	4	Running speed in seconds

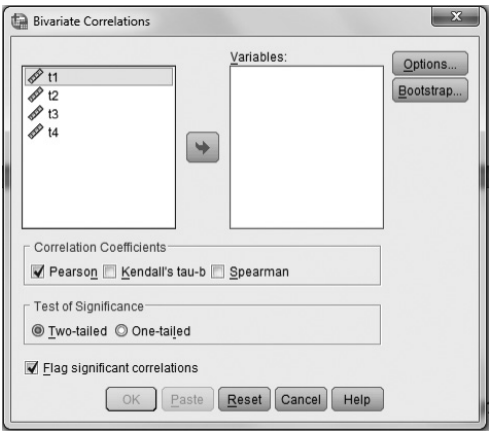
8.4.2 Testing Assumptions

8.4.2.1 Independence of Observations

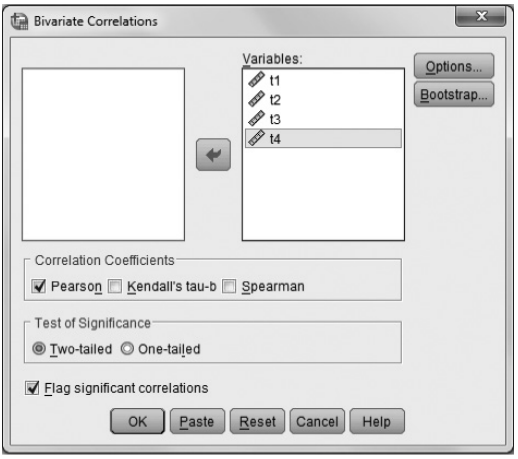
For the present example, independence of observations will be tested using correlation analysis between a pair of dependent variables.

8.4.2.1.1 Windows Method

1. From the menu bar, click **Analyze**, then **Correlate**, and then **Bivariate…**. The following **Bivariate Correlations** window will open.



2. Transfer the **T1**, **T2**, **T3**, and **T4** variables to the **Variables:** field by clicking (highlight) them and then clicking . By default, SPSS will employ **Pearson correlation analysis**, and a **two-tailed test of significance** (both fields are checked).



3. Click OK to complete the analysis. See Table 8.1 for the results.

8.4.2.1.2 SPSS Syntax Method

```
CORRELATIONS T1 T2 T3 T4
/MISSING = PAIRWISE.
```

8.4.2.1.3 SPSS Output

TABLE 8.1

Correlation Coefficients between T1, T2, T3, and T4

Correlations					
		t1	**t2**	**t3**	**t4**
t1	Pearson Correlation	1	−.142	.904	.882
	Sig. (2-tailed)		.819	.035	.048
	N	5	5	5	5
t2	Pearson Correlation	−.142	1	−.184	−.308
	Sig. (2-tailed)	.819		.767	.614
	N	5	5	5	5
t3	Pearson Correlation	.904	−.184	1	.639
	Sig. (2-tailed)	.035	.767		.246
	N	5	5	5	5
t4	Pearson Correlation	.882	−.308	.639	1
	Sig. (2-tailed)	.048	.614	.246	
	N	5	5	5	5

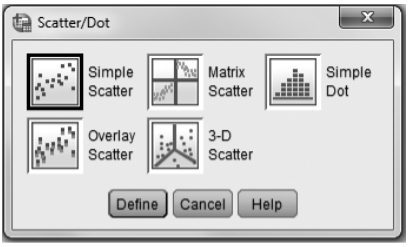
For this case, the probability that the nature of the study will lead to correlated observations is high. That is, all five men ran in four different distance events and, as such, one can reasonably presume that their performance in one running event will be affected by their performance in the other events. The high probability of correlated observations suggests that the correlation coefficients should be tested at a more stringent level of significance, e.g., $p < .01$. At this level of significance, none of the dependent variables are significantly correlated and as such the assumption of independence of observations has not been violated.

8.4.2.2 Linearity

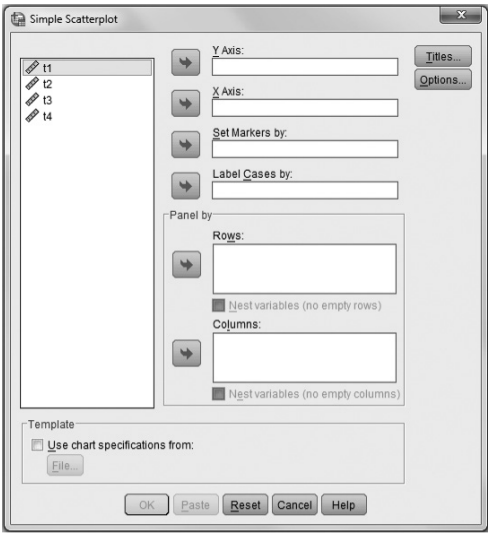
Scatterplots will be used to test the assumption of linearity.

8.4.2.2.1 Windows Method

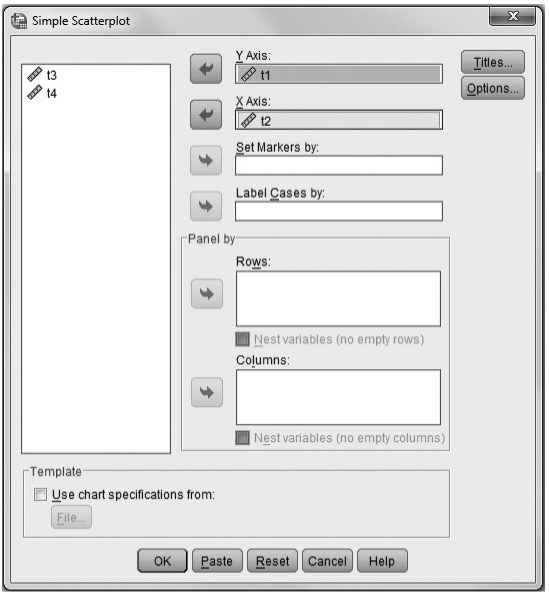
1. From the menu bar, click **Graphs**, then **Legacy Dialogs**, and then **Scatter/Dot…**. The following **Scatter/Dot** window will open. Click (highlight) the Simple Scatter icon.



2. Click Define to open the **Simple Scatterplot** window below.



3. Transfer the **T1** variable to the **Y Axis:** field by clicking (highlight) the variable and then clicking . Transfer the **T2** variable to the **X Axis:** field by clicking (highlight) the variable and then clicking .



4. Click OK to complete the analysis. To obtain scatterplots between T1 and T3, T1 and T4, T2 and T3, T2 and T4, and T3 and T4, repeat steps 1 through 4 for each pair of variables. See Figure 8.1 for the results.

8.4.2.2.2 SPSS Syntax Method

```
GRAPH
/SCATTERPLOT(BIVAR) = T1 WITH T2.
GRAPH
/SCATTERPLOT(BIVAR) = T1 WITH T3.
GRAPH
/SCATTERPLOT(BIVAR) = T1 WITH T4.
GRAPH
/SCATTERPLOT(BIVAR) = T2 WITH T3.
GRAPH
/SCATTERPLOT(BIVAR) = T2 WITH T4.
GRAPH
/SCATTERPLOT(BIVAR) = T3 WITH T4.
```

8.4.2.2.3 Scatterplots

The scatterplots indicate a linear relationship only between the variables T2 and T3. Thus, the assumption of linearity between pairs of dependent

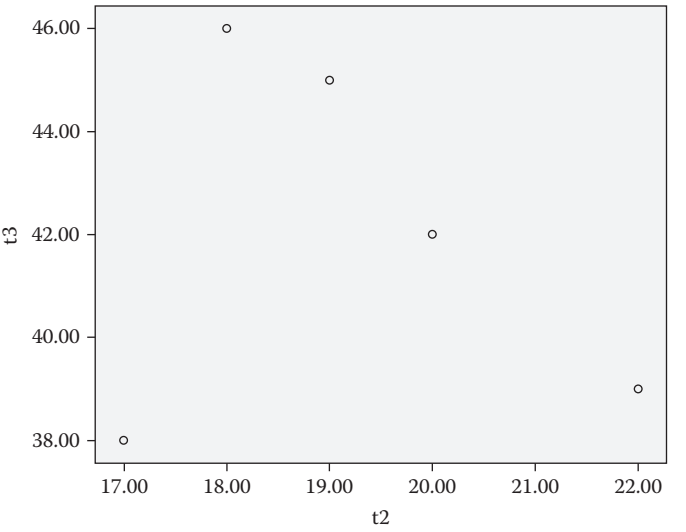


FIGURE 8.1
Scatterplot between T2 and T3.

variables is violated. This is not unexpected given the very small sample size ($N = 5$) employed in this example.

8.4.2.3 Homogeneity of Covariance Matrices

This homogeneity assumption is tested with Box's *M* test; it tests the hypothesis that the covariance matrices of the dependent variables are significantly different across levels of the independent variable. This test is very sensitive so a significance level of .001 is recommended. In the present case, the homogeneity of covariance matrices assumption cannot be tested as there is no between-groups variable.

8.4.2.4 Normality

As there are no specific tests for multivariate normality, univariate normality for the four variables will be tested instead. For the present example, normality will be tested using the normal Q–Q plot and the *z* test for skewness.

8.4.2.4.1 Windows Method

With the data set opened, follow step 1 through step 5 in Section 6.4.2.1.1 in Chapter 6. Treat **T1, T2, T3,** and **T4** as the dependent variables. See Table 8.2 for the results.

8.4.2.4.2 SPSS Syntax Method

```
EXAMINE VARIABLES = T1 T2 T3 T4
/PLOT BOXPLOT NPPLOT
/COMPARE GROUPS
```

```
/STATISTICS DESCRIPTIVES
/CINTERVAL 95
/MISSING LISTWISE
/NOTOTAL.
```

8.4.2.4.3 SPSS Output

TABLE 8.2

Explore Analysis (Selected) Output

Descriptives

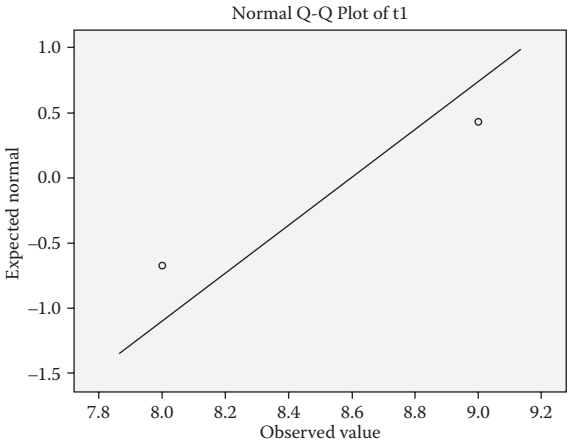
			Statistic	Std. Error
t1	Mean		8.6000	.24495
	95% confidence interval for mean	Lower bound	7.9199	
		Upper bound	9.2801	
	5% trimmed mean		8.6111	
	Median		9.0000	
	Variance		.300	
	Std. deviation		.54772	
	Minimum		8.00	
	Maximum		9.00	
	Range		1.00	
	Interquartile range		1.00	
	Skewness		−.609	.913
	Kurtosis		−3.333	2.000
t2	Mean		19.2000	.86023
	95% confidence interval for mean	Lower bound	16.8116	
		Upper bound	21.5884	
	5% trimmed mean		19.1667	
	Median		19.0000	
	Variance		3.700	
	Std. deviation		1.92354	
	Minimum		17.00	
	Maximum		22.00	
	Range		5.00	
	Interquartile range		3.50	
	Skewness		.590	.913
	Kurtosis		−.022	2.000
t3	Mean		42.0000	1.58114
	95% confidence interval for mean	Lower bound	37.6101	
		Upper bound	46.3899	
	5% trimmed mean		42.0000	
	Median		42.0000	
	Variance		12.500	
	Std. deviation		3.53553	
	Minimum		38.00	
	Maximum		46.00	
	Range		8.00	
	Interquartile range		7.00	
	Skewness		.000	.913
	Kurtosis		−2.608	2.000

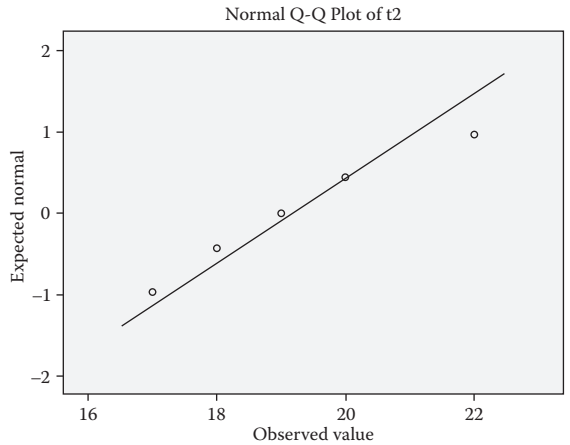
TABLE 8.2 (*Continued*)

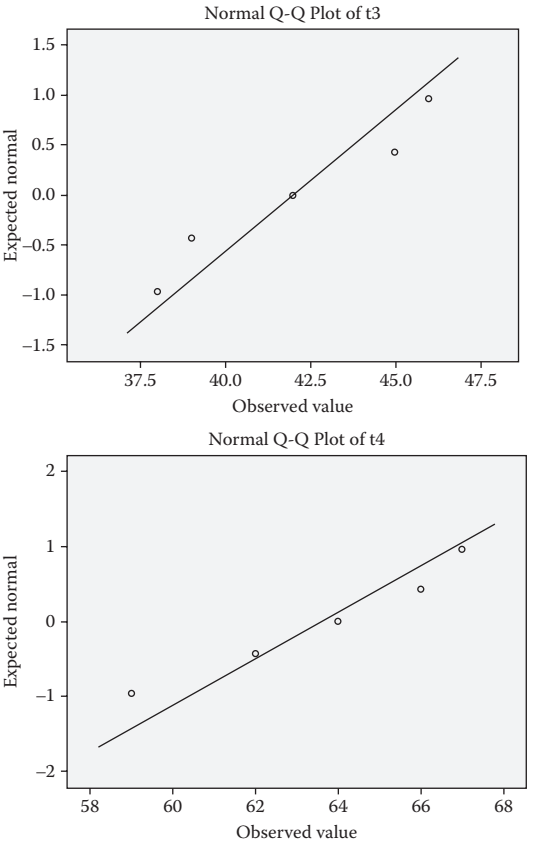
Explore Analysis (Selected) Output

Descriptives

			Statistic	Std. Error
t4	Mean		63.6000	1.43527
	95% confidence interval for mean	Lower bound	59.6151	
		Upper bound	67.5849	
	5% trimmed mean		63.6667	
	Median		64.0000	
	Variance		10.300	
	Std. deviation		3.20936	
	Minimum		59.00	
	Maximum		67.00	
	Range		8.00	
	Interquartile range		6.00	
	Skewness		−.608	.913
	Kurtosis		−.681	2.000







8.4.2.4.4 Interpretation

A simple diagnostic test for normality is based on the skewness and kurtosis values. The statistical *z* value for the skewness value is calculated as:

$$Z_{\text{skewness}} = \frac{\text{skewness}}{\sqrt{\text{s.e. skewness}}}$$

The statistical *z* value for the kurtosis value is calculated as:

$$Z_{\text{kurtosis}} = \frac{\text{kurtosis}}{\sqrt{\text{s.e. kurtosis}}}$$

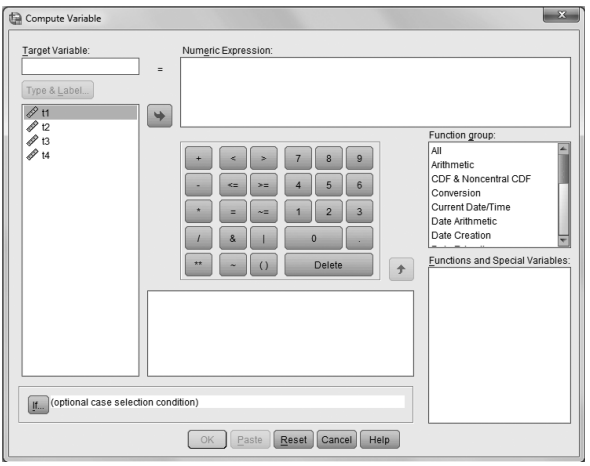
If the calculated *z* value exceeds the specified critical probability value, then the distribution is nonnormal. For example, a calculated *z* value exceeding ±2.58 will result in a rejection of the assumption of normality at the 0.01 critical probability (alpha) level. A calculated *z* value exceeding ±1.96 will result in a rejection of the assumption of normality at the .05 alpha level. Based on the obtained skewness statistics, the *z* values for the variables of

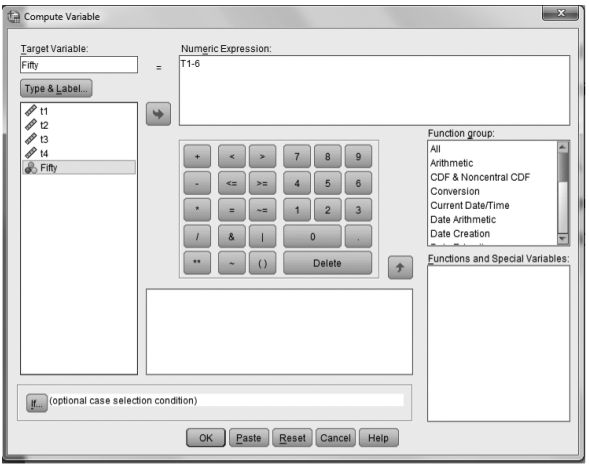
T1, **T2**, **T3**, and **T4** are −0.637, 0.617, 0, and −0.636, respectively, which are less than ± 2.58 ($p > .01$). Therefore, it can be concluded that the distributions of these variables do not vary significantly from normality.

Another diagnostic test for normality is a visual check of the **Normal Q–Q Plot**, which compares the cumulative distribution of the observed values with the expected values derived from the normal distribution. The normal distribution forms a straight diagonal line, and if a variable's distribution is normal, the data distribution will fall more or less along the diagonal. Inspection of the normal Q-Q plots shows very little departure from normality for the four variables.

8.4.3 Windows Method: GLM Multivariate Analysis

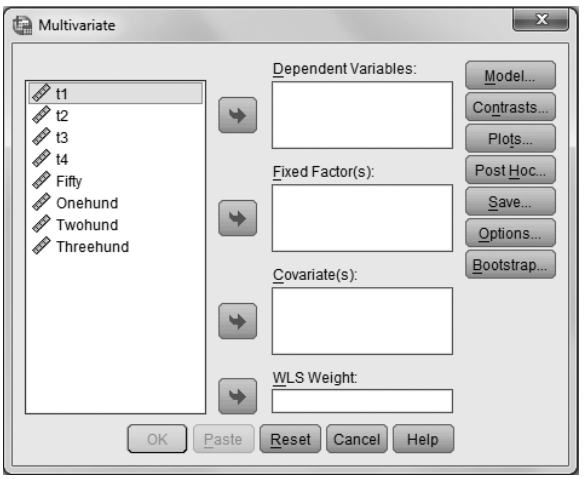
1. In order to test the hypothesis that a set of means is equal to 0, the normative values must first be subtracted from the observed scores. From the menu bar, click **Transform**, and then **Compute Variable**. The following **Compute Variable** window will open.
2. To create the first variable representing the difference between the recorded time for the 50-yard dash and its normative time of 6 seconds, type the name of the variable **Fifty** in the **Target Variable:** field. In the **Numeric Expression:** field, type the expression **T1-6**. Next, click OK. This will create a new variable called **Fifty**, which represents the difference between the observed score (T1: time in seconds to run 50-yard dash) and its normative value (6 seconds).
3. To create the other three variables representing the 100-yard dash (**ONEHUND**), the 200-yard dash (**TWOHUND**), and the 300-yard dash (**THREEHUN**), follow the steps presented in step 2 above. Note that for each of the three computed variables (ONEHUND, TWOHUND, THREEHUN), its normative value has been subtracted



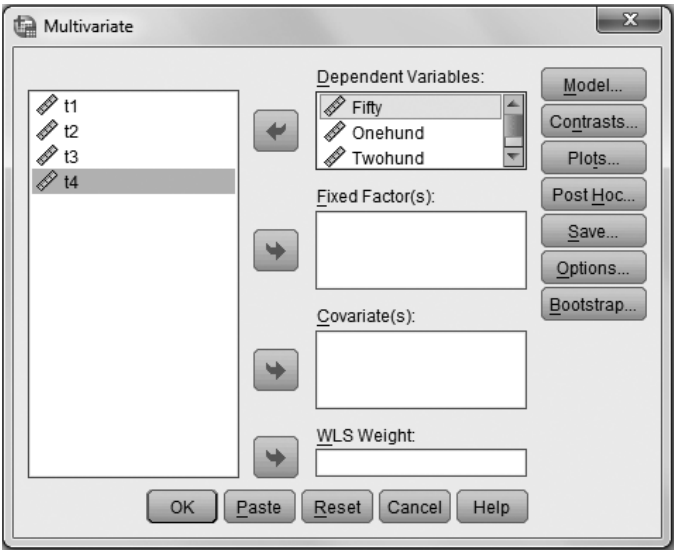


from its observed score (i.e., T2 – 12, T3 – 25, T4 – 40). Successful completion of this procedure will create the four new variables FIFTY, ONEHUND, TWOHUND, and THREEHUN.

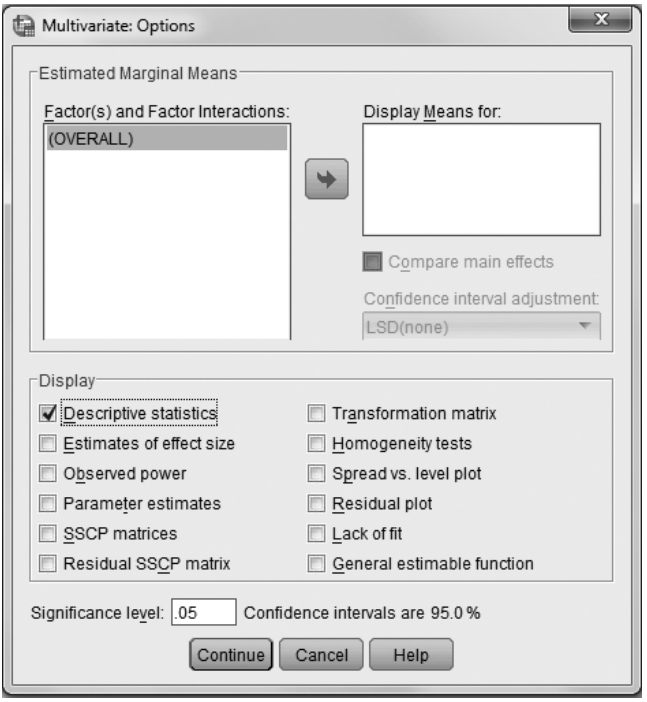
4. The next step is to run the one-sample **General Linear Model** (**GLM**) **Multivariate Analysis** to test the hypothesis that the means of the four computed variables (FIFTY, ONEHUND, TWOHUND, and THREEHUN) do not differ from a set of constants. From the menu bar, click **Analyze**, then **General Linear Model**, and then **Multivariate**. The following **Multivariate** window will open.



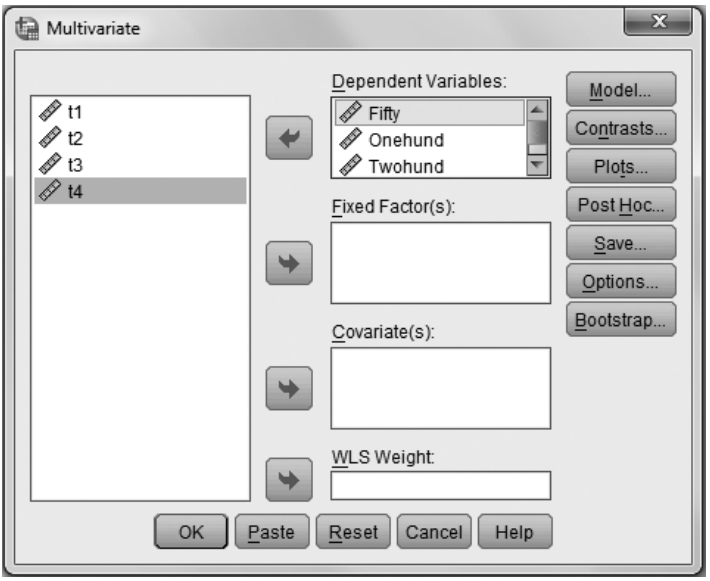
5. Transfer the four variables FIFTY, ONEHUND, TWOHUND, and THREEHUN to the **Dependent Variables:** field by clicking (highlight) these four variables and then clicking ▶.



6. To obtain descriptive statistics for these four variables, click Options... . This will open the **Multivariate: Options** window. Check the **Descriptive statistics** cell and then click Continue.



7. This will open the **Multivariate** window. Click OK to complete the analysis. See Table 8.3 for the results.



8.4.4 SPSS Syntax Method

```
COMPUTE FIFTY = (T1-6).
COMPUTE ONEHUND = (T2-12).
COMPUTE TWOHUND = (T3-25).
COMPUTE THREEHUN = (T4-40).

GLM FIFTY ONEHUND TWOHUND THREEHUN
/PRINT = DESCRIPTIVES.
```

8.4.5 SPSS Output

TABLE 8.3

One-Sample GLM Output

Descriptive Statistics

	Mean	Std. Deviation	N
FIFTY	2.6000	.5477	5
ONEHUND	7.2000	1.9235	5
TWOHUND	17.0000	3.5355	5
THREEHUN	23.6000	3.2094	5

TABLE 8.3 (*Continued*)

One-Sample GLM Output

Multivariate Tests[b]

Effect		Value	*F*	Hypothesis d*f*	Error d*f*	Sig.
Intercept	Pillai's Trace	.999	232.563[a]	4.000	1.000	.049
	Wilks' Lambda	.001	232.563[a]	4.000	1.000	.049
	Hotelling's Trace	930.250	232.563[a]	4.000	1.000	.049
	Roy's Largest Root	930.250	232.563[a]	4.000	1.000	.049

[a] Exact statistic
[b] Design: Intercept

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	d*f*	Mean Square	*F*	Sig.
Corrected Model	FIFTY	.000[a]	0	.	.	.
	ONEHUND	.000[a]	0	.	.	.
	TWOHUND	.000[a]	0	.	.	.
	THREEHUN	.000[a]	0	.	.	.
Intercept	FIFTY	33.800	1	33.800	112.667	.000
	ONEHUND	259.200	1	259.200	70.054	.001
	TWOHUND	1445.000	1	1445.000	115.600	.000
	THREEHUN	2784.800	1	2784.800	270.369	.000

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	d*f*	Mean Square	*F*	Sig.
Error	FIFTY	1.200	4	.300		
	ONEHUND	14.800	4	3.700		
	TWOHUND	50.000	4	12.500		
	THREEHUN	41.200	4	10.300		
Total	FIFTY	35.000	5			
	ONEHUND	274.000	5			
	TWOHUND	1495.000	5			
	THREEHUN	2826.000	5			
Corrected Total	FIFTY	1.200	4			
	ONEHUND	14.800	4			
	TWOHUND	50.000	4			
	THREEHUN	41.200	4			

[a] R Squared = .000 (Adjusted R Squared = .000)

8.4.6 Results and Interpretation

The **Descriptive Statistics** table presents means and standard deviations for the four running events after their normative values have been subtracted. The sample exceeds (i.e., is poorer than) the norm for all four running events.

The **Multivariate Tests** table tests the hypothesis that the four sample means do not differ from the specified set of constants. Of the three multivariate tests (Pillai's, Hotelling's, Wilks', Roy's), Pillai's trace is the most powerful (i.e., the ability to detect differences if they exist) and the most robust (i.e., the significance level based on it is reasonably correct even when the assumptions are violated). Since the observed significance level is small ($p < .05$), the null hypothesis that the sample means do not differ from the specified constants is rejected, multivariate Pillai $F(1,4) = 232.56$, $p < .05$.

When the null hypothesis of no difference is rejected, it is often informative to examine the univariate **Tests of Between-Subjects Effects** to identify which variables yielded significant differences. The univariate test results (**Intercept**) show that the means for all running events differ significantly from their specified standards ($p < .01$).

8.5 Example 2: GLM Multivariate Analysis: Two-Sample Test

Following from the previous example, suppose that the researcher also recorded running times for the four events for five women. The running times for the two samples for the four events are shown below. The hypothesis that men and women do not differ on the four running events will be tested.

	Running Events			
	50 (Yard)	**100 (Yard)**	**200 (Yard)**	**300 (Yard)**
Men				
s1	9	20	42	67
s2	9	19	45	66
s3	8	17	38	62
s4	9	18	46	64
s5	8	22	39	59
Women				
s1	15	30	41	77
s2	14	29	44	76
s3	13	27	39	72
s4	14	28	56	74
s5	13	32	49	69

8.5.1 Data Entry Format

The data set has been saved under the name **EX8b.SAV**.

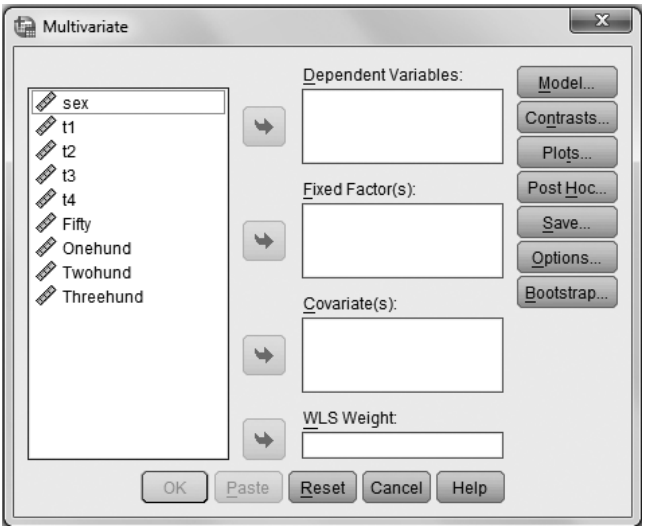
Variables	Column(s)	Code
Sex	1	1 = men, 2 = women
T1	2	Running speed in seconds
T2	3	Running speed in seconds
T3	4	Running speed in seconds
T4	5	Running speed in seconds

8.5.2 Testing Assumptions

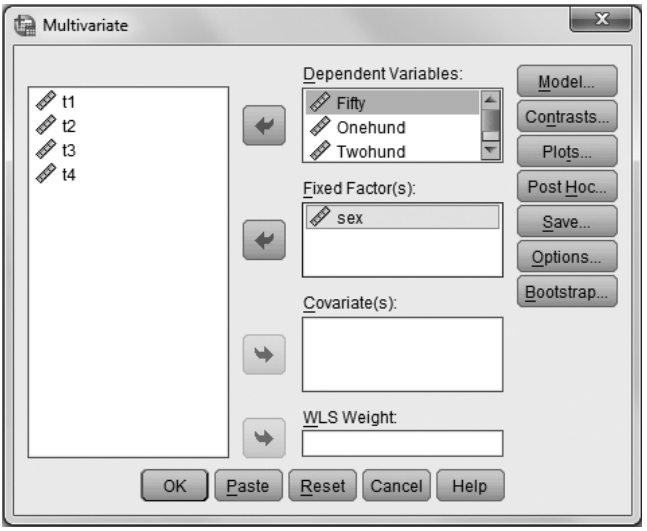
To test the assumptions of (1) independence of observations, (2) linearity, (3) homogeneity of covariance matrices, and (4) normality, follow the steps given in Section 8.4.2.

8.5.3 Windows Method: GLM Multivariate Analysis: Two-Sample Test

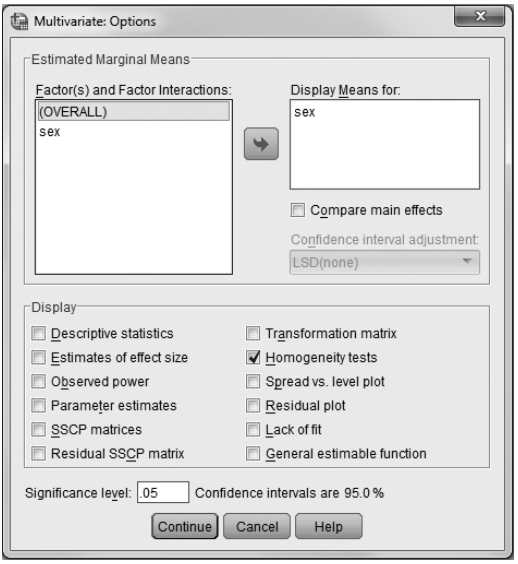
1. To create the four variables **FIFTY**, **ONEHUND**, **TWOHUND**, and **THREEHUN**, repeat steps 1 to 3 in Section 8.4.3.
2. From the menu bar, click **Analyze**, then **General Linear Model**, and then **Multivariate**. The following **Multivariate** window will open.



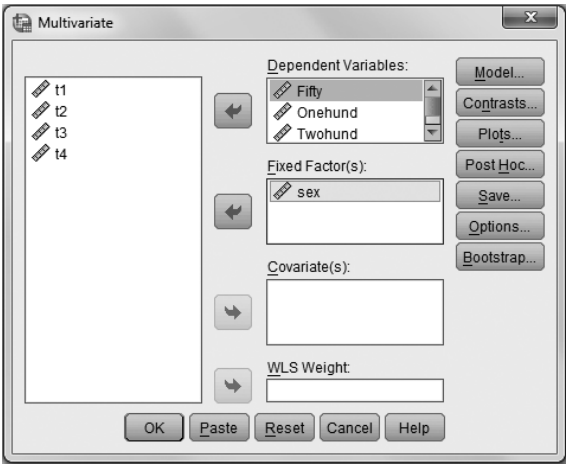
3. Transfer the four variables FIFTY, ONEHUND, TWOHUND, and THREEHUN to the **Dependent Variables:** field by clicking (highlight) these four variables and then clicking ▶. Transfer the **SEX** variable to the **Fixed Factor(s):** field by clicking (highlight) this variable and clicking ▶.



4. To obtain descriptive statistics for the four dependent variables for males and females, click Options... . When the **Multivariate: Options** window opens, click (highlight) **SEX** in the **Factor(s) and Factor Interactions:** field and then click ▶ to transfer this factor to the **Display Means for:** field. To test for the assumption of homogeneity of covariance matrices, check the **Homogeneity test** cell. Click Continue to return to the **Multivariate** window.



5. When the **Multivariate** window opens, click OK to complete the analysis. See Table 8.4 for the results.



8.5.4 SPSS Syntax Method

```
COMPUTE FIFTY = (T1-6).
COMPUTE ONEHUND = (T2-12).
COMPUTE TWOHUND = (T3-25).
COMPUTE THREEHUN = (T4-40).

GLM FIFTY ONEHUND TWOHUND THREEHUN BY SEX
/EMMEANS = TABLES(SEX)
/PRINT HOMOGENEITY.
```

8.5.5 SPSS Output

TABLE 8.4

Two-Sample GLM Output

General Linear Model

Between-Subject Factors

		Value Label	*N*
sex	1.00	male	5
	2.00	female	5

Box's Test of Equality of Covariance Matrices[a]

Box's M	18.722
F	.805
d*f*1	10
d*f*2	305.976
Sig.	.624

(*Continued*)

TABLE 8.4 (*Continued*)

Two-Sample GLM Output

Multivariate Tests[b]

Effect		Value	*F*	Hypothesis d*f*	Error d*f*	Sig.
Intercept	Pillai's Trace	.997	358.769[a]	4.000	5.000	.000
	Wilks' Lambda	.003	358.769[a]	4.000	5.000	.000
	Hotelling's Trace	287.015	358.769[a]	4.000	5.000	.000
	Roy's Largest Root	287.015	358.769[a]	4.000	5.000	.000

Multivariate Tests[b]

Effect		Value	*F*	Hypothesis d*f*	Error d*f*	Sig.
sex	Pillai's Trace	.971	41.742[a]	4.000	5.000	.000
	Wilks' Lambda	.029	41.742[a]	4.000	5.000	.000
	Hotelling's Trace	33.394	41.742[a]	4.000	5.000	.000
	Roy's Largest Root	33.394	41.742[a]	4.000	5.000	.000

Levene's Test of Equality of Error Variances[a]

	F	d*f*1	d*f*2	Sig.
FIFTY	.640	1	8	.447
ONEHUND	.000	1	8	1.000
TWOHUND	2.429	1	8	.158
THREEHUN	.000	1	8	1.000

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	d*f*	Mean Square	*F*	Sig.
Corrected Model	FIFTY	67.600[a]	1	67.600	135.200	.000
	ONEHUND	250.000[b]	1	250.000	67.568	.000
	TWOHUND	36.100[c]	1	36.100	1.220	.302
	THREEHUN	250.000[d]	1	250.000	24.272	.001
Intercept	FIFTY	270.400	1	270.400	540.800	.000
	ONEHUND	1488.400	1	1488.400	402.270	.000
	TWOHUND	3572.100	1	3572.100	120.679	.000
	THREEHUN	8179.600	1	8179.600	794.136	.000
Sex	FIFTY	67.600	1	67.600	135.200	.000
	ONEHUND	250.000	1	250.000	67.568	.000
	TWOHUND	36.100	1	36.100	1.220	.302
	THREEHUN	250.000	1	250.000	24.272	.001

TABLE 8.4 (*Continued*)

Two-Sample GLM Output

Multivariate Tests[b]

Error	FIFTY	4.000	8	.500
	ONEHUND	29.600	8	3.700
	TWOHUND	236.800	8	29.600
	THREEHUN	82.400	8	10.300
Total	FIFTY	342.000	10	
	ONEHUND	1768.000	10	
	TWOHUND	3845.000	10	
	THREEHUN	8512.000	10	
Corrected Total	FIFTY	71.600	9	
	ONEHUND	279.600	9	
	TWOHUND	272.600	9	
	THREEHUN	332.400	9	

Estimated Marginal Means

Sex

Dependent Variable	Sex	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
FIFTY	male	2.600	.316	1.871	3.329
	female	7.800	.316	7.071	8.529
ONEHUND	male	7.200	.860	5.216	9.184
	female	17.200	.860	15.216	19.184
TWOHUND	male	17.000	2.433	11.389	22.611
	female	20.800	2.433	15.189	26.411
THREEHUN	male	23.600	1.435	20.290	26.910
	female	33.600	1.435	30.290	36.910

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

[a] Design: Intercept + sex

[b] Exact statistic

[c] Design: Intercept + sex

[d] Design: Intercept + sex

[e] R Squared = .944 (Adjusted R Squared = .937)

[f] R Squared = .894 (Adjusted R Squared = .881)

[g] R Squared = .132 (Adjusted R Squared = .024)

[h] R Squared = .752 (Adjusted R Squared = .721)

8.5.6 Results and Interpretation

Box's Test of Equality of Covariance Matrices (Box's *M* test) tests the assumption of homogeneity of covariance matrices. As the test is not significant, this homogeneity assumption has not been violated, $F(10,305.976) = .805$, $p > .001$.

Levene's Test of Equality of Error Variances presents the univariate tests for homogeneity of variance for each of the four dependent variables FIFTY, ONEHUND, TWOHUND, and THREEHUN. All four *F* tests are not significant indicating that the assumption of homogeneity of variance for all four dependent variables has not been violated ($p > .001$).

The **Estimated Marginal Means** table presents the mean running times for men and women, for each of the four running events after the normative values have been subtracted. Both samples exceed the norms for all four running events.

The **Multivariate Tests** table (Pillai's, Hotelling's, Wilks', Roy's) tests the hypothesis that men and women do not differ significantly on overall running time. The significance level is based on the *F* distribution with 4 and 5 degrees of freedom. The observed significance levels for all four multivariate tests are small ($p < .001$), so the null hypothesis that men and women performed similarly on the running events is rejected (e.g., multivariate Pillai $F(4,5) = 41.74$, $p < .001$).

The **Tests of Between-Subjects Effects** table presents the univariate test of sex difference for the individual running events. The obtained *F* values are equivalent to those obtained from the one-way analysis of variance. ***In the case where the comparison is between two groups, the F values are the square of the two-sample t values.*** The results indicate significant sex differences in running times for three of the four running events (FIFTY: male: $M = 2.60$, female: $M = 7.80$; ONEHUND: male: $M = 7.20$, female: $M = 17.20$; THREEHUN: male: $M = 23.60$, female: $M = 33.60$); there is no significant sex difference for the 200-yard dash.

8.6 Example 3: GLM: 2 × 2 × 4 Factorial Design

In Example 2, the researcher used the GLM two-sample test to test for sex differences in running times across four running events. Suppose that in addition to sex differences, the researcher was also interested in whether the subjects' ethnicity (white versus nonwhite) would make a difference in their running times. In particular, the researcher was interested in the interaction between subjects' sex and ethnicity in influencing their running times. The running times for the four groups (men-white, men-nonwhite, women-white, women-nonwhite) for the four events are presented below.

	Running Events			
	50 (Yard)	**100 (Yard)**	**200 (Yard)**	**300 (Yard)**
Men				
White				
s1	9	20	42	67
s2	9	19	45	66
s3	8	17	38	62
s4	9	18	46	64
s5	8	22	39	59
Nonwhite				
s1	6	17	39	64
s2	6	16	42	63
s3	5	14	35	59
s4	6	15	43	61
s5	5	19	36	56
Women				
White				
s1	15	30	41	77
s2	14	29	44	76
s3	13	27	39	72
s4	14	28	56	74
s5	13	32	49	69
Nonwhite				
s1	20	35	46	82
s2	19	34	49	81
s3	18	32	44	77
s4	19	32	61	79
s5	18	37	54	74

8.6.1 Data Entry Format

The data set has been saved under the name **EX8c.SAV**.

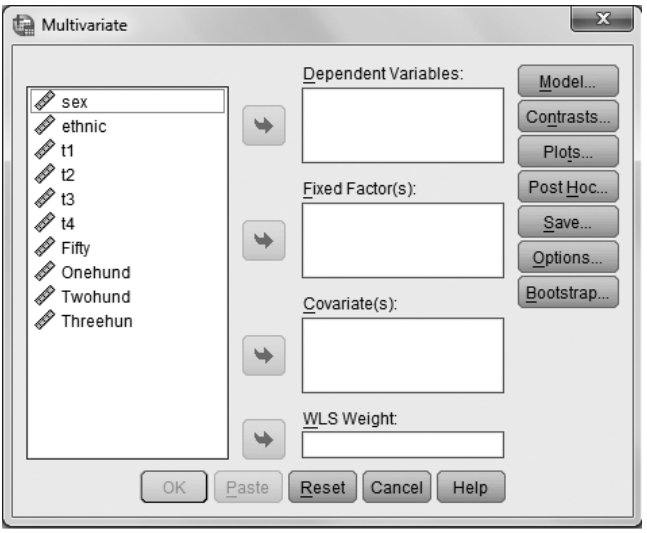
Variables	Column(s)	Code
Sex	1	1 = men, 2 = women
Ethnic	2	1 = white, 2 = nonwhite
T1	3	Running speed in seconds
T2	4	Running speed in seconds
T3	5	Running speed in seconds
T4	6	Running speed in seconds

8.6.2 Testing Assumptions

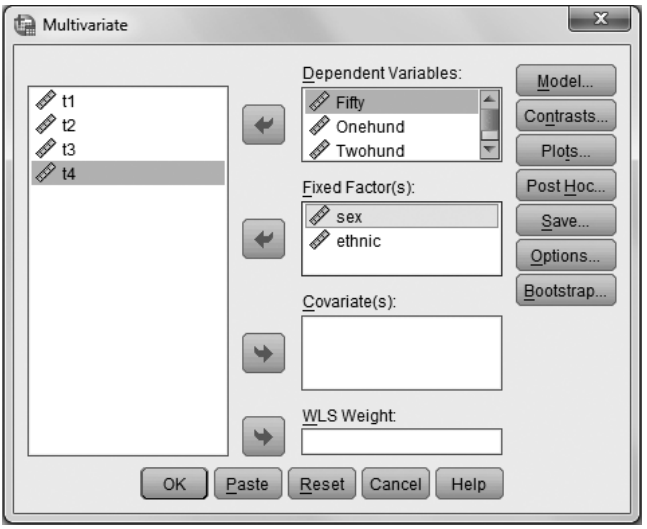
To test the assumptions of (1) independence of observations, (2) linearity, (3) homogeneity of covariance matrices, and (4) normality, follow the steps presented in Section 8.4.2

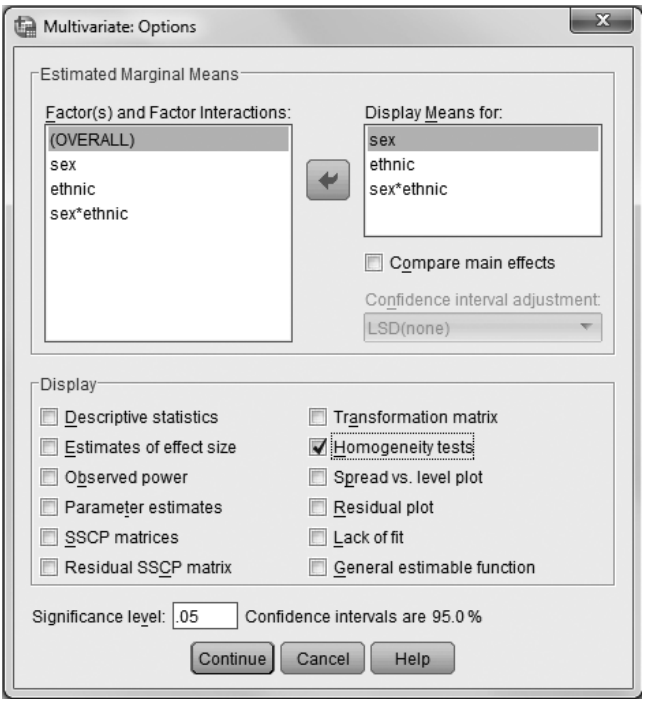
8.6.3 Windows Method: GLM: 2 × 2 × 4 Factorial Design

1. To create the four variables **FIFTY**, **ONEHUND**, **TWOHUND**, and **THREEHUN**, repeat steps 1 to 3 in Section 8.4.3.
2. From the menu bar, click **Analyze**, then **General Linear Model**, and then **Multivariate**. The following **Multivariate** window will open.

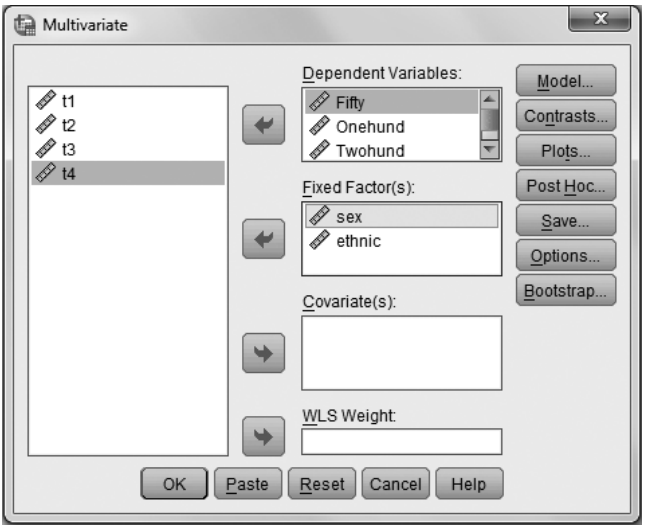


3. Transfer the four variables FIFTY, ONEHUND, TWOHUND, and THREEHUN to the **Dependent Variables:** field by clicking (highlight) these four variables and then clicking ▶. Transfer the **SEX** and **ETHNIC** variables to the **Fixed Factor(s):** field by clicking (highlight) these variables and then clicking ▶.
4. To obtain the estimated marginal means for the four dependent variables for the four experimental groups (men-white, men-nonwhite, women-white, women-nonwhite), click Options... . When the **Multivariate: Options** window opens, click (highlight) **SEX**, **ETHNIC**, and **SEX*ETHNIC** in the **Factor(s) and Factor Interactions:** field, and then click ▶ to transfer these factors and factor interaction to the **Display Means for:** field. To test for the assumption of homogeneity of covariance matrices, check the **Homogeneity tests** cell. Click Continue to return to the **Multivariate** window.





5. When the **Multivariate** window opens, click OK to complete the analysis. See Table 8.5 for the results.



8.6.4 SPSS Syntax Method

```
COMPUTE FIFTY = (T1-6).
COMPUTE ONEHUND = (T2-12).
COMPUTE TWOHUND = (T3-25).
COMPUTE THREEHUN = (T4-40).

GLM FIFTY ONEHUND TWOHUND THREEHUN BY SEX ETHNIC
/PRINT HOMOGENEITY
/EMMEANS = TABLES(SEX)
/EMMEANS = TABLES(ETHNIC)
/EMMEANS = TABLES(SEX*ETHNIC).
```

8.6.5 SPSS Output

TABLE 8.5

$2 \times 2 \times 4$ GLM Output

General Linear Model			
Between-Subject Factors			
		Value Label	***N***
sex	1.00	male	10
	2.00	female	10
ethnic	1.00	WHITE	10
	2.00	NON-WHITE	10

TABLE 8.5 (*Continued*)

2 × 2 × 4 GLM Output

Box's Test of Equality of Covariance Matrices[a]

Box's M	37.361
F	.634
df1	30
df2	703.847
Sig.	.937

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.

[a] Design: Intercept + sex + ethnic + sex * ethnic

Multivariate Tests[b]

Effect		Value	*F*	Hypothesis d*f*	Error d*f*	Sig.
Intercept	Pillai's Trace	.996	890.965[a]	4.000	13.000	.000
	Wilks' Lambda	.004	890.965[a]	4.000	13.000	.000
	Hotelling's Trace	274.143	890.965[a]	4.000	13.000	.000
	Roy's Largest Root	274.143	890.965[a]	4.000	13.000	.000
sex	Pillai's Trace	.992	392.434[a]	4.000	13.000	.000
	Wilks' Lambda	.008	392.434[a]	4.000	13.000	.000
	Hotelling's Trace	120.749	392.434[a]	4.000	13.000	.000
	Roy's Largest Root	120.749	392.434[a]	4.000	13.000	.000
ethnic	Pillai's Trace	.658	6.254[a]	4.000	13.000	.005
	Wilks' Lambda	.342	6.254[a]	4.000	13.000	.005
	Hotelling's Trace	1.924	6.254[a]	4.000	13.000	.005
	Roy's Largest Root	1.924	6.254[a]	4.000	13.000	.005
sex * ethnic	Pillai's Trace	.968	99.122[a]	4.000	13.000	.000
	Wilks' Lambda	.032	99.122[a]	4.000	13.000	.000
	Hotelling's Trace	30.499	99.122[a]	4.000	13.000	.000
	Roy's Largest Root	30.499	99.122[a]	4.000	13.000	.000

[a] Exact statistic

[b] Design: Intercept+sex + ethnic + sex * ethnic

Levene's Test of Equality of Error Variances[a]

	F	d*f*1	d*f*2	Sig.
FIFTY	.427	3	16	.737
ONEHUND	.028	3	16	.994
TWOHUND	1.620	3	16	.224
THREEHUN	.000	3	16	1.000

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

[a] Design: Intercept + sex + ethnic + sex * ethnic

TABLE 8.5 (*Continued*)

2 × 2 × 4 GLM Output

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	d*f*	Mean Square	*F*	Sig.
Corrected Model	FIFTY	508.200[a]	3	169.400	338.800	.000
	ONEHUND	1046.150[b]	3	348.717	89.415	.000
	TWOHUND	389.200[c]	3	129.733	4.383	.020
	THREEHUN	1065.000[d]	3	355.000	34.466	.000
Intercept	FIFTY	649.800	1	649.800	1299.600	.000
	ONEHUND	3200.450	1	3200.450	820.628	.000
	TWOHUND	7527.200	1	7527.200	254.297	.000
	THREEHUN	16936.200	1	16936.200	1644.291	.000
sex	FIFTY	423.200	1	423.200	846.400	.000
	ONEHUND	966.050	1	966.050	247.705	.000
	TWOHUND	304.200	1	304.200	10.277	.006
	THREEHUN	980.000	1	980.000	95.146	.000
ethnic	FIFTY	5.000	1	5.000	10.000	.006
	ONEHUND	4.050	1	4.050	1.038	.323
	TWOHUND	5.000	1	5.000	.169	.687
	THREEHUN	5.000	1	5.000	.485	.496
sex * ethnic	FIFTY	80.000	1	80.000	160.000	.000
	ONEHUND	76.050	1	76.050	19.500	.000
	TWOHUND	80.000	1	80.000	2.703	.120
	THREEHUN	80.000	1	80.000	7.767	.013
Error	FIFTY	8.000	16	.500		
	ONEHUND	62.400	16	3.900		
	TWOHUND	473.600	16	29.600		
	THREEHUN	164.800	16	10.300		
Total	FIFTY	1166.000	20			
	ONEHUND	4309.000	20			
	TWOHUND	8390.000	20			
	THREEHUN	18166.000	20			
Corrected Total	FIFTY	516.200	19			
	ONEHUND	1108.550	19			
	TWOHUND	862.800	19			
	THREEHUN	1229.800	19			

[a] R Squared = .985 (Adjusted R Squared = .982)
[b] R Squared = .944 (Adjusted R Squared = .933)
[c] R Squared = .451 (Adjusted R Squared = .348)
[d] R Squared = .866 (Adjusted R Squared = .841)

TABLE 8.5 (*Continued*)

2 × 2 × 4 GLM Output

Estimated Marginal Means

1. Sex

Dependent Variable	Sex	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
FIFTY	male	1.100	.224	.626	1.574
	female	10.300	.224	9.826	10.774
ONEHUND	male	5.700	.624	4.376	7.024
	female	19.600	.624	18.276	20.924
TWOHUND	male	15.500	1.720	11.853	19.147
	female	23.300	1.720	19.653	26.947
THREEHUN	male	22.100	1.015	19.949	24.251
	female	36.100	1.015	33.949	38.251

2. Ethnic

Dependent Variable	Ethnic	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
FIFTY	WHITE	5.200	.224	4.726	5.674
	NON-WHITE	6.200	.224	5.726	6.674
ONEHUND	WHITE	12.200	.624	10.876	13.524
	NON-WHITE	13.100	.624	11.776	14.424
TWOHUND	WHITE	18.900	1.720	15.253	22.547
	NON-WHITE	19.900	1.720	16.253	23.547
THREEHUN	WHITE	28.600	1.015	26.449	30.751
	NON-WHITE	29.600	1.015	27.449	31.751

3. Sex * Ethnic

Dependent Variable	Sex	Ethnic	Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
FIFTY	male	WHITE	2.600	.316	1.930	3.270
		NON-WHITE	−.400	.316	−1.070	.270
	female	WHITE	7.800	.316	7.130	8.470
		NON-WHITE	12.800	.316	12.130	13.470
ONEHUND	male	WHITE	7.200	.883	5.328	9.072
		NON-WHITE	4.200	.883	2.328	6.072
	female	WHITE	17.200	.883	15.328	19.072
		NON-WHITE	22.000	.883	20.128	23.872
TWOHUND	male	WHITE	17.000	2.433	11.842	22.158
		NON-WHITE	14.000	2.433	8.842	19.158
	female	WHITE	20.800	2.433	15.642	25.958
		NON-WHITE	25.800	2.433	20.642	30.958
THREEHUN	male	WHITE	23.600	1.435	20.557	26.643
		NON-WHITE	20.600	1.435	17.557	23.643
	female	WHITE	33.600	1.435	30.557	36.643
		NON-WHITE	38.600	1.435	35.557	41.643

8.6.6 Results and Interpretation

Box's Test of Equality of Covariance Matrices (Box's *M* test) is not significant, and therefore it can be concluded that the homogeneity assumption has not been violated, $F(30,703.847) = .634$, $p > .001$.

Levene's Test of Equality of Error Variances presents the univariate tests for homogeneity of variance for each of the four dependent variables FIFTY, ONEHUND, TWOHUND, and THREEHUN. All four *F* tests are not significant indicating that the assumption of homogeneity of variance for all four dependent variables has not been violated ($p > .001$)

The **Estimated Marginal Means** table presents the means of the overall running times for the four running events as a function of subjects' **SEX**, **ETHNICITY**, and **SEX*ETHNICITY** interaction.

The **Multivariate Tests** table presents the multivariate tests of significance for the main effects of the between-groups variables **SEX** and **ETHNICITY**, and the **SEX*ETHNICITY** interaction. For all three effects, the observed significance levels for the four multivariate tests (Pillai's, Wilks', Hotelling's, Roy's) are small. Hence, their associated null hypotheses (no sex difference, no ethnicity difference, no sex*ethnicity interaction) are rejected. The results of the multivariate tests have been converted to approximate *F* values and can be interpreted in the same way as *F* values from one-way ANOVA are interpreted. The multivariate tests for both **SEX** (multivariate Pillai $F(4,13) = 392.43$, $p < .001$) and **ETHNICITY** (multivariate Pillai $F(4,13) = 6.25$, $p < .01$) are statistically significant, indicating that men ($M = 11.1$) when compared to women ($M = 22.32$), and whites ($M = 13.18$) when compared to nonwhites ($M = 17.2$) differed significantly in their overall running times.

The **Tests of Between-Subjects Effects** table can be examined for significant **SEX** and **ETHNICITY** differences for each of the four running events. For **SEX**, the results show significant sex differences for all four running events ($p < .01$) (FIFTY: men: $M = 1.10$, women: $M = 10.30$; ONEHUND: men: $M = 5.70$, women: $M = 19.60$; TWOHUND: men: $M = 15.50$, women: $M = 23.30$; THREEHUN: men: $M = 22.10$, women: $M = 36.10$). For **ETHNICITY**, the results show significant ethnicity difference for only the 50-yard dash (white: $M = 5.20$, non-white: $M = 6.20$; $F(1,16) = 10.00$, $p < .01$)

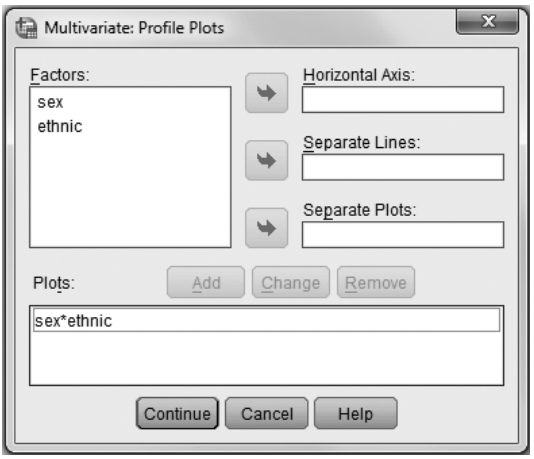
Given that the **SEX*ETHNICITY** interaction is significant, the **Tests of Between-Subjects Effects** table can be examined to see which of the four running events are significantly affected by this interaction. The univariate *F* tests for the **SEX*ETHNICITY** interaction are identical to the *F* ratios generated from the **SEX*ETHNICITY** interaction in a two-way ANOVA. The univariate results show the significant **SEX*ETHNICITY** interaction effect for three of the four running events ($p < .01$). Thus, the subjects' performances on the 50-yard dash, the 100-yard dash, and the 300-yard dash are dependent on the joint effects of their sex and ethnicity.

To interpret the **SEX*ETHNICITY** interactions, it would be useful to graph the means of the four running events (FIFTY, ONEHUND, TWOHUND,

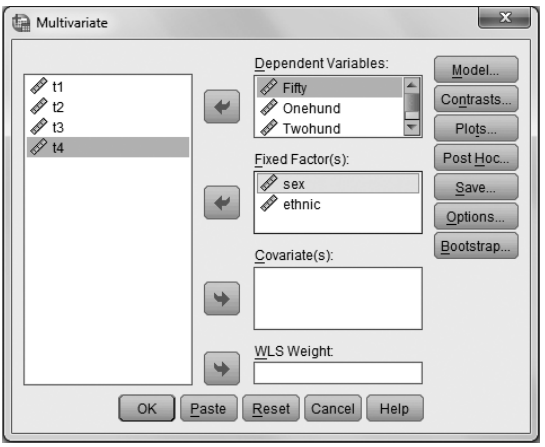
THREEHUN) from the estimated marginal means table. Sections 8.6.7 and 8.6.8 will generate the graphs presented in Figures 8.2 through 8.5.

8.6.7 Windows Method (Profile Plot)

1. Open the **Multivariate** window and click Plots... . The following **Multivariate: Profile Plots** window will open. Transfer the **SEX** variable to the **Horizontal Axis:** field by clicking (highlight) the variable and then clicking ▶. Transfer the **ETHNIC** variable to the **Separate Lines:** field by clicking (highlight) the variable and then clicking ▶. Next, click Add to transfer the **SEX*ETHNIC** interaction to the **Plots:** field. When this is done, click Continue.



2. When the **Multivariate** window opens, click OK to plot the graphs.



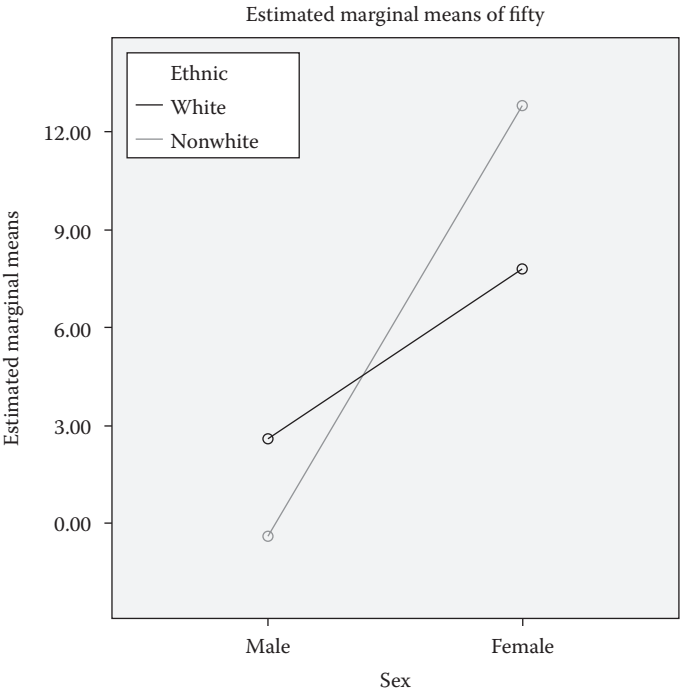


FIGURE 8.2
Sex × Ethnicity interaction for 50-yard dash.

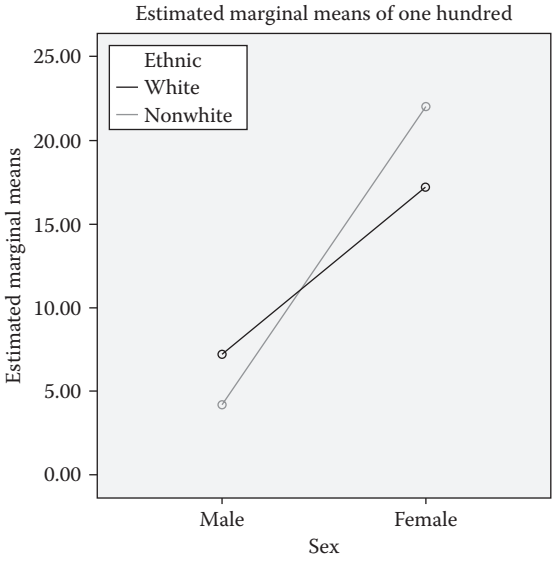


FIGURE 8.3
Sex × Ethnicity interaction for 100-yard dash.

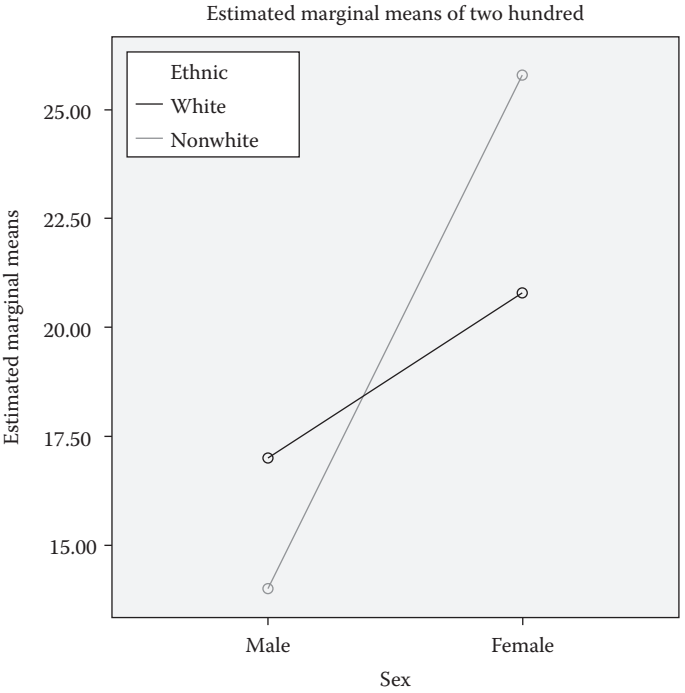


FIGURE 8.4
Sex × Ethnicity interaction for 200-yard dash.

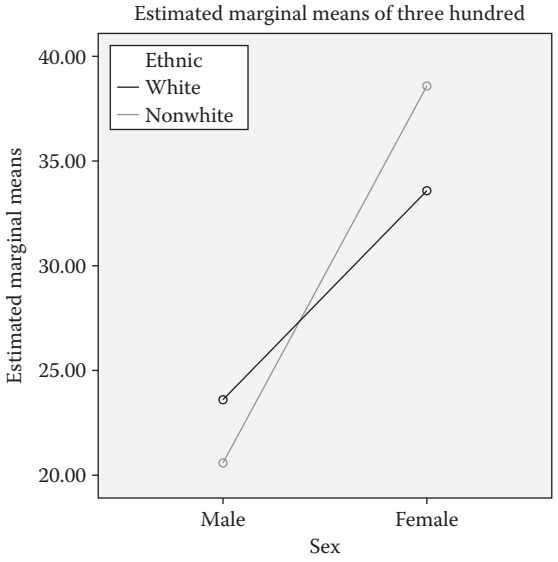


FIGURE 8.5
Sex × Ethnicity interaction for 300-yard dash.

8.6.8 SPSS Syntax Method (Profile Plot)

```
GRAPH
/LINE(MULTIPLE)MEAN(FIFTY) BY SEX BY ETHNIC.
GRAPH
/LINE(MULTIPLE)MEAN(ONEHUND) BY SEX BY ETHNIC.
GRAPH
/LINE(MULTIPLE)MEAN(TWOHUND) BY SEX BY ETHNIC.
GRAPH
/LINE(MULTIPLE)MEAN(THREEHUN) BY SEX BY ETHNIC.
```

8.6.9 Results and Interpretation

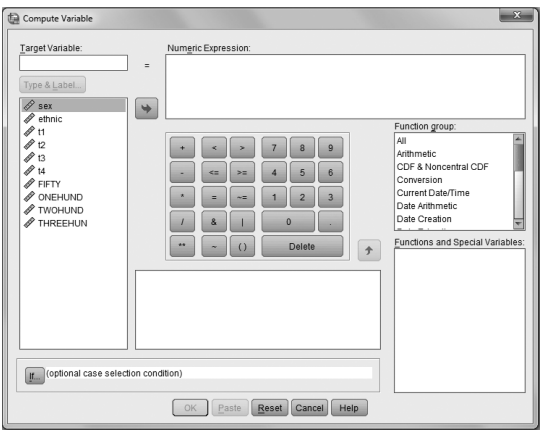
Using the 100-yard dash as an example, the **SEX*ETHNICITY** interaction can be interpreted as follows. The effect of subjects' sex on the running times in the 100-yard dash is dependent on the ethnicity of the subjects, such that for men, nonwhites ran faster than whites; for women, the effect is opposite with whites running faster than nonwhites.

Post hoc comparison analysis can also be used to identify which of the four groups (male-white, male-nonwhite, female-white, female-nonwhite) generated from the **SEX*ETHNICITY** interaction is significantly different from each other in the 100-yard dash. The following **Windows Method** and the **Syntax File Method** will accomplish this analysis.

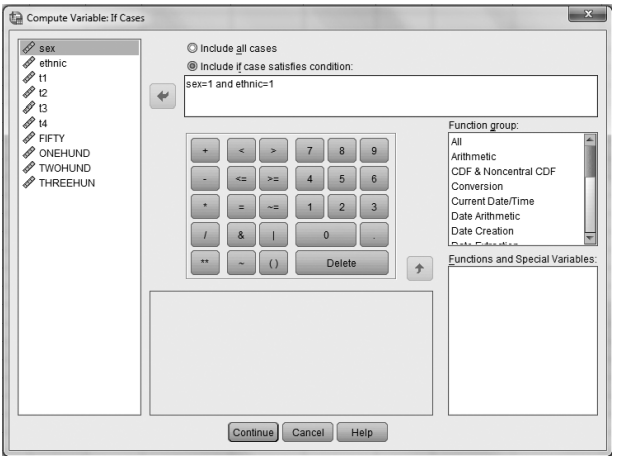
8.6.10 Windows Method (Data Transformation)

8.6.10.1 Data Transformation

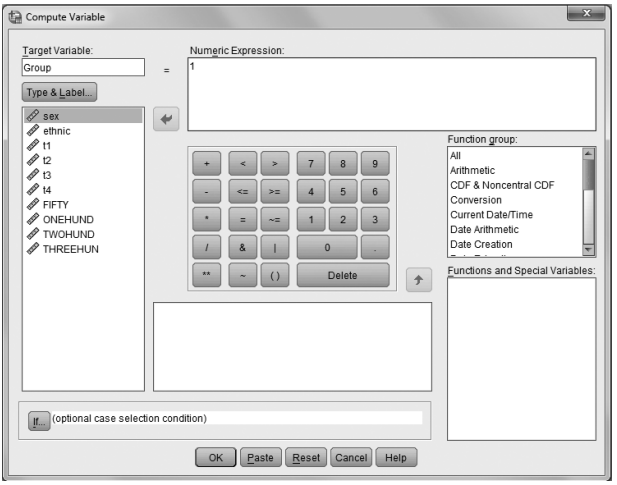
1. The first step is to create a new grouping variable called **GROUP** that contains the four levels (male-white, male-nonwhite, female-white, female-nonwhite) generated by the **SEX*ETHNIC** interaction. From the menu bar, click **Transform** and then **Compute Variable.** The following **Compute Variable** window will open.



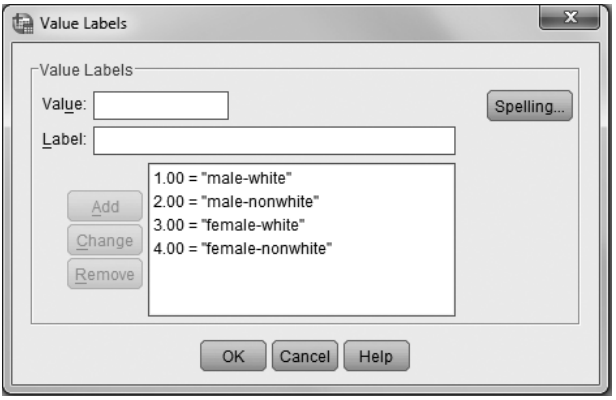
2. Click If... to open the **Compute Variable: If Cases** window below. Ensure that the **Include if case satisfies the condition:** button is checked. To create the first group (**male-white**), type **sex = 1 and ethnic = 1** in the field. Click Continue.



3. Since the **male-white** group is the first of four levels within a new grouping variable called **GROUP**, this level will be coded **1** within the **GROUP** variable. When the next window opens, type **GROUP** in the **Target Variable:** field and **1** in the **Numeric Expression:** field. Click OK to create the first level of **male-white** (coded **1**) within the new grouping variable of the **GROUP**.
4. Repeat steps 1 to 3 to create the other three levels: **male-nonwhite** (coded **2**), **female-white** (coded **3**), and **female-nonwhite** (coded **4**).



5. To aid interpretation of the obtained results, **Value Labels** in the data set should be activated and labels attached to the numerical codes for the four levels. To do this, open the data set, and under **Variable View**, click the **Values** field for the **GROUP** variable. Type in the value labels as shown in the **Value Labels** window below.

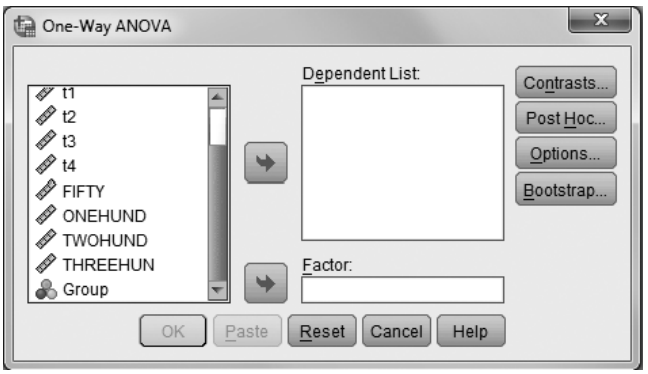


8.6.10.2 Post Hoc Comparisons

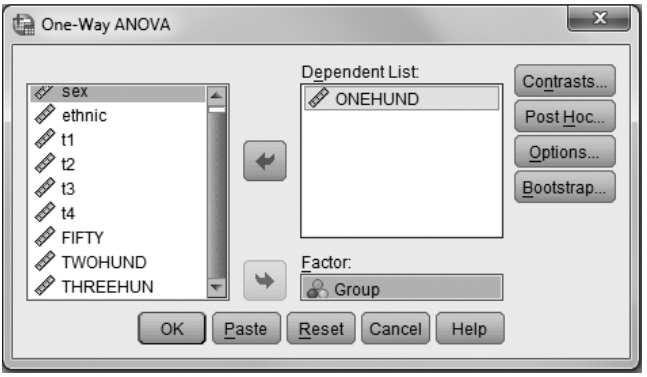
Once the four levels (male-white, male-nonwhite, female-white, female-nonwhite) have been created, **Scheffé** post hoc comparisons can be conducted to test for differences (simple effects) between these four levels.

8.6.10.2.1 Post Hoc Comparisons (Windows Method)

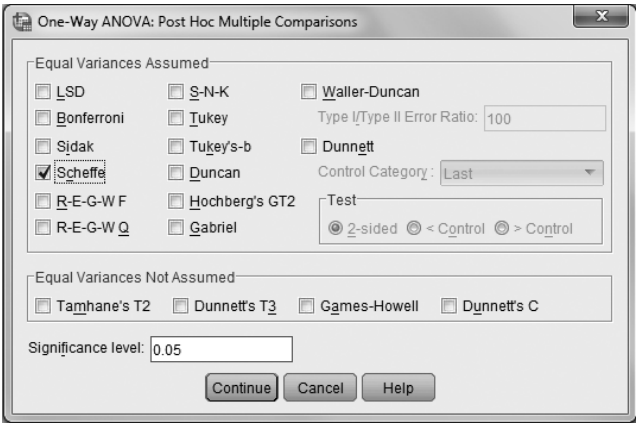
From the menu bar, click **Analyze**, then **Compare Means**, and then **One-Way ANOVA**. The following **One-Way ANOVA** window will open. Note that the list of variables now includes the newly created variable **GROUP**, which contains the four levels male-white, male-nonwhite, female-white, and female-nonwhite.



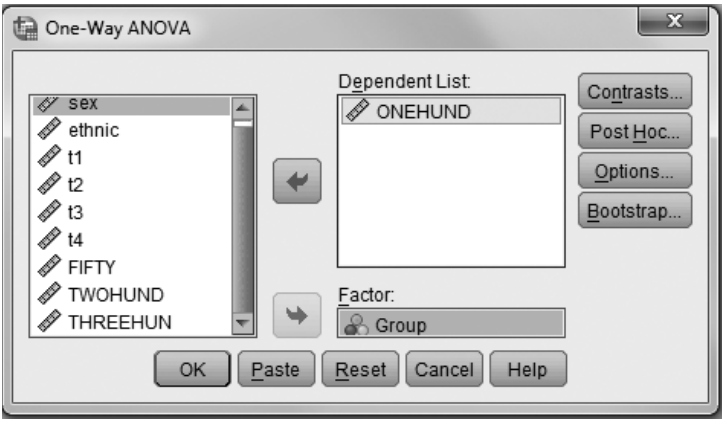
1. Transfer the dependent variable **ONEHUND** to the **Dependent List:** field by clicking (highlight) the variable and then clicking ▶. Transfer the independent variable **GROUP** to the **Factor:** field by clicking (highlight) the variable and then clicking ▶. To do post hoc comparisons for the running times in the 100-yard dash between the four levels (male-white, male-nonwhite, female-white, female-nonwhite), click Post Hoc....



2. When the following **One-Way ANOVA: Post Hoc Multiple Comparisons** window opens, check the **Scheffe** box to run the Scheffé post hoc test. Next, click Continue.



3. When the **One-Way ANOVA** window opens, run the analysis by clicking OK. See Table 8.6 for the results.



8.6.10.2.2 Post Hoc Comparisons (SPSS Syntax Method)

```
IF (SEX EQ 1 AND ETHNIC EQ 1) GROUP = 1.
IF (SEX EQ 1 AND ETHNIC EQ 2) GROUP = 2.
IF (SEX EQ 2 AND ETHNIC EQ 1) GROUP = 3.
IF (SEX EQ 2 AND ETHNIC EQ 2) GROUP = 4.
VALUE LABELS GROUP 1 'MALE-WHITE' 2 'MALE-NONWHITE'
3 'FEMALE-WHITE' 4 'FEMALE-NONWHITE'.
ONEWAY ONEHUND BY GROUP (1,4)
/RANGES = SCHEFFE(.05).
```

8.6.11 SPSS Output

TABLE 8.6

Scheffé Post Hoc Comparisons Output

Multiple Comparisons

Dependent Variable: Onehund

Scheffe

(I) Group	(J) Group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
MALE-WHITE	MALE-NONWHITE	3.0000	1.2490	.167	−.8933	6.8933
	FEMALE-WHITE	−10.0000*	1.2490	.000	−13.8933	−6.1067
	FEMALE-NONWHITE	−14.8000*	1.2490	.000	−18.6933	−10.9067
MALE-NONWHITE	MALE-WHITE	−3.0000	1.2490	.167	−6.8933	.8933
	FEMALE-WHITE	−13.0000*	1.2490	.000	−16.8933	−9.1067
	FEMALE-NONWHITE	−17.8000*	1.2490	.000	−21.6933	−13.9067

TABLE 8.6 (*Continued*)

Scheffé Post Hoc Comparisons Output

Multiple Comparisons						
Dependent Variable: Onehund						
Scheffe						
					95% Confidence Interval	
(I) Group	**(J) Group**	**Mean Difference (I-J)**	**Std. Error**	**Sig.**	**Lower Bound**	**Upper Bound**
FEMALE-WHITE	MALE-WHITE	10.000*	1.2490	.000	6.1067	13.8933
	MALE-NONWHITE	13.0000*	1.2490	.000	9.1067	16.8933
	FEMALE-NONWHITE	−4.8000*	1.2490	.013	−8.6933	−.9067
FEMALE-NONWHITE	MALE-WHITE	14.8000*	1.2490	.000	10.9067	18.6933
	MALE-NONWHITE	17.8000*	1.2490	.000	13.9067	21.6933
	FEMALE-WHITE	4.8000*	1.2490	.013	.9067	8.6933

* The mean difference is significant at the .05 level.

8.6.12 Results and Interpretation

Results from the Scheffé comparisons indicate that white females ($M = 17.20$) and nonwhite females ($M = 22.00$) ran significantly slower than white males ($M = 7.20$) and nonwhite males ($M = 4.20$). While there was no significant difference in running times between white males and nonwhite males, the results show that nonwhite females ran significantly slower than white females.

9

General Linear Model: Repeated Measures Analysis

9.1 Aim

When a subject is tested on the same variable over time, it is a repeated measures design. While the advantages of repeated measurements are obvious (e.g., they require fewer subjects per experiment, and they eliminate between-subjects differences from the experimental error), they violate the most important assumption of multivariate analysis, *independence*—the same subject is measured repeatedly on the same variable. GLM for repeated measures is a special procedure that can account for this dependence and still test for differences across individuals for the set of dependent variables.

9.2 Assumption

Sphericity—The relationship between pairs of variables is similar. For example, if five subjects were given the same sequential treatments (A, B, C, and D), then the assumption of sphericity states that the variance of the differences between treatment A and B equals the variance of the difference between A and C, which equals the variance of the differences between A and D, which equals the variance of the differences between B and C…. The sphericity assumption is tested by Mauchly's test as part of the repeated measures analysis. If the Mauchly's test result is significant, then there are differences between pairs of variables and the sphericity assumption is not satisfied.

9.3 Example 1: GLM: One-Way Repeated Measures

An investigator is interested in examining how exposure to different degrees of temperature influences problem-solving ability. Five subjects were asked to solve a series of mathematical problems under four different temperature conditions. The temperature in the room in which the subjects were tested was originally set at 35°C, with the temperature dropping by 5° every five minutes with a minimum of 20°C. Thus, each subject was asked to solve a set of mathematical problems under the four temperature conditions of 35°C, 30°C, 25°C, and 20°C. The investigator expected that the greatest number of errors would be made when the room temperature was 35°C, fewer errors would be made when the room temperature was 30°C, still fewer when the room temperature was 25°C, and the fewest errors would be made when the room temperature was 20°C. The data below represent the number of errors made by five subjects across the four conditions.

	Room Temperature (°C)			
Subjects	35	30	25	20
s1	7	5	2	3
s2	8	7	5	4
s3	6	5	3	3
s4	8	8	4	2
s5	5	4	3	2

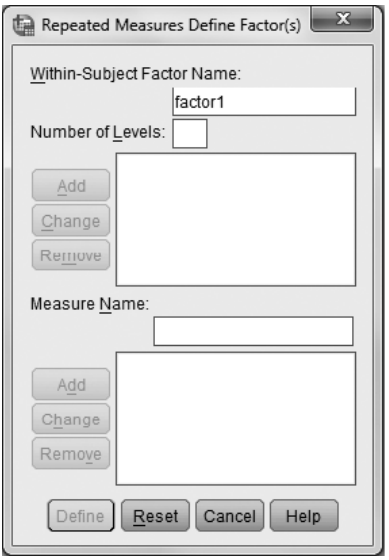
9.3.1 Data Entry Format

The data set has been saved under the name **EX9a.SAV**.

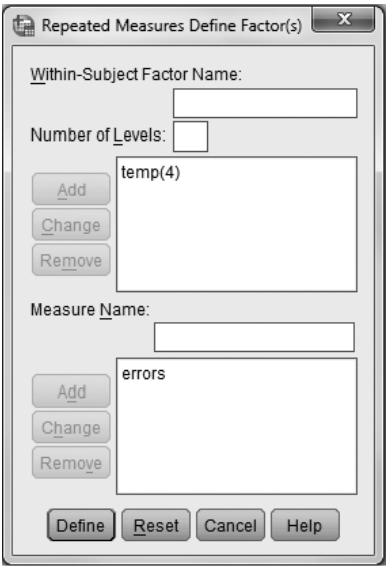
Variables	Column(s)	Code
Temp1	1	Number of errors
Temp2	2	Number of errors
Temp3	3	Number of errors
Temp4	4	Number of errors

9.3.2 Windows Method

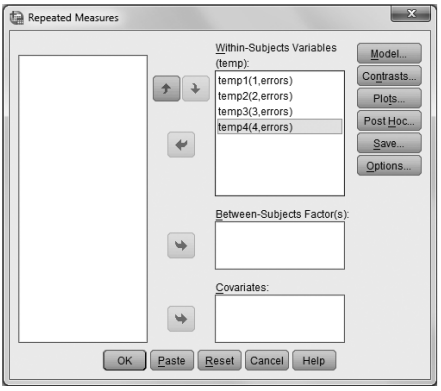
1. From the menu bar, click **Analyze**, then **General Linear Model**, and then **Repeated Measures**. The following **Repeated Measures Define Factor(s)** window will open.



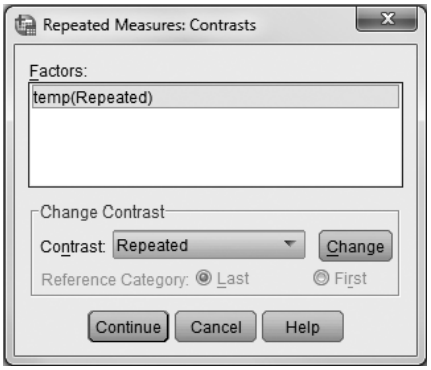
2. In the **Within-Subject Factor Name:** field, type the name **TEMP** to denote the name of the within-subject factor. In the **Number of Levels:** field, type **4** to denote the four levels of the within-subject factor **TEMP** (Temp1, Temp2, Temp3, Temp4). Click Add to transfer the **TEMP** factor to the **ADD** field. In the **Measure Name**: field, type the name of the dependent measure, **ERRORS**, and then click Add.



3. Click Define to open the **Repeated Measures** window below. Transfer the four variables Temp1, Temp2, Temp3 and Temp4 to the **Within-Subjects Variables (temp):** field by clicking these variables (highlight) and then clicking ▶.

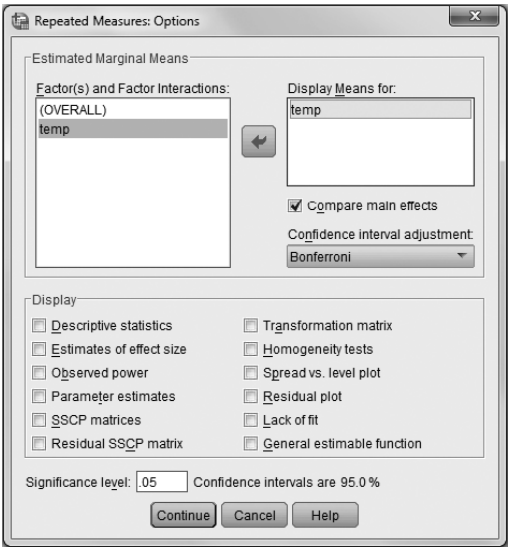


4. In order to test for differences among the four levels (Temp1, Temp2, Temp3, Temp4) of the within-subjects factor **TEMP**, click Contrasts... to open the **Repeated Measures: Contrasts** window below. Select a contrast from the **Contrast:** drop-down list and right-click the mouse button on the contrast to see its description. For this example, choose **Repeated** as the contrast and click Change. This contrast compares the mean of each level (except the last) to the mean of the subsequent level. Click Continue to return to the **Repeated Measures** window.

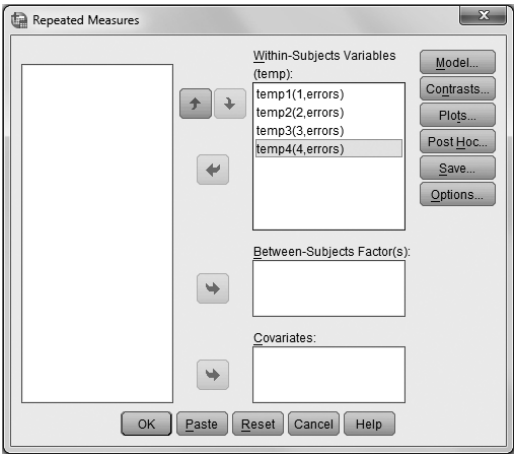


5. In the **Repeated Measures** window, click Options... to open the **Repeated Measures: Options** window that follows. In order to display means for the four levels of **TEMP** (Temp1, Temp2, Temp3, Temp4), click (highlight) **TEMP** in the **Factor(s) and Factor Interactions:** field, and then click ▶ to transfer this variable to the **Display Means for:** field.

In order to compute **pairwise comparisons** of the means for Temp1, Temp2, Temp3, and Temp4, check the **Compare main effects** cell. However, because of inflated Type I error rate owing to the multiple comparisons, a **Bonferroni** type adjustment should be made. From the **Confidence interval adjustment:** field, select **Bonferroni** from the drop-down list. Click Continue to return to the **Repeated Measures** window.



6. When the **Repeated Measures** window opens, click OK to complete the analysis. See Table 9.1 for the results.



9.3.3 SPSS Syntax Method

```
GLM TEMP1 TO TEMP4
/WSFACTOR = TEMP 4 REPEATED
/MEASURE = ERRORS
/EMMEANS = TABLES(TEMP) COMPARE ADJ(BONFERRONI).
```

Note: As the investigator expected fewer errors to be made when the room temperature dropped from 35°C to 20°C, specific comparisons or contrasts between the temperature conditions can be made. In this example, the **Repeated** contrast has been used to compare the mean of each level (except the last) to the mean of the subsequent level. SPSS offers the following contrast types for GLM repeated measures analysis:

Deviation. Compares the mean of each level (except a reference category) to the mean of all of the levels (grand mean). The levels of the factor can be in any order.

Simple. Compares the mean of each level to the mean of a specified level. This type of contrast is useful when there is a control group. You can choose the first or last category as the reference.

Difference. Compares the mean of each level (except the first) to the mean of previous levels.

Helmert. Compares the mean of each level of the factor (except the last) to the mean of subsequent levels.

Repeated. Compares the mean of each level (except the last) to the mean of the subsequent level.

Polynomial. Compares the linear effect, quadratic effect, cubic effect, and so forth. The first degree of freedom contains the linear effect across all categories; the second degree of freedom, the quadratic effect; and so on. These contrasts are often used to estimate polynomial trends.

9.3.4 SPSS Output

TABLE 9.1

GLM One-Way Repeated Measures Output

Within-Subjects Factors	
Measure: MEASURE_1	
TEMP	**Dependent Variable**
1	TEMP1
2	TEMP2
3	TEMP3
4	TEMP4

TABLE 9.1 (*Continued*)

GLM One-Way Repeated Measures Output

Multivariate Tests[b]

Effect		Value	F	Hypothesis df	Error df	Sig.
TEMP	Pillai's Trace	.985	43.500[a]	3.000	2.000	.023
	Wilks' Lambda	.015	43.500[a]	3.000	2.000	.023
	Hotelling's Trace	65.250	43.500[a]	3.000	2.000	.023
	Roy's Largest Root	65.250	43.500[a]	3.000	2.000	.023

[a] Exact statistic
[b] Design: Intercept
Within-Subjects Design: TEMP

Mauchly's Test of Sphericity[b]

Measure: MEASURE_1

Within-Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[a] Greenhouse-Geisser	Huynh-Feldt	Lower-Bound
TEMP	.054	7.965	5	.182	.571	.958	.333

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identify matrix.
[a] May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed i Tests of Within-Subjects Effects table.
[b] Design: Intercept
Within Subjects Design: TEMP

Tests of Within-Subjects Effect

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
TEMP	Sphericity Assumed	54.600	3	18.200	24.539	.000
	Greenhouse-Geisser	54.600	1.714	31.853	24.539	.001
	Huynh-Feldt	54.600	2.874	18.995	24.539	.000
	Lower-Bound	54.600	1.000	54.600	24.539	.008
Error(TEMP)	Sphericity Assumed	8.900	12	.742		
	Greenhouse-Geisser	8.900	6.857	1.298		
	Huynh-Feldt	8.900	11.498	.774		
	Lower-Bound	8.900	4.000	2.225		

(*Continued*)

TABLE 9.1 (*Continued*)

GLM One-Way Repeated Measures Output

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	TEMP	Type III Sum of Squares	df	Mean Square	F	Sig.
TEMP	Level 1 vs. Level 2	5.000	1	5.000	10.000	.034
	Level 2 vs. Level 3	28.800	1	28.800	22.154	.009
	Level 3 vs. Level 4	1.800	1	1.800	1.385	.305
Error(TEMP)	Level 1 vs. Level 2	2.000	4	.500		
	Level 2 vs. Level 3	5.200	4	1.300		
	Level 3 vs. Level 4	5.200	4	1.300		

Tests of Between-Subjects Effects

Measure: MEASURE_1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig
Intercept	110.450	1	110.450	105.820	.001
Error	4.175	4	1.044		

Estimated Marginal Means

2. TEMP

Estimates

Measure: ERRORS

			95% Confidence Interval	
TEMP	Mean	Std. Error	Lower Bound	Upper Bound
1	6.800	.583	5.181	8.419
2	5.800	.735	3.760	7.840
3	3.400	.510	1.984	4.816
4	2.800	.374	1.761	3.839

Pairwise Comparisons

Measure: ERRORS

					95% Confidence Interval for Difference[a]	
(I) TEMP	(J) TEMP	Mean Difference (I-J)	Std. Error	Sig.[a]	Lower Bound	Upper Bound
1	2	1.000	.316	.205	−.534	2.534
	3	3.400*	.510	.016	.926	5.874
	4	4.000*	.548	.011	1.343	6.657

TABLE 9.1 (*Continued*)

GLM One-Way Repeated Measures Output

Pairwise Comparisons

Measure: ERRORS

(I) TEMP	(J) TEMP	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
2	1	−1.000	.316	.205	−2.534	.534
	3	2.400	.510	.056	−.074	4.874
	4	3.000	.775	.108	−.758	6.758
3	1	−3.400*	.510	.016	−5.874	−.926
	2	−2.400	.510	.056	−4.874	.074
	4	.600	.510	1.000	−1.874	3.074
4	1	−4.000*	.548	.011	−6.657	−1.343
	2	−3.000	.775	.108	−6.758	.758
	3	−.600	.510	1.000	−3.074	1.874

Based on estimated marginal means

* The mean difference is significant at the .05 level.

[a] Adjustment for multiple comparisons: Bonferroni.

Multivariate Tests

	Value	F	Hypothesis df	Error df	Sig.
Pillai's Trace	.985	43.500[a]	3.000	2.000	.023
Wilks' Lambda	.015	43.500[a]	3.000	2.000	.023
Hotelling's Trace	65.250	43.500[a]	3.000	2.000	.023
Roy's Largest Root	65.250	43.500[a]	3.000	2.000	.023

Each F tests the multivariate effect of TEMP. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

[a] Exact statistic

9.3.5 Choosing Tests of Significance

The experiment involves testing the effect of the within-subjects variable **TEMP** on the number of errors made. In interpreting the results of the analysis, the SPSS output allows the researcher to interpret either the **Multivariate Tests** results, or the **Tests of Within-Subjects Effects** results**.** Choosing between the two depends on the outcome of **Mauchly's Test of Sphericity**, which tests the hypothesis that the covariance matrix has a constant variance on the diagonal and zeros off the diagonal, i.e, the correlations between the variables are all roughly the same (primary assumption underlying the *repeated measures* design).

- If **Mauchly's Test of Sphericity** is not significant (i.e., the assumption about the characteristics of the variance-covariance matrix is not violated), the **Tests of Within-Subjects Effects** can be used.
- If **Mauchly's Test of Sphericity** is significant, the **Multivariate Tests** should be applied.

However, if the researcher chooses to interpret the **Tests of Within-Subjects Effects** averaged *F* tests, then an adjustment to the numerator and denominator degrees of freedom must be made. Both the numerator and denominator degrees of freedom must be multiplied by the **Huynh-Feldt epsilon** (presented as part of the **Tests of Within-Subjects Effects** table), and the significance of the *F* ratio evaluated with the new degrees of freedom.

9.3.6 Results and Interpretation

The results from the analysis indicate that the **Mauchly's Test of Sphericity** is not significant ($p = .158$). Therefore, the **Tests of Within-Subjects Effects** table can be interpreted. The result indicates that the within-subjects variable of **TEMP** is highly significant, $F(3,12) = 24.54$, $p < .001$. That is, the number of errors made by the subjects differed significantly as a function of the four temperature conditions. This is supported by the decrease in the average number of errors made as the room temperature decreased from 35°C to 20°C.

Had the results of **Mauchly's Test of Sphericity** been significant, and had the researcher decided to interpret the **Tests of Within-Subjects Effects**, then an adjustment to the numerator and denominator degrees of freedom must be made. This is accomplished by multiplying these two values by the Huynh-Feldt epsilon. Therefore, the adjusted numerator degrees of freedom is $3 \times 0.95816 = 2.87$; the adjusted denominator degrees of freedom is $12 \times 0.95816 = 11.50$. The *F* ratio of 24.54 must then be evaluated with these new degrees of freedom. The significance of each test when sphericity is assumed, or when the Huynh-Feldt epsilon is used, is displayed under the **Tests of Within-Subjects Effects** table.

As the within-subjects variable **TEMP** is statistically significant, results from the **Repeated** contrast can be interpreted to determine which variables contributed to the overall difference. These results are presented under the heading of **Tests of Within-Subjects Contrasts**. The first contrast is between Level 1 (Temp1) versus Level 2 (Temp2), which is statistically significant, $F(1,4) = 10.00$, $p < .05$. This means that the mean number of errors made in the 35°C condition ($M = 6.80$) is significantly greater than the mean number of errors made in the 30°C condition ($M = 5.80$). The second contrast is between Level 2 (Temp2) versus Level 3 (Temp3), which is statistically significant, $F(1,4) = 22.15$, $p < .01$. This means that the mean number of errors made in the 30°C condition ($M = 5.80$) is significantly greater than the mean number of errors made in the 25°C condition ($M = 3.40$). The third contrast is between Level 3 (Temp3) versus Level 4 (Temp4) and is non-significant, $F(1,4) = 1.39$, $p > .05$. This means that the difference in the mean number of errors made in the 25°C ($M = 3.40$) and 20°C ($M = 2.80$) conditions is due to chance variation.

The **Pairwise Comparisons** table presents all pairwise comparisons (with Bonferroni adjustment) between the four levels. The results indicate that the number of errors generated under Temp1 (35°C) is significantly larger than the number of errors generated in both Temp3 (25°C) and Temp4 (20°C). There is no significant difference in the number of errors generated under Temp1 (35°C) and Temp2 (30°C).

The **Multivariate Tests** statistics table (Pillai's trace, Wilks' lambda, Hotelling's trace, Roy's largest root) indicates that the overall difference in the number of errors generated across the four levels of temperature is statistically significant ($p < .05$).

9.4 Example 2: GLM: Two-Way Repeated Measures (Doubly Multivariate Repeated Measures)

In the previous example, subjects were tested on one variable across four trials. In the following example, subjects will be tested on two variables, that is, there will be two within-subjects variables.

An experimenter wants to study the effect of pain on memorization. Specifically, he wants to study the relationship between strength of electrical shock (low versus high shocks) and type of items (easy, moderate, difficult) on retention. The scores recorded for each subject are the subject's total number of correct responses on the easy, moderate, and difficult items when he/she is tested under low and high shocks.

	Low Shock			High Shock		
Subject	**Easy**	**Moderate**	**Difficult**	**Easy**	**Moderate**	**Difficult**
s1	83	88	100	75	60	40
s2	100	92	83	71	54	41
s3	100	97	100	50	47	43
s4	89	87	92	74	58	47
s5	100	96	90	83	72	62
s6	100	89	100	100	86	70
s7	92	94	92	72	63	44
s8	100	95	89	83	74	62
s9	50	60	54	33	25	38
s10	100	97	100	72	64	50

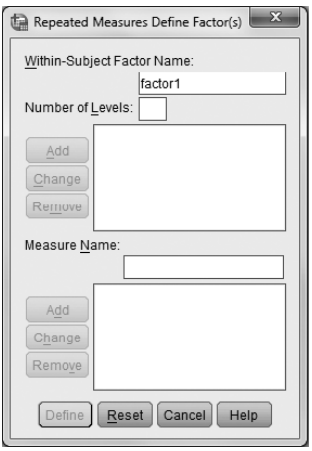
9.4.1 Data Entry Format

The data set has been saved under the name **EX9b.SAV**.

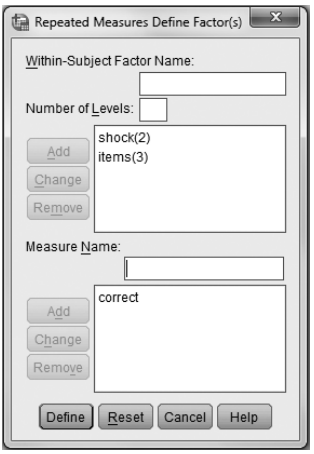
Variables	Column(s)	Code
LS_E	1	Number of correct responses
LS_M	2	Number of correct responses
LS_D	3	Number of correct responses
HS_E	4	Number of correct responses
HS_M	5	Number of correct responses
HS_D	6	Number of correct responses

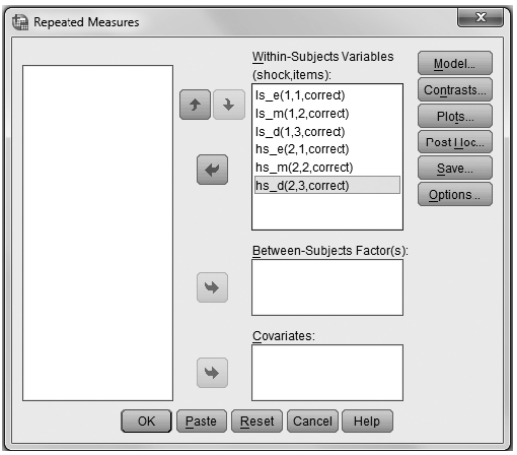
9.4.2 Windows Method

1. From the menu bar, click **Analyze**, then **General Linear Model**, and then **Repeated Measures**. The following **Repeated Measures Define Factor(s)** window will open.

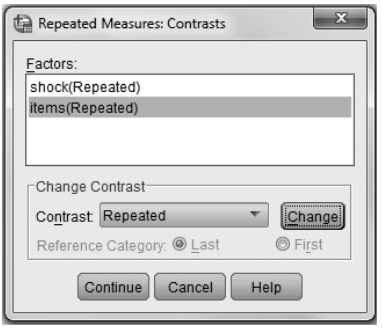


2. In the **Within-Subject Factor Name:** field, type **SHOCK** to denote the name of the first within-subject factor. In the **Number of Levels:** field, type **2** to denote the two levels of the **SHOCK** factor (Low Shock, High Shock). Click Add to transfer the **SHOCK** factor to the **ADD** field. Repeat the above procedure to create the second within-subject factor, **ITEMS**. For this factor, there are three levels (easy, moderate, difficult). In the **Measure Name**: field, type the name of the dependent measure, **CORRECT**, and then click Add.
3. Click Define to open the **Repeated Measures** window. Transfer the six variables LS_E, LS_M, LS_D, HS_E, HS_M, and HS_D (generated from the 2 (SHOCK) × 3 (ITEMS) factorial combination) to the **Within-Subjects Variables (shock, items):** cell by clicking these variables (highlight) and then clicking ▶.
4. In order to test for differences among the six levels of LS_E, LS_M, LS_D, HS_E, HS_M, and HS_D (generated from the factorial



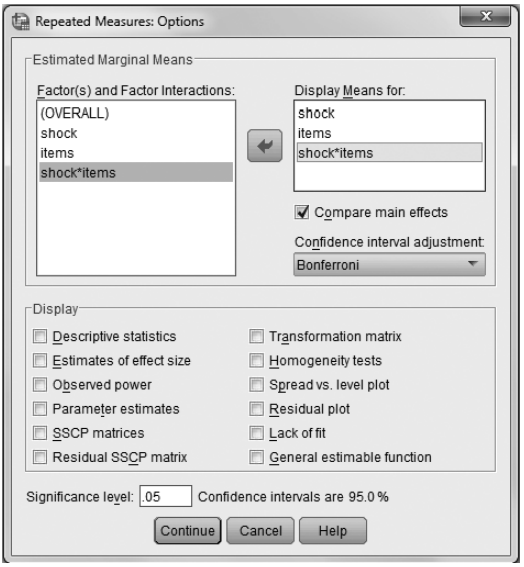


combination of the two within-subject factors **SHOCK** and **ITEMS**), click Contrasts... to open the following **Repeated Measures: Contrasts** window. Select **Repeated** as the contrast from the **Contrast:** drop-down list and click Change. This contrast compares (1) the overall mean number of correct responses obtained under the low shock and high shock conditions (i.e., collapsing across the three types of items (Easy, Moderate, Difficult)), (2) the overall mean number of correct responses obtained under the easy, moderate, and difficult item conditions (i.e., collapsing across the two levels of shock (low shock, high shock)), and (3) the overall mean number of correct responses obtained under the easy, moderate, and difficult item conditions across the two levels of shock (i.e., the **SHOCK*ITEMS** interaction). Click Continue to return to the **Repeated Measures** window.

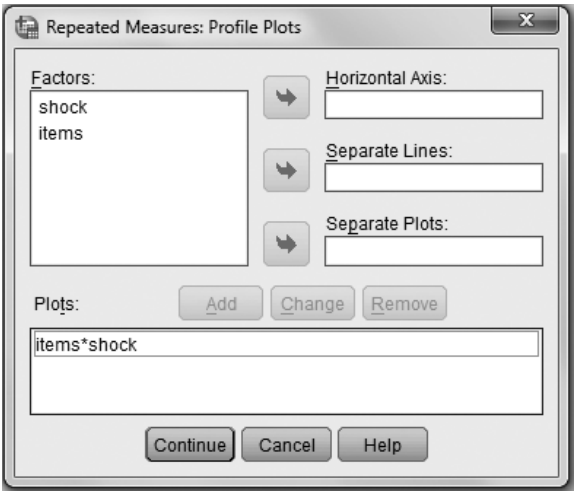


5. In the **Repeated Measures** window, click Options... to open the **Repeated Measures: Options** window below. In order to display means for the two levels of **SHOCK,** the three levels of **ITEMS**, and the **SHOCK*ITEMS** interaction, click (highlight) **SHOCK**, **ITEMS**, and **SHOCK*ITEMS** in the **Factor(s) and Factor Interactions:** field, and then click ▶ to transfer these variables to the **Display Means for:** field.

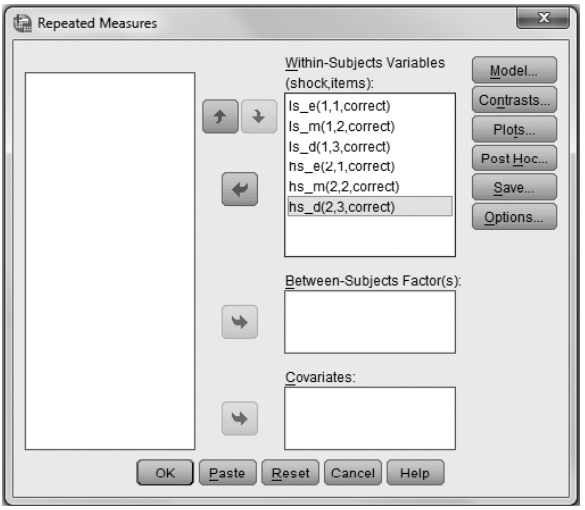
   In order to compute pairwise comparisons of the means for the **SHOCK** (low shock versus high shock) and **ITEMS** variables (easy versus moderate, easy versus difficult, moderate versus difficult), check the **Compare main effects** field. However, because of inflated Type I error rate owing to the multiple comparisons, a **Bonferroni** type adjustment should be made. From the **Confidence interval adjustment:** field, select **Bonferroni** from the drop-down list. Click Continue to return to the **Repeated Measures** window.



6. In order to aid interpretation of the **SHOCK*ITEMS** interaction effect, it would be useful to graph the means of this interaction. In the **Repeated Measures** window, click Plots... to open the **Repeated Measures: Profile Plots** window below. Transfer the **ITEMS** factor to the **Horizontal Axis:** field by clicking it (highlight) and then clicking ▶. Transfer the **SHOCK** factor to the **Separate Lines:** field by clicking it (highlight) and then clicking ▶. Transfer this profile plot to the **Plots:** field by clicking Add. Click Continue to return to the **Repeated Measures** window.



7. When the **Repeated Measures** window opens, click OK to complete the analysis. See Table 9.2 for the results.



9.4.3 SPSS Syntax Method

```
GLM LS_E TO HS_D
/WSFACTOR = SHOCK 2 REPEATED ITEMS 3 REPEATED
/MEASURE = CORRECT
/PLOT = PROFILE(ITEMS*SHOCK)
/EMMEANS = TABLES(SHOCK) COMPARE ADJ(BONFERRONI)
/EMMEANS = TABLES(ITEMS) COMPARE ADJ(BONFERRONI)
/EMMEANS = TABLES(SHOCK*ITEMS).
```

9.4.4 SPSS Output

TABLE 9.2

Doubly Multivariate Repeated Measures Output

General Linear Model

Within-Subjects Factors		
Measure: CORRECT		
SHOCK	**ITEMS**	**Dependent Variable**
1	1	is_e
	2	is_m
	3	is_d
2	1	hs_e
	2	hs_m
	3	hs_d

Multivariate Tests[b]

Effect		Value	F	Hypothesis df	Error df	Sig.
SHOCK	Pillai's Trace	.887	70.792[a]	1.000	9.000	.000
	Wilks' Lambda	.113	70.792[a]	1.000	9.000	.000
	Hotelling's Trace	7.866	70.792[a]	1.000	9.000	.000
	Roy's Largest Root	7.866	70.792[a]	1.000	9.000	.000
ITEMS	Pillai's Trace	.735	11.092[a]	2.000	8.000	.005
	Wilks' Lambda	.265	11.092[a]	2.000	8.000	.005
	Hotelling's Trace	2.773	11.092[a]	2.000	8.000	.005
	Roy's Largest Root	2.773	11.092[a]	2.000	8.000	.005
SHOCK * ITEMS	Pillai's Trace	.745	11.684[a]	2.000	8.000	.004
	Wilks' Lambda	.255	11.684[a]	2.000	8.000	.004
	Hotelling's Trace	2.921	11.684[a]	2.000	8.000	.004
	Roy's Largest Root	2.921	11.684[a]	2.000	8.000	.004

[a] Exact statistic

[b] Design: Intercept

Within-Subjects Design: SHOCK+ITEMS+SHOCK*ITEMS

TABLE 9.2 (*Continued*)

Doubly Multivariate Repeated Measures Output

Mauchly's Test Sphericity[b]

Measure: CORRECT

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[a] Greenhouse-Geisser	Huynh-Feldt	Lower-Bound
SHOCK	1.000	.000	0		1.000	1.000	1.000
ITEMS	.372	7.913	2	.019	.614	.662	.500
SHOCK * ITEMS	.454	6.322	2	.042	.647	.709	.500

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variable is proportional to an identity matrix.

[a] May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

[b] Design: Intercept

Within-Subjects Design: SHOCK+ITEMS+SHOCK*ITEMS

Tests of Within-Subjects Effects

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
SHOCK	Sphericity Assumed	13380.267	1	13380.267	70.792	.000
	Greenhouse-Geisser	13380.267	1.000	13380.267	70.792	.000
	Huynh-Feldt	13380.267	1.000	13380.267	70.792	.000
	Lower-Bound	13380.267	1.000	13380.267	70.792	.000
Error(SHOCK)	Sphericity Assumed	1701.067	9	189.007		
	Greenhouse-Geisser	1701.067	9.000	189.007		
	Huynh-Feldt	1701.067	9.000	189.007		
	Lower-Bound	1701.067	9.000	189.007		
ITEMS	Sphericity Assumed	1329.033	2	664.517	20.367	.000
	Greenhouse-Geisser	1329.033	1.228	1081.911	20.367	.001
	Huynh-Feldt	1329.033	1.323	1004.334	20.367	.000
	Lower-Bound	1329.033	1.000	1329.033	20.367	.001
Error(ITEMS)	Sphericity Assumed	587.300	18	32.628		
	Greenhouse-Geisser	587.300	11.056	53.122		
	Huynh-Feldt	587.300	11.910	49.313		
	Lower-Bound	587.300	9.000	65.256		

(*Continued*)

TABLE 9.2 (*Continued*)

Doubly Multivariate Repeated Measures Output

Tests of Within-Subjects Effects

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
SHOCK * ITEMS	Sphericity Assumed	1023.433	2	511.717	12.683	.000
	Greenhouse-Geisser	1023.433	1.293	791.247	12.683	.003
	Huynh-Feldt	1023.433	1.419	721.313	12.683	.002
	Lower-Bound	1023.433	1.000	1023.433	12.683	.006
Error(SHOCK * ITEMS)	Sphericity Assumed	726.233	18	40.346		
	Greenhouse-Geisser	726.233	11.641	62.386		
	Huynh-Feldt	726.233	12.770	56.872		
	Lower-Bound	726.233	9.000	80.693		

Tests of Within-Subjects Contrasts

Measure: CORRECT

Source	SHOCK	ITEMS	Type III Sum of Squares	df	Mean Square	F	Sig.
SHOCK	Level 1 vs. Level 2		8920.178	1	8920.178	70.792	.000
Error(SHOCK)	Level 1 vs. Level 2		1134.044	9	126.005		
ITEMS		Level 1 vs. Level 2	416.025	1	416.025	23.369	.001
		Level 2 vs. Level 3	255.025	1	255.025	11.757	.008
Error(ITEMS)		Level 1 vs. Level 2	160.225	9	17.803		
		Level 2 vs. Level 3	195.225	9	21.692		
SHOCK * ITEMS	Level 1 vs. Level 2	Level 1 vs. Level 2	828.100	1	828.100	18.969	.002
		Level 2 vs. Level 3	1232.100	1	1232.100	6.050	.036
Error(SHOCK * ITEMS)	Level 1 vs. Level 2	Level 1 vs. Level 2	392.900	9	43.656		
		Level 2 vs. Level 3	1832.900	9	203.656		

Tests of Between-Subjects Effects

Measure: CORRECT

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig
Intercept	56801.344	1	56801.344	353.185	.000
Error	1447.433	9	160.826		

TABLE 9.2 (*Continued*)

Doubly Multivariate Repeated Measures Output

Estimated Marginal Means

1. SHOCK

Estimates

Measure: CORRECT

			95% Confidence Interval	
SHOCK	**Mean**	**Std. Error**	**Lower Bound**	**Upper Bound**
1	90.300	4.097	81.033	99.567
2	60.433	4.657	49.899	70.967

Pairwise Comparisons

Measure: CORRECT

					95% Confidence Interval for Difference[a]	
(I) SHOCK	**(J) SHOCK**	**Mean Difference (I-J)**	**Std. Error**	**Sig.[a]**	**Lower Bound**	**Upper Bound**
1	2	29.867*	3.550	.000	21.837	37.897
2	1	−29.867*	3.550	.000	−37.897	−21.837

Based on estimated marginal means.

* The mean difference is significant at the .05 level.

[a] Adjustment for multiple comparisons: Bonferroni.

Multivariate Tests

	Value	**F**	**Hypothesis df**	**Error df**	**Sig.**
Pillai's Trace	.887	70.792[a]	1.000	9.000	.000
Wilks' Lambda	.113	70.792[a]	1.000	9.000	.000
Hotelling's Trace	7.866	70.792[a]	1.000	9.000	.000
Roy's Largest Root	7.866	70.792[a]	1.000	9.000	.000

Each F tests the multivariate effect of SHOCK. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

[a] Exact statistic

2. ITEMS

Estimates

Measure: CORRECT

			95% Confidence Interval	
ITEMS	**Mean**	**Std. Error**	**Lower Bound**	**Upper Bound**
1	81.350	4.974	70.097	92.603
2	74.900	4.009	65.831	83.969
3	69.850	3.270	62.452	77.248

(*Continued*)

TABLE 9.2 (*Continued*)

Doubly Multivariate Repeated Measures Output

Pairwise Comparisons

Measure: CORRECT

(I) ITEMS	(J) ITEMS	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a] Lower Bound	Upper Bound
1	2	6.450*	1.334	.003	2.536	10.364
	3	11.500*	2.416	.003	4.412	18.588
2	1	−6.450*	1.334	.003	−10.364	−2.536
	3	5.050*	1.473	.023	.730	9.370
3	1	−11.500*	2.416	.003	−18.588	−4.412
	2	−5.050*	1.473	.023	−9.370	−.730

Based on estimated marginal means.

* The mean difference is significant at the .05 level.

[a] Adjustment for multiple comparisons: Bonferroni.

Multivariate Tests

	Value	F	Hypothesis df	Error df	Sig.
Pillai's trace	.735	11.092[a]	2.000	8.000	.005
Wilks' lambda	.265	11.092[a]	2.000	8.000	.005
Hotelling's trace	2.773	11.092[a]	2.000	8.000	.005
Roy's largest root	2.773	11.092[a]	2.000	8.000	.005

Each F tests the multivariate effect of ITEMS. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

[a] Exact statistic

3. SHOCK * ITEMS

Measure: CORRECT

SHOCK	ITEMS	Mean	Std. Error	95% Confidence Interval Lower Bound	Upper Bound
1	1	91.400	4.983	80.128	102.672
	2	89.500	3.481	81.626	97.374
	3	90.000	4.415	80.013	99.987
2	1	71.300	5.812	58.152	84.448
	2	60.300	5.243	48.439	72.161
	3	49.700	3.506	41.770	57.630

9.4.5 Results and Interpretation

The **Multivariate Tests** involving the first within-subjects variable **SHOCK** yielded highly significant results for all four multivariate tests (Pillai's, Hotelling's, Wilks', Roy's), $F(1,9) = 70.79$, $p < .001$. *Note that for this variable, the* ***Mauchly's Test of Sphericity*** *statistic cannot be calculated as there are only two levels of this variable and as such there is no variance-covariance matrix to be analyzed*. From the **Estimated Marginal Means** table, it can be seen that on average, subjects obtained more correct responses under the low shock condition ($M = 90.30$) than in the high shock condition ($M = 60.43$).

For the second within-subjects variable **ITEMS**, **Mauchly's Test of Sphericity** yielded a value of 0.37 (chi-square approximate value of 7.91 with two degrees of freedom) and is significant ($p < .05$). As the assumption of sphericity is violated, the **Multivariate Statistics** should be interpreted (if the researcher chooses to interpret the **Tests of Within-Subjects Effects** statistics, then an adjustment must be made to the numerator and denominator degrees of freedom, using the Huynh-Feldt epsilon; the **Tests of Within-Subjects Effects** table displays the significance of each test when sphericity is assumed, or when the Huynh-Feldt epsilon is used).

All four multivariate tests of significance (Pillai's, Hotelling's, Wilks', Roy's) indicate that the within-subjects variable **ITEMS** is significant, $F(2,8) = 11.09$, $p < .01$. From the **Estimated Marginal Means** table, it can be seen that on average, subjects obtained the most number of correct responses for the easy items ($M = 81.35$), fewer correct responses for the moderate items ($M = 74.90$), and the fewest correct responses for the difficult items ($M = 69.85$).

For the **SHOCK*ITEMS** interaction effect, **Mauchly's Test of Sphericity** is significant, indicating that the assumption of sphericity is violated. Thus, the **Multivariate Tests** of significance will be interpreted. All four multivariate tests of significance (Pillai's, Hotelling's, Wilks', Roy's) indicate that the interaction effect is significant, $F(2,8) = 11.68$, $p < .01$, suggesting that the number of correct responses made across the three levels of difficulty of the items is dependent on the strength of the electrical shock. In order to assist interpretation of the interaction effect, it is useful to analyze the graph shown in Figure 9.1.

Figure 9.1 shows that the number of correct responses obtained across the three difficulty levels of the **ITEMS** variable is dependent on the strength of the electrical **SHOCK**. Under the low shock condition (coded 1), subjects obtained a similar number of correct responses across the three difficulty levels. However, under the high shock condition (coded 2), the number of correct responses obtained decreased steadily from easy to moderate to difficult.

The **Tests of Within-Subjects Contrasts** table can be used to compare the levels of the two within-subjects variables. For the **SHOCK** variable, the contrast (LEVEL 1 vs. LEVEL 2) compares the number of correct responses made under the **low shock** condition ($M = 90.30$) with the number of correct responses made under the **high shock** condition ($M = 60.43$), averaged across

Profile Plots

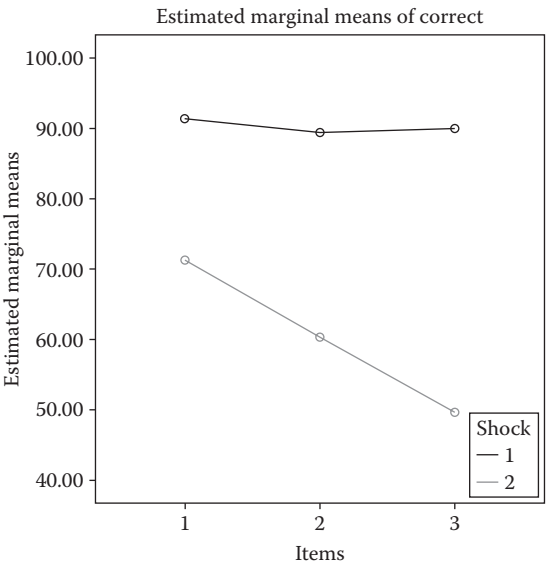


FIGURE 9.1
Items × Shock interaction.

the three levels of **ITEMS** (see the estimated marginal means for SHOCK). This contrast is significant, $F(1,9) = 70.79$, $p < .001$.

For the **ITEMS** variable, the first contrast (LEVEL 1 vs. LEVEL 2) compares the number of correct responses made under the **easy items** condition ($M = 81.35$) with the number of correct responses made under the **moderate items** condition ($M = 74.90$), averaged across the two levels of **SHOCK**. This contrast is significant, $F(1,9) = 23.37$, $p < .01$. The second contrast (LEVEL 2 vs. LEVEL 3) compares the number of correct responses made under the **moderate items** condition ($M = 74.9$) with the number of correct responses made under the **difficult items** condition ($M = 69.85$), averaged across the two levels of **SHOCK** (see the estimated marginal means for ITEMS). This contrast is significant, $F(1,9) = 11.76$, $p < .01$.

The **SHOCK*ITEMS** contrast tests the hypothesis that the mean of the specified ITEMS contrast is the same across the two SHOCK levels. The first contrast (LEVEL 1 vs. LEVEL 2) is significant, F (1,9) = 18,97, $p < .01$, which indicates that the mean difference in the number of correct responses made between **EASY-ITEMS** and **MODERATE-ITEMS** is not the same across the two levels of SHOCK.

Mean Difference (Easy-Items vs. Moderate-Items)	
Low shock	1.90 (91.4–89.5)
High shock	11.00 (71.3–60.3)

In conjunction with Figure 9.1, the results indicate that the decrease in the number of correct responses made between Easy-Items and Moderate-Items are different under the two shock conditions; the decrease is greater under the high shock condition (11.00) than under the low shock condition (1.90).

The second contrast (LEVEL 2 vs. LEVEL 3) is also significant, $F(1,9) = 6.05$, $p < .05$, which indicates that the mean difference in the number of correct responses made between **MODERATE-ITEMS** and **DIFFICULT-ITEMS** is not the same across the two levels of SHOCK.

Mean Difference (Moderate-Items vs. Difficult-Items)	
Low shock	–0.50 (89.50–90.00)
High shock	10.60 (60.30–49.70)

In conjunction with Figure 9.1, the results indicate that under the low shock condition, there was a mean average increase of 0.50 correct responses, while under the high shock condition there was an average decrease of 10.60 correct responses.

The **Pairwise Comparisons** table presents all pairwise comparisons between the two levels of the **SHOCK** variable and the three levels of the **ITEMS** variable (with Bonferroni adjustment). For the **SHOCK** variable, the results indicate that the number of correct responses obtained under the low shock condition ($M = 90.30$) is significantly greater ($p < .001$) than the number of correct responses obtained under the high shock condition ($M = 60.43$). For the **ITEMS** variable, the results indicate that (i) the number of correct responses obtained under the easy condition ($M = 81.35$) is significantly greater ($p < .01$) than the number of correct responses obtained under the moderate ($M = 74.90$) and difficult ($M = 69.85$) conditions, and (ii) the number of correct responses obtained under the moderate condition ($M = 74.90$) is significantly greater ($p < 05$) than the number of correct responses obtained under the difficult condition ($M = 69.85$).

9.5 Example 3: GLM: Two-Factor Mixed Design (One Between-Groups Variable and One Within-Subjects Variable)

An experimenter wishes to determine the effects of positive reinforcement on rate of learning. The subjects are randomly assigned to three groups, with the first receiving "effort + ability" feedback, the second receiving "effort" feedback only, and the third group serving as the control. All subjects were given three trials, and the number of correct responses per trial is recorded.

Group	Trial1	Trial2	Trial3
Control			
s1	3	3	4
s2	3	4	4
s3	1	3	4
s4	1	2	3
s5	2	3	5
s6	4	5	6
s7	4	5	6
s8	1	4	5
s9	1	4	5
s10	2	3	3
s11	2	2	4
Effort			
s12	2	3	6
s13	1	4	5
s14	3	6	7
s15	2	4	6
s16	1	2	3
s1	4	5	5
s18	1	3	4
s19	1	2	4
s20	2	3	4
s21	3	5	6
s22	4	6	7
Effort + Ability			
s23	2	4	8
s24	3	6	3
s25	4	7	7
s26	1	7	4
s27	3	7	12
s28	1	4	2
s29	1	5	3
s30	2	5	6
s31	3	6	6
s32	1	5	12
s33	4	9	7

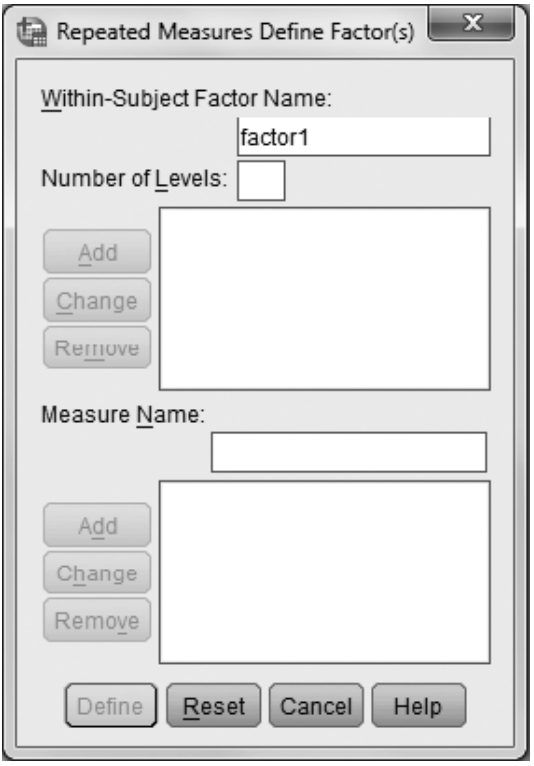
9.5.1 Data Entry Format

The data set has been saved under the name **EX9c.SAV**.

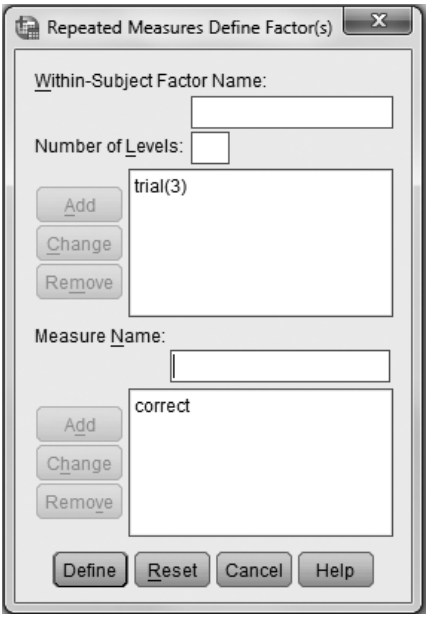
Variables	Column(s)	Code
Group	1	1 = control, 2 = effort, 3 = effort + ability
Trial1	2	Number of correct responses
Trial2	3	Number of correct responses
Trial3	4	Number of correct responses

9.5.2 Windows Method

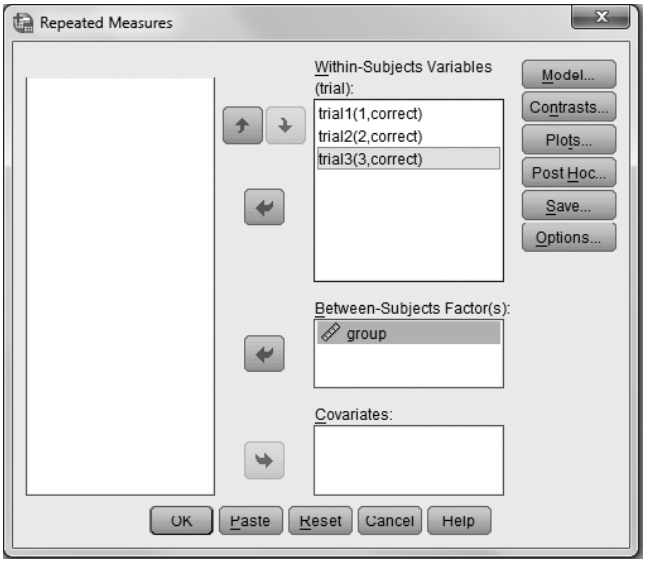
1. From the menu bar, click **Analyze**, then **General Linear Model**, and then **Repeated Measures**. The following **Repeated Measures Define Factor(s)** window will open.



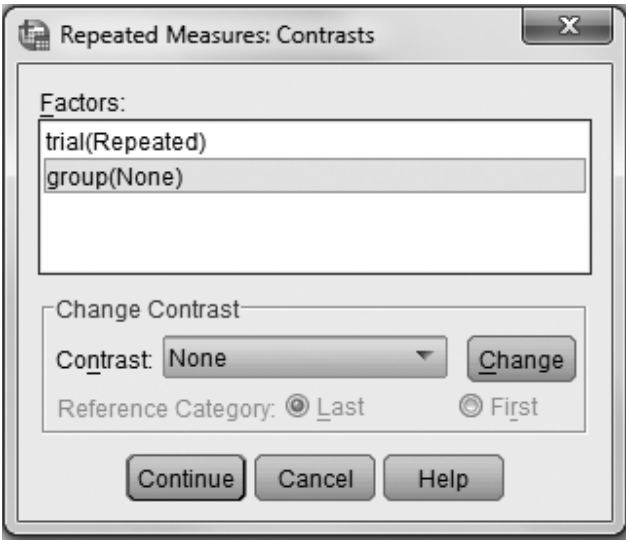
2. In the **Within-Subject Factor Name:** field, type **TRIAL** to denote the name of the within-subject factor. In the **Number of Levels:** field, type **3** to denote the three levels of the **TRIAL** factor (Trial1, Trial2, Trial3). Click Add to transfer the **TRIAL** factor to the **ADD** field. In the **Measure Name**: field, type the name of the dependent measure, **CORRECT**, and then click Add.



3. Click Define to open the **Repeated Measures** window below. Transfer Trial1, Trial2, and Trial3 to the **Within-Subjects Variables (trial):** field by clicking these variables (highlight) and then clicking ▶. Next, transfer the **GROUP** variable to the **Between-Subjects Factor(s):** field by clicking it (highlight) and then clicking ▶.

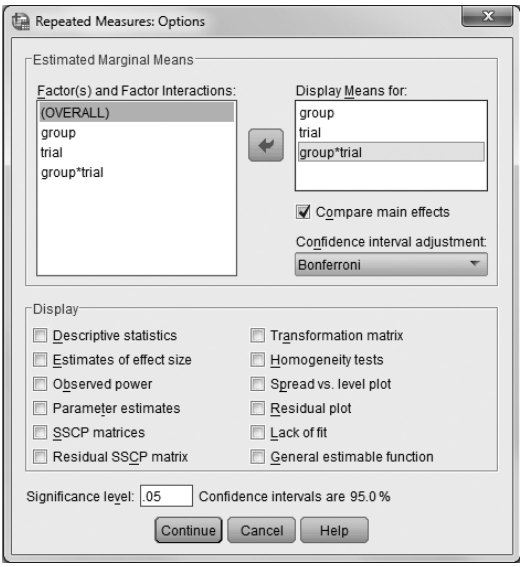


4. In order to test for differences among the three trial levels (Trial1, Trial2, Trial3), click Contrasts... to open the **Repeated Measures: Contrasts** window below. Select **Repeated** as the contrast from the **Contrast:** drop-down list and click Change. This contrast compares (1) the overall mean number of correct responses obtained under the Trial1, Trial2, and Trial3 conditions (i.e., collapsing across the three groups of Control, Effort, and Effort + Ability), and (2) the overall mean number of correct responses across the three trial conditions for the three groups (i.e., **TRIAL*GROUP** interaction). Select **None** as the contrast for the **GROUP** variable. Click Continue to return to the **Repeated Measures** window.

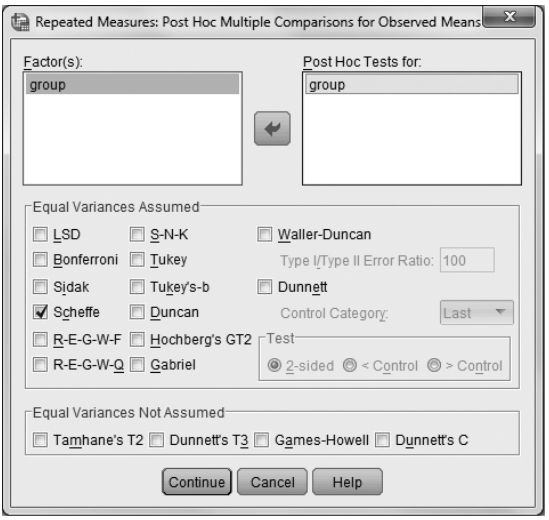


5. In the **Repeated Measures** window, click Options... to open the following **Repeated Measures: Options** window. In order to display means for the three **GROUPS** (Control, Effort, Effort + Ability), the three **TRIAL** levels (Trial1, Trial2, Trial3), and the **GROUP*TRIAL** interaction levels, click (highlight) **GROUP, TRIAL,** and **GROUP*TRIAL** in the **Factor(s) and Factor Interactions:** field, and then click ▶ to transfer these variables to the **Display Means for:** field.

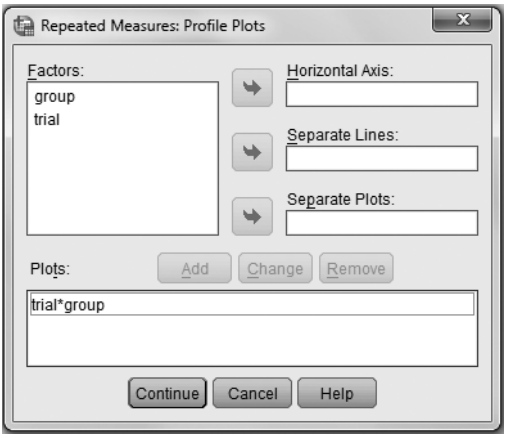
   In order to compute pairwise comparisons of the means for **TRIAL** (Trial1, Trial2, Trial3), check the **Compare main effects** field. However, because of inflated Type I error rate owing to the multiple comparisons, a **Bonferroni** type adjustment should be made. From the **Confidence interval adjustment** cell, select **Bonferroni** from the drop-down list. Click Continue to return to the **Repeated Measures** window.



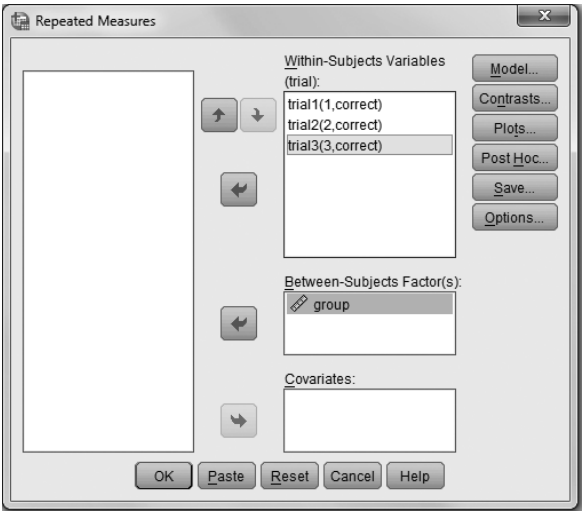
6. In order to test for differences between the three groups (Control, Effort, Effort + Ability), conduct post hoc comparisons by clicking Post Hoc... in the **Repeated Measures** window. This will open the **Repeated Measures: Post Hoc Multiple Comparisons for Observed Means** window. Transfer the **GROUP** variable to the **Post Hoc Tests for:** field by clicking it (highlight) and then clicking ▶. Check the **Scheffé** test field to run the Scheffé post hoc test. Next, click Continue to return to the **Repeated Measures** window.



7. In order to aid interpretation of the **TRIAL*GROUP** interaction effect, it would be useful to graph the means of this interaction. In the **Repeated Measures** window, click Plots... to open the **Repeated Measures: Profile Plots** window below. Transfer the **TRIAL** factor to the **Horizontal Axis:** field by clicking it (highlight) and then clicking ➔. Transfer the **GROUP** factor to the **Separate Lines:** field by clicking it (highlight) and then clicking ➔. Click Add to transfer this profile plot to the **Plots:** field. Click Continue to return to the **Repeated Measures** window.



8. When the **Repeated Measures** window opens, click OK to complete the analysis. See Table 9.3 for the results.



9.5.3 SPSS Syntax Method

```
GLM TRIAL1 TO TRIAL3 BY GROUP
/WSFACTOR = TRIAL 3 REPEATED
/MEASURE = CORRECT
/PLOT = PROFILE(TRIAL*GROUP)
/POSTHOC = GROUP(SCHEFFE)
/EMMEANS = TABLES(GROUP)
/EMMEANS = TABLES(TRIAL) COMPARE ADJ(BONFERRONI)
/EMMEANS = TABLES(GROUP*TRIAL).
```

9.5.4 SPSS Output

TABLE 9.3

Two-Factor Mixed Design Output

General Linear Model

Within-Subjects Factors

Measure: CORRECT

TRIAL	Dependent Variable
1	Trial 1
2	Trial 2
3	Trial 3

Between-Subjects Factors

		Value Label	N
Group	1.00	Control	11
	2.00	Effort	11
	3.00	Effort+Ability	11

Multivariate Tests[c]

Effect		Value	F	Hypothesis df	Error df	Sig.
TRIAL	Pillai's Trace	.867	94.318[a]	2.000	29.000	.000
	Wilks' Lambda	.133	94.318[a]	2.000	29.000	.000
	Hotelling's Trace	6.505	94.318[a]	2.000	29.000	.000
	Roy's Largest Root	6.505	94.318[a]	2.000	29.000	.000
TRIAL * group	Pillai's Trace	.553	5.738	4.000	60.000	.001
	Wilks' Lambda	.451	7.099[a]	4.000	58.000	.000
	Hotelling's Trace	1.210	8.470	4.000	56.000	.000
	Roy's Largest Root	1.202	18.037[a]	2.000	30.000	.000

[a] Exact statistic

[b] The statistic is an upper bound on F that yields a lower bound on the significance level.

[c] Design: Intercept+group

Within-Subjects Design: TRIAL

TABLE 9.3 (*Continued*)

Two-Factor Mixed Design Output

Mauchly's Test of Sphericity[b]

Measure: CORRECT

Within-Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[a] Greenhouse-Geisser	Huynh-Feldt	Lower-Bound
TRIAL	.507	19.712	2	.000	.670	.736	.500

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identify matrix.

[a] May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

[b] Design: Intercept+group
Within-Subjects Design: TRIAL

Tests of Within-Subjects Effects

Measure: CORRECT

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
TRIAL	Sphericity Assumed	170.081	2	85.040	52.619	.000
	Greenhouse-Geisser	170.081	1.339	126.986	52.619	.000
	Huynh-Feldt	170.081	1.472	115.515	52.619	.000
	Lower-Bound	170.081	1.000	170.081	52.619	.000
TRIAL * group	Sphericity Assumed	18.949	4	4.737	2.931	.028
	Greenhouse-Geisser	18.949	2.679	7.074	2.931	.050
	Huynh-Feldt	18.949	2.945	6.435	2.931	.045
	Lower-Bound	18.949	2.000	9.475	2.931	.069
Error(TRIAL)	Sphericity Assumed	96.970	60	1.616		
	Greenhouse-Geisser	96.970	40.181	2.413		
	Huynh-Feldt	96.970	44.171	2.195		
	Lower-Bound	96.970	30.000	3.232		

Tests of Within-Subjects Contrasts

Measure: CORRECT

Source	TRIAL	Type III Sum of Squares	df	Mean Square	F	Sig.
TRIAL	Level 1 vs. Level 2	161.485	1	161.485	167.579	.000
	Level 2 vs. Level 3	27.273	1	27.273	6.347	.017
TRIAL * group	Level 1 vs. Level 2	34.606	2	17.303	17.956	.000
	Level 2 vs. Level 3	3.818	2	1.909	.444	.645
Error(TRIAL)	Level 1 vs. Level 2	28.909	30	.964		
	Level 2 vs. Level 3	128.909	30	4.297		

(*Continued*)

TABLE 9.3 (*Continued*)

Two-Factor Mixed Design Output

Tests of Between-Subjects Effects

Measure: CORRECT

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig
Intercept	525.337	1	525.337	333.102	.000
Group	13.017	2	6.508	4.127	.026
Error	47.313	30	1.577		

Estimated Marginal Means

1. Group

Measure: CORRECT

			95% Confidence Interval	
Group	Mean	Std. Error	Lower Bound	Upper Bound
Control	3.364	.379	2.590	4.137
Effort	3.758	.379	2.984	4.531
Effort+Ability	4.848	.379	4.075	5.622

2. TRIAL

Estimates

Measure: CORRECT

			95% Confidence Interval	
TRIAL	Mean	Std. Error	Lower Bound	Upper Bound
1	2.212	.205	1.794	2.630
2	4.424	.235	3.945	4.904
3	5.333	.380	4.557	6.109

Pairwise Comparisons

Measure: CORRECT

(I) TRIAL	(J) TRIAL	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	−2.212*	.171	.000	−2.645	−1.779
	3	−3.121*	.367	.000	−4.051	−2.191
2	1	2.212*	.171	.000	1.779	2.645
	3	−.909	.361	.052	−1.824	.006
3	1	3.121*	.367	.000	2.191	4.051
	2	.909	.361	.052	−.006	1.824

Based on estimated marginal means.

* The mean difference is significant at the .05 level.

[a] Adjustment for multiple comparisons: Bonferroni.

TABLE 9.3 (*Continued*)

Two-Factor Mixed Design Output

Multivariate Tests

	Value	F	Hypothesis df	Error df	Sig.
Pillai's trace	.867	94.318[a]	2.000	29.000	.000
Wilks' lambda	.133	94.318[a]	2.000	29.000	.000
Hotelling's trace	6.505	94.318[a]	2.000	29.000	.000
Roy's largest root	6.505	94.318[a]	2.000	29.000	.000

Each F tests the multivariate effect of TRIAL. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

[a] Exact statistic

3. Group * TRIAL

Measure: CORRECT

				95% Confidence Interval	
Group	TRIAL	Mean	Std. Error	Lower Bound	Upper Bound
Control	1	2.182	.354	1.458	2.906
	2	3.455	.407	2.624	4.285
	3	4.455	.658	3.111	5.799
Effort	1	2.182	.354	1.458	2.906
	2	3.909	.407	3.079	4.739
	3	5.182	.658	3.838	6.526
Effort+Ability	1	2.273	.354	1.549	2.997
	2	5.909	.407	5.079	6.739
	3	6.364	.658	5.020	7.708

Post Hoc Tests Group

Multiple Comparisons

Measure: CORRECT

Scheffe

		Mean Difference	Std.		95% Confidence Interval	
(I) Group	(J) Group	(I-J)	Error	Sig.	Lower Bound	Upper Bound
Control	Effort	−.3939	.53549	.765	−1.7729	.9850
	Effort+Ability	−1.4848*	.53549	.033	−2.8638	−.1059
Effort	control	.3939	.53549	.765	−.9850	1.7729
	Effort+Ability	−1.0909	.53549	.143	−2.4699	.2881
Effort+Ability	Control	1.4848*	.53549	.033	.1059	2.8638
	Effort	1.0909	.53549	.143	−.2881	2.4699

Based on observed means.

* The mean difference is significant at the .05 level

(*Continued*)

TABLE 9.3 (*Continued*)

Two-Factor Mixed Design Output

Homogeneous Subsets			
CORRECT			
Scheffe[a,b,c]			
		Subset	
Group	N	1	2
Control	11	3.3636	
Effort	11	3.7576	3.7576
Effort+Ability	11		4.8485
Sig.		.765	.143

Means for groups in homogeneous subsets are displayed.
Based on Type III Sum of Squares.
The error term is Mean Square(Error) = 1.577.
[a] Uses Harmonic Mean Sample Size = 11.000.
[b] The group sizes are unequal. The harmonic mean of the group sizes is used. Type I error levels are not guaranteed.
[c] Alpha = .05

9.5.5 Results and Interpretation

9.5.5.1 Within-Subjects Effects

The **Multivariate Tests** test the within-subjects effects of **TRIAL** and the **TRIAL*GROUP** interaction. The decision to interpret either the **Multivariate Tests** statistics or the **Tests of Within-Subjects Effects** depends on the outcome of **Mauchly's Test of Sphericity** test. In this instance, the Mauchly value is 0.507 (chi-square approximate value of 19.71 with two degrees of freedom) and is significant ($p < .001$). Thus, the assumption of sphericity is violated and the multivariate statistics should be interpreted. (Again if the researcher chooses to interpret the **Tests of Within-Subjects Effects**, then an adjustment must be made to the numerator and denominator degrees of freedom, using the Huynh-Feldt epsilon. The **Tests of Within-Subjects Effects** table displays the significance of each test when sphericity is assumed, or when the Huynh-Feldt epsilon is used.)

In the **Multivariate Tests** table, the main effect for the within-subjects variable **TRIAL** is presented first and is significant ($p < .001$), on the basis of all four multivariate tests of significance (Pillai's, Wilks', Hotelling's, Roy's). From the cell means presented in the **Estimated Marginal Means** table of **TRIAL**, the results indicate that the subjects made the least correct responses in Trial1 ($M = 2.21$), more in Trial2 ($M = 4.42$), and the most correct responses in Trial3 ($M = 5.33$), averaged across the three groups.

The **Tests of Within-Subjects Contrasts** present the contrasts between the responses obtained across the three trials. The first contrast, LEVEL 1 vs. LEVEL 2, compares the number of correct responses made in Trial1

($M = 2.21$) with those made in Trial2 ($M = 4.42$), and is statistically significant, $F(1,30) = 167.58$, $p < .001$. The second contrast, LEVEL 2 vs. LEVEL 3, compares the number of correct responses made in Trial2 ($M = 4.42$) with the number of correct responses made in Trial3 ($M = 5.33$). The average value of this contrast is statistically significant, $F(1,30) = 6.35$, $p < .05$.

While the **Tests of Within-Subjects Contrasts** test for differences between adjacent levels within the within-subjects variable **TRIAL**, the **Pairwise Comparisons** (with Bonferroni correction) test for differences between *all* three levels. The comparison results presented in the **Pairwise Comparisons** table (under **Estimated Marginal Means**) indicate that the number of correct responses made in Trial1 is significantly different (lower) from Trial2 and Trial3 ($p < .001$). There is no significant difference between Trial2 and Trial3.

For the **TRIAL*GROUP** interaction, all four multivariate tests (Pillai's, Hotelling's, Wilks', Roy's) indicate that this interaction is statistically significant ($p < .05$), suggesting that the number of correct responses made across the three trials is dependent on the type of treatment groups (i.e., Control, Effort, Effort + Ability). In order to aid interpretation of the interaction effect, it would be useful to examine the graph presented in Figure 9.2. Please note that the graph was plotted from the mean number of correct responses presented in the **Estimated Marginal Means** table of **GROUP*TRIAL** interaction.

Figure 9.2 shows that the subjects' performances across the three tests are dependent on the type of feedback they received. While there is a general increase in the number of correct responses made across the three trials for

Profile Plots

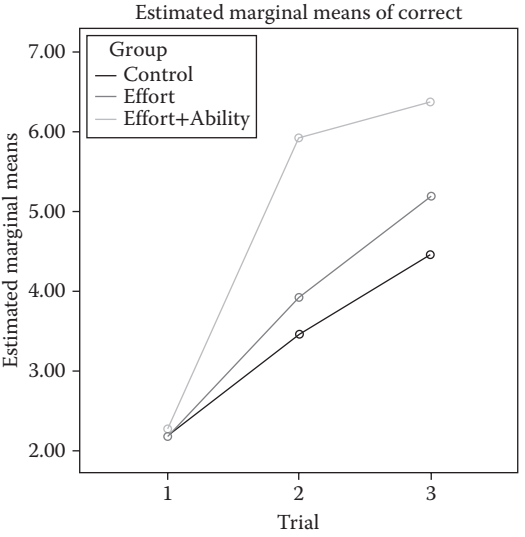


FIGURE 9.2
Trial × Group interaction.

all three groups, the rate of increase is greater for the Effort + Ability group from Trial1 to Trial2 than for the other two treatment groups.

The **Tests of Within-Subjects Contrasts** present the contrasts between the responses obtained across the three trials for the three groups. The first contrast, LEVEL 1 vs. LEVEL 2, is significant, $F(1,30) = 17.96$, $p < .001$, which indicates that the mean difference in the number of correct responses made between Trial1 and Trial2 is not the same for the three groups.

Mean Difference (Trial1 vs. Trial2)	
Control	−1.27 (2.18–3.45)
Effort	−1.73 (2.18–3.91)
Effort+Ability	−3.64 (2.27–5.91)

In conjunction with Figure 9.2, the results indicate that the increase in the number of correct responses made between Trial1 and Trial2 is different across the three groups; the increase is greatest for the Effort + Ability group (−3.64), less for the Effort group (−1.73), and least for the Control group (−1.27).

The second contrast, LEVEL 2 vs. LEVEL 3, is not significant ($p > .05$), which indicates that the mean difference in the number of correct responses made in Trial2 and Trial3 is the same for the three groups.

Mean Difference (Trial2 vs. Trial3)	
Control	−1.00 (3.45–4.45)
Effort	−1.27 (3.91–5.18)
Effort + Ability	−0.45 (5.91–6.36)

In conjunction with Figure 9.2, the results indicate that the increase in the number of correct responses made between Trial2 and Trial3 is similar across the three groups.

9.5.5.2 Between-Groups Effects

The **Tests of Between-Subjects Effects** is equivalent to a one-way analysis of variance. The results indicate that the between-groups variable **GROUP** is statistically significant, $F(2,30) = 4.13$, $p < .05$. From the cell means presented in the **Estimated Marginal Means** table of **GROUP**, the results indicate that the Control group made the least correct responses ($M = 3.36$), with the Effort group making more correct responses ($M = 3.76$), and the Effort+Ability group making the most correct responses ($M = 4.85$), averaged across the three trials.

The post hoc Scheffé multiple comparisons indicate that only the Control group ($M = 3.36$) and the Effort + Ability group ($M = 4.85$) differed significantly ($p < .05$), averaged across the three trials.

9.6 Example 4: GLM: Three-Factor Mixed Design (Two Between-Groups Variables and One Within-Subjects Variable)

In addition to investigating the effect of a drug assumed to increase learning ability, a researcher also wanted to investigate the effect of a subject's gender on performance over a series of trials. This study incorporates (1) the between-groups variables of drug-present versus drug-absent, and male versus female subjects, and (2) the within-subjects variables of Trial1, Trial2, Trial3, and Trial4. Twenty subjects have been randomly assigned to the four groups. The number of errors per trial is recorded.

	Trials			
Subjects	**T1**	**T2**	**T3**	**T4**
Drug-present				
Male				
s1	14	10	10	8
s2	16	9	10	5
s3	18	16	12	7
s4	17	14	11	8
s5	22	19	17	10
Female				
s6	18	19	17	15
s7	25	24	20	16
s8	37	35	32	24
s9	21	18	15	14
s10	29	29	28	26
Drug-absent				
Male				
s11	17	15	10	7
s12	24	12	9	9
s13	28	20	14	12
s14	30	22	16	12
s15	22	16	9	4
Female				
s16	42	34	30	22
s17	29	28	26	22
s18	26	25	20	16
s19	30	30	25	14
s20	26	27	20	12

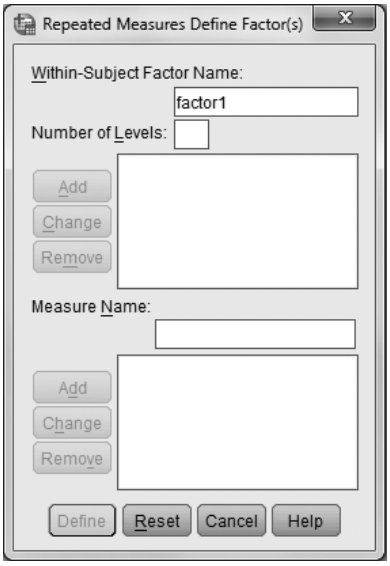
9.6.1 Data Entry Format

The data set has been saved under the name **EX9d.SAV.**

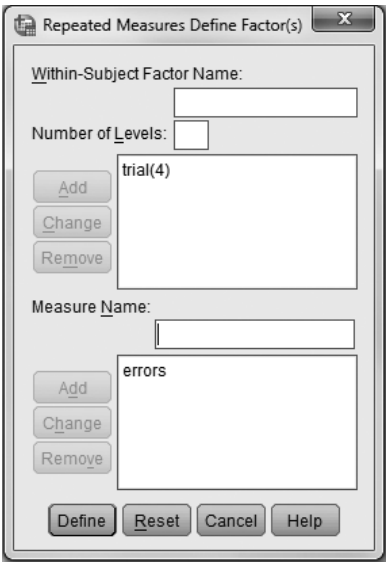
Variables	Column(s)	Code
Drug	1	1 = present, 2 = absent
Sex	2	1 = male, 2 = female
Trial1	3	Number of errors
Trial2	4	Number of errors
Trial3	5	Number of errors
Trial4	6	Number of errors

9.6.2 Windows Method

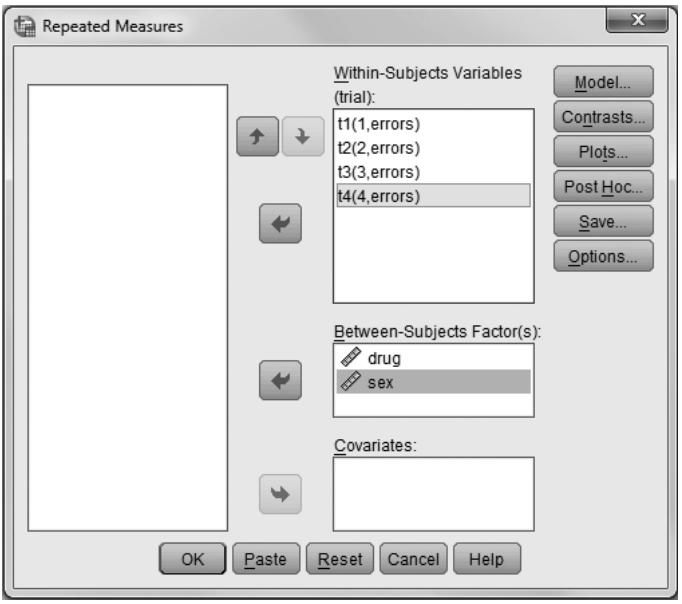
1. From the menu bar, click **Analyze**, then **General Linear Model**, and then **Repeated Measures**. The following **Repeated Measures Define Factor(s)** window will open.



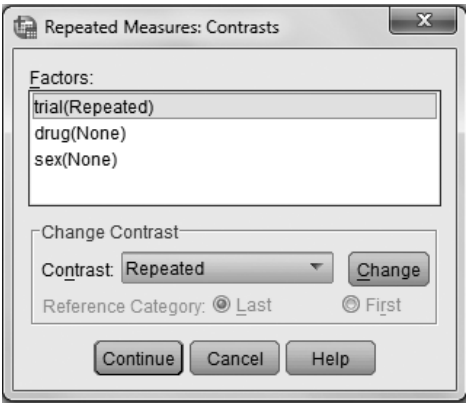
2. In the **Within-Subject Factor Name:** field, type **TRIAL** to denote the name of the within-subject factor. In the **Number of Levels:** field, type **4** to denote the four levels of the **TRIAL** factor (Trial1, Trial2, Trial3, Trial4). Click Add to transfer the **TRIAL** factor to the **ADD** field. In the **Measure Name**: field, type the name of the dependent measure, **ERRORS**, and then click Add.



3. Click Define to open the **Repeated Measures** window below. Transfer the four variables Trial1, Trial2, Trial3, and Trial4 to the **Within-Subjects Variables (trial):** field by clicking these variables (highlight) and then clicking ▶. Next, transfer the **DRUG** and **SEX** variables to the **Between-Subjects Factor(s):** field by clicking these variables (highlight) and then clicking ▶.



4. In order to test for differences among the four trial levels (Trial1, Trial2, Trial3, Trial4), click Contrasts... to open the **Repeated Measures: Contrasts** window below. Select **Repeated** as the contrast from the **Contrast:** drop-down list and click Change. This contrast compares (1) the overall mean number of error responses obtained under the Trial1, Trial2, Trial3, and Trial4 conditions (i.e., collapsing across the **DRUG** and **SEX** variables), (2) the overall mean number of error responses across the four trial conditions for the two **DRUG** groups (i.e., **TRIAL*DRUG** interaction—collapsing across the **SEX** variable), (3) the overall mean number of error responses across the four trial conditions for the two **SEX** groups (i.e., **TRIAL* SEX** interaction—collapsing across the **DRUG** variable), and (4) the overall mean number of error responses across the four trial conditions as a function of the **DRUG*SEX** interaction (i.e., **TRIAL*DRUG*SEX** interaction). Select **None** as the contrast for the **DRUG** and **SEX** variables.

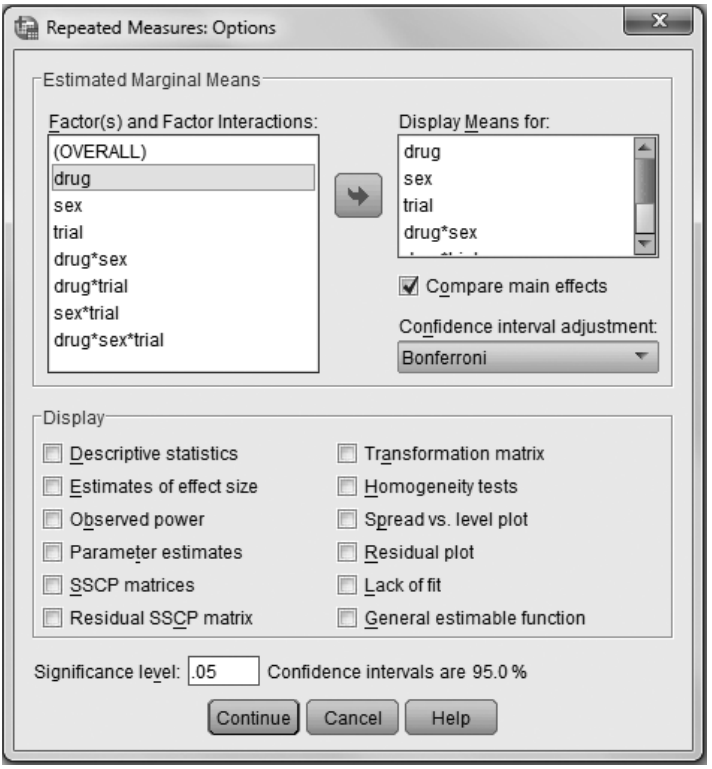


Click Continue to return to the **Repeated Measures** window.

5. In the **Repeated Measures** window, click Options... to open the following **Repeated Measures: Options** window. In order to display the mean number of errors for the two groups of **DRUG**, the two groups of **SEX**, the four **TRIAL** levels (Trial1, Trial2, Trial3, Trial4), the **DRUG*SEX** interaction, the **DRUG*TRIAL** interaction, the **SEX*TRIAL** interaction, and the **DRUG*SEX*TRIAL** interaction, click (highlight) **DRUG**, **SEX**, **TRIAL**, **DRUG*SEX**, **DRUG*TRIAL**, **SEX*TRIAL**, and **DRUG*SEX*TRIAL** in the **Factor(s) and Factor Interactions:** field, and then click ▶ to transfer these variables to the **Display Means for:** field.

   In order to compute pairwise comparisons of the means for **TRIAL** (Trial1, Trial2, Trial3, Trial4), check the **Compare main effects** field. However, because of inflated Type I error rate owing to the multiple

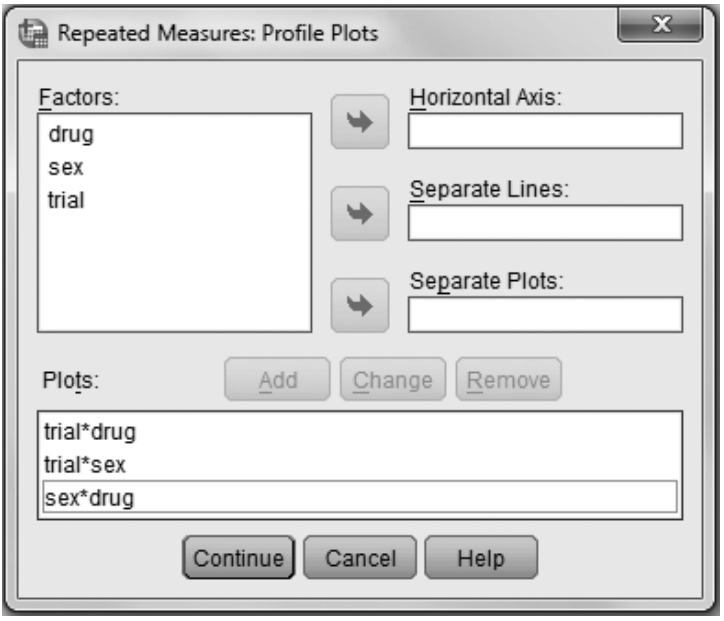
comparisons, a **Bonferroni** type adjustment should be made. From the **Confidence interval adjustment** cell, select **Bonferroni** from the drop-down list. Click Continue to return to the **Repeated Measures** window.



6. In order to aid interpretation of the **DRUG*SEX**, **DRUG*TRIAL**, and **SEX*TRIAL** interaction effects, it would be useful to graph the means of these interactions. In the **Repeated Measures** window, click Plots... to open the following **Repeated Measures: Profile Plots** window.
   - For the **DRUG*SEX** interaction, transfer the **SEX** factor to the **Horizontal Axis:** field by clicking it (highlight) and then clicking ▶. Transfer the **DRUG** factor to the **Separate Lines:** field by clicking it (highlight) and then clicking ▶. Click Add to transfer this profile plot to the **Plots:** field.
   - For the **TRIAL*DRUG** interaction, transfer the **TRIAL** factor to the **Horizontal Axis:** field by clicking it (highlight) and then clicking ▶. Transfer the **DRUG** factor to the **Separate Lines:**

field by clicking it (highlight) and then clicking ▶. Click Add to transfer this profile plot to the **Plots:** field.

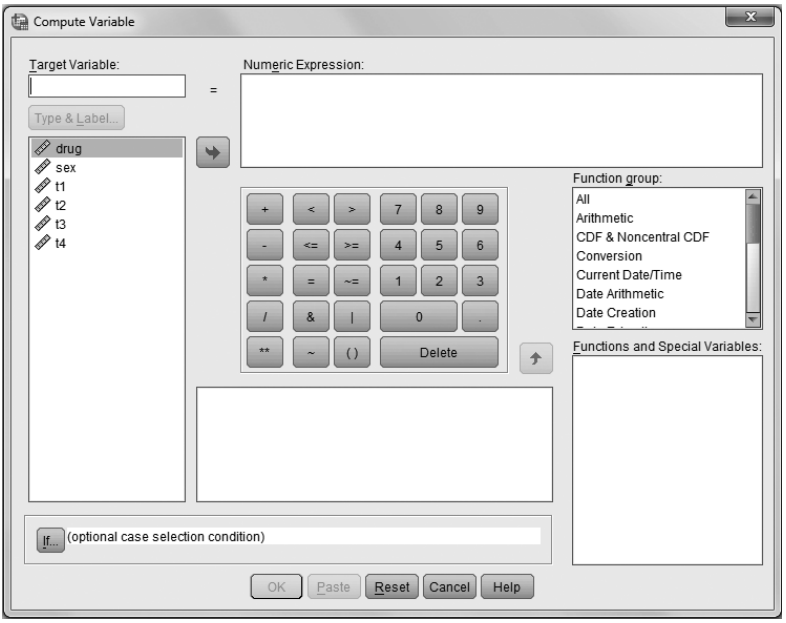
- For the **TRIAL*SEX** interaction, transfer the **TRIAL** factor to the **Horizontal Axis:** field by clicking it (highlight) and then clicking ▶. Transfer the **SEX** factor to the **Separate Lines:** field by clicking it (highlight) and then clicking ▶. Click Add to transfer this profile plot to the **Plots:** field.
- Click Continue to return to the **Repeated Measures** window.



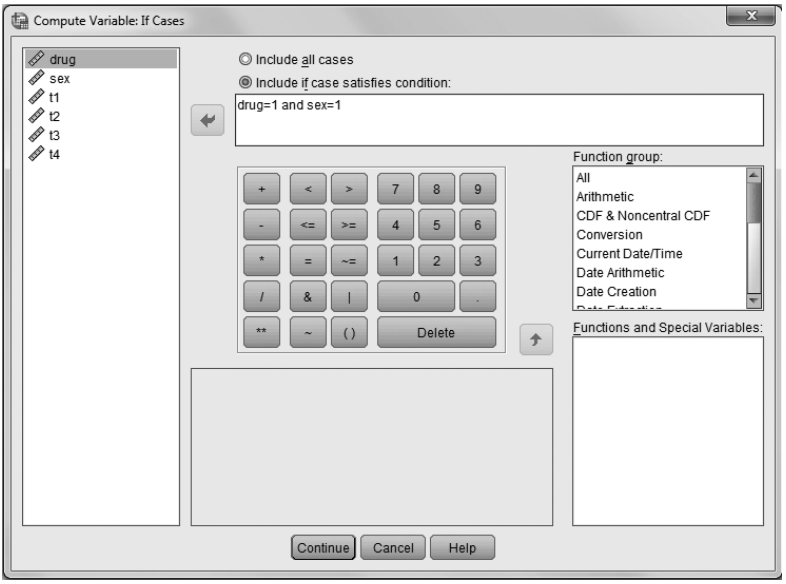
7. In order to interpret the full three-way interaction (**TRIAL*DRUG*SEX**), this interaction should also be graphed. This three-way interaction entails plotting the within-subjects factor of **TRIAL** against the factorial combination of the two between-groups factors **DRUG** and **SEX,** that is, against the four groups of **drug-present male, drug-present female, drug-absent male, drug-absent female**. To do this, some data transformation must be carried out.

7a. ***Data Transformation***

7a.1 The first step is to create a new grouping variable called **GROUP** that contains the four levels generated by the **DRUG*SEX** interaction (**drug-present male, drug-present female, drug-absent male, drug-absent female**). From the menu bar, click **Transform** and then **Compute Variable** to open the following **Compute Variable** window.

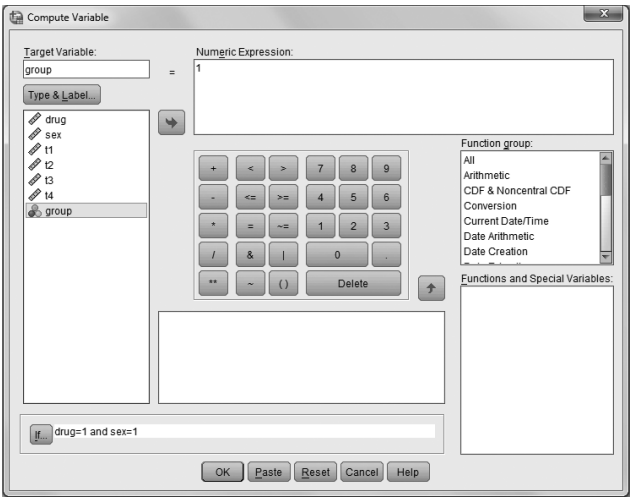


7a.2 Click If... to open the **Compute Variable: If Cases** window below. Ensure that the **Include if case satisfies condition:** field is checked. To create the first level (**drug-present male**), type **drug = 1 and sex = 1** in the field. Click Continue.



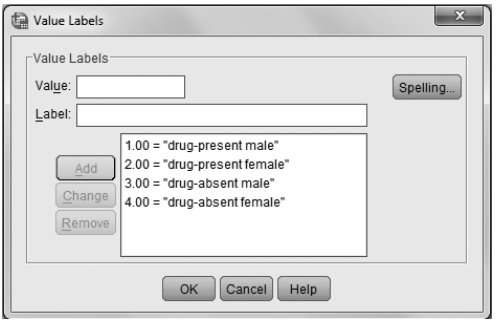
7a.3 Since the **drug-present male** level is the first of four levels within a new variable called **GROUP**, this level will be coded **1** within the **GROUP** variable.

When the **Compute Variable** window opens, type **GROUP** in the **Target Variable:** field, and **1** in the **Numeric Expression:** cell. Click OK to create the **drug-present male level** (coded 1) within the new grouping variable **GROUP**.

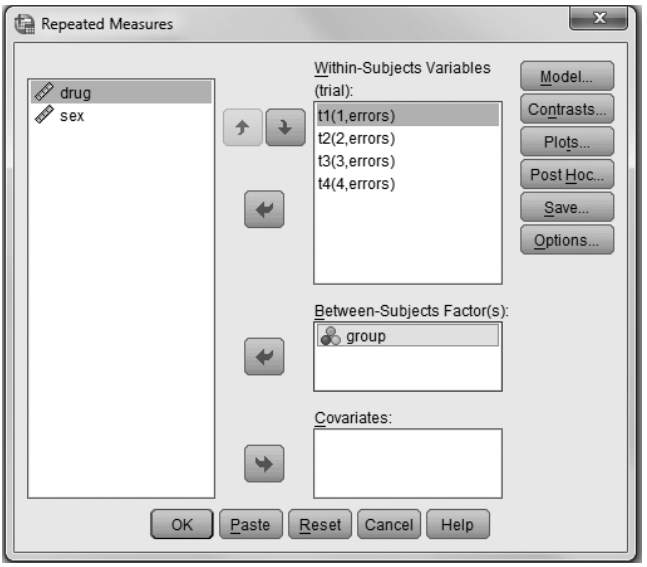


7a.4 Repeat steps 7a.1–7a.3 to create the other three levels: **drug-present female** (coded 2), **drug-absent male** (coded 3), **drug-absent female** (coded 4).

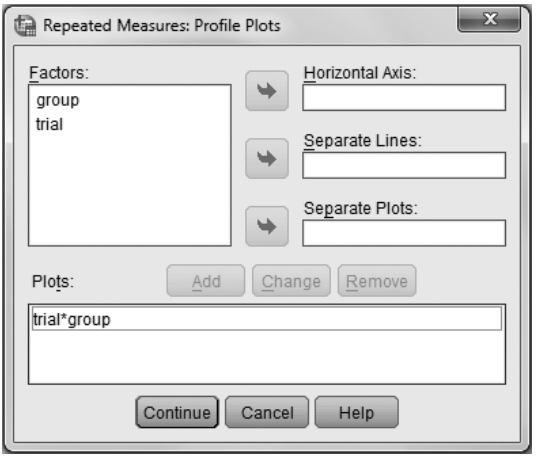
7a.5 To aid interpretation of the obtained results, **Value Labels** in the data set should be activated and labels attached to the numerical codes for the four levels. To do this, open the data set, and under **Variable View**, click the **Values** cell for the newly created **GROUP** variable. Type the value labels as indicated in the **Value Labels** window below. Click OK.



8. Open the **Repeated Measures** window by repeating step 1 to step 3. In the **Repeated Measures** window, transfer the **GROUP** variable to the **Between-Subjects Factor(s):** field.

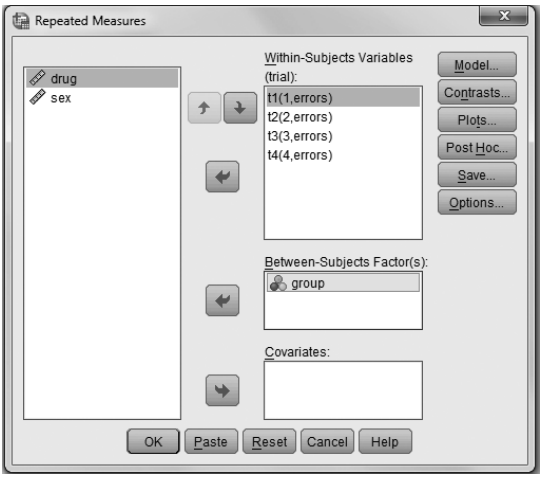


9. Click Plots... to open the **Repeated Measures: Profile Plots** window below.



To graph the full three-way **DRUG*SEX*TRIAL** interaction, transfer the **TRIAL** factor to the **Horizontal Axis:** field by clicking it (highlight) and then clicking ▶. Transfer the newly created **GROUP** factor to the **Separate Lines:** field by clicking it (highlight) and then

clicking ▶. Click Add to transfer this profile plot to the **Plots:** field. Click Continue to return to the **Repeated Measures** window.



10. When the **Repeated Measures** window opens, click OK to complete the analysis. See Table 9.4 for the results.

9.6.3 SPSS Syntax Method

```
GLM TRIAL1 TO TRIAL4 BY DRUG SEX
/WSFACTOR = TRIAL 4 REPEATED
/MEASURE = ERRORS
/EMMEANS = TABLES(DRUG)
/EMMEANS = TABLES(SEX)
/EMMEANS = TABLES(DRUG*SEX)
/EMMEANS = TABLES(TRIAL) COMPARE ADJ(BONFERRONI)
/EMMEANS = TABLES(DRUG*TRIAL)
/EMMEANS = TABLES(SEX*TRIAL)
/EMMEANS = TABLES(DRUG*SEX*TRIAL)
/PLOT = PROFILE(DRUG*SEX)
/PLOT = PROFILE(TRIAL*DRUG)
/PLOT = PROFILE(TRIAL*SEX).
IF (DRUG EQ 1 AND SEX EQ 1) GROUP = 1.
IF (DRUG EQ 2 AND SEX EQ 1) GROUP = 2.
IF (DRUG EQ 1 AND SEX EQ 2) GROUP = 3.
IF (DRUG EQ 2 AND SEX EQ 2) GROUP = 4.
VALUE LABELS GROUP 1 'DRUG PRESENT-MALE'
                   2 'DRUG ABSENT-MALE'
                   3 'DRUG PRESENT-FEMALE'
                   4 'DRUG ABSENT-FEMALE'.
```

```
GLM TRIAL1 TO TRIAL4 BY GROUP
/WSFACTOR = TRIAL 4 REPEATED
/MEASURE = ERRORS
/PLOT = PROFILE(TRIAL*GROUP)
/EMMEANS = TABLES(GROUP*TRIAL).
```

9.6.4 SPSS Output

TABLE 9.4

Three-Factor Mixed Design Output

General Linear Model

Within-Subjects Factors

Measure: ERRORS

TRIAL	Dependent Variable
1	trial1
2	trial2
3	trial3
4	trial4

Between-Subjects Factors

		Value Label	N
drug	1.00	present	10
	2.00	absent	10
sex	1.00	male	10
	2.00	female	10

Multivariate Tests[b]

Effect		Value	F	Hypothesis df	Error df	Sig.
TRIAL	Pillai's Trace	.927	59.694[a]	3.000	14.000	.000
	Wilks' Lambda	.073	59.694[a]	3.000	14.000	.000
	Hotelling's Trace	12.792	59.694[a]	3.000	14.000	.000
	Roy's Largest Root	12.792	59.694[a]	3.000	14.000	.000
TRIAL * drug	Pillai's Trace	.639	8.259[a]	3.000	14.000	.002
	Wilks' Lambda	.361	8.259[a]	3.000	14.000	.002
	Hotelling's Trace	1.770	8.259[a]	3.000	14.000	.002
	Roy's Largest Root	1.770	8.259[a]	3.000	14.000	.002

(*Continued*)

TABLE 9.4 (*Continued*)

Three-Factor Mixed Design Output

Multivariate Tests[b]

Effect		Value	F	Hypothesis df	Error df	Sig.
TRIAL * sex	Pillai's Trace	.488	4.448[a]	3.000	14.000	.021
	Wilks' Lambda	.512	4.448[a]	3.000	14.000	.021
	Hotelling's Trace	.953	4.448[a]	3.000	14.000	.021
	Roy's Largest Root	.953	4.448[a]	3.000	14.000	.021
TRIAL * drug * sex	Pillai's Trace	.483	4.366[a]	3.000	14.000	.023
	Wilks' Lambda	.517	4.366[a]	3.000	14.000	.023
	Hotelling's Trace	.936	4.366[a]	3.000	14.000	.023
	Roy's Largest Root	.936	4.366[a]	3.000	14.000	.023

[a] Exact statistic

[b] Design: Intercept+drug+sex+drug * sex
Within-Subjects Design: TRIAL

Mauchly's Test of Sphericity[b]

Measure: ERRORS

Within-Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[a] Greenhouse-Geisser	Huynh-Feldt	Lower-Bound
TRIAL	.298	17.816	5	.003	.674	.917	.333

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identify matrix.

[a] May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

[b] Design: Intercept+drug+sex+drug * sex
Within-Subjects Design: TRIAL

Tests of Within-Subjects Effects

Measure: ERRORS

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
TRIAL	Sphericity Assumed	1430.137	3	476.712	111.295	.000
	Greenhouse-Geisser	1430.137	2.022	707.140	111.295	.000
	Huynh-Feldt	1430.137	2.751	519.912	111.295	.000
	Lower-Bound	1430.137	1.000	1430.137	111.295	.000

TABLE 9.4 (*Continued*)

Three-Factor Mixed Design Output

Tests of Within-Subjects Effects

Measure: ERRORS

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
TRIAL * drug	Sphericity Assumed	112.538	3	37.513	8.758	.000
	Greenhouse-Geisser	112.538	2.022	55.645	8.758	.001
	Huynh-Feldt	112.538	2.751	40.912	8.758	.000
	Lower-Bound	112.538	1.000	112.538	8.758	.009
TRIAL * sex	Sphericity Assumed	55.038	3	18.346	4.283	.009
	Greenhouse-Geisser	55.038	2.022	27.214	4.283	.022
	Huynh-Feldt	55.038	2.751	20.008	4.283	.012
	Lower-Bound	55.038	1.000	55.038	4.283	.055
TRIAL * drug * sex	Sphericity Assumed	21.438	3	7.146	1.668	.186
	Greenhouse-Geisser	21.438	2.022	10.600	1.668	.204
	Huynh-Feldt	21.438	2.751	7.793	1.668	.191
	Lower-Bound	21.438	1.000	21.438	1.668	.215
Error(TRIAL)	Sphericity Assumed	205.600	48	4.283		
	Greenhouse-Geisser	205.600	32.359	6.354		
	Huynh-Feldt	205.600	44.012	4.671		
	Lower-Bound	205.600	16.000	12.850		

Tests of Within-Subjects Contrasts

Measure: ERRORS

Source	TRIAL	Type III Sum of Squares	df	Mean Square	F	Sig.
TRIAL	Level 1 vs. Level 2	238.050	1	238.050	29.664	.000
	Level 2 vs. Level 3	252.050	1	252.050	90.018	.000
	Level 3 vs. Level 4	387.200	1	387.200	63.737	.000
TRIAL * drug	Level 1 vs. Level 2	22.050	1	22.050	2.748	.117
	Level 2 vs. Level 3	42.050	1	42.050	15.018	.001
	Level 3 vs. Level 4	5.000	1	5.000	.823	.378

(*Continued*)

TABLE 9.4 (*Continued*)

Three-Factor Mixed Design Output

Tests of Within-Subjects Contrasts

Measure: ERRORS

Source	TRIAL	Type III Sum of Squares	df	Mean Square	F	Sig.
TRIAL * sex	Level 1 vs. Level 2	84.050	1	84.050	10.474	.005
	Level 2 vs. Level 3	.050	1	.050	.018	.895
	Level 3 vs. Level 4	12.800	1	12.800	2.107	.166
TRIAL * drug * sex	Level 1 vs. Level 2	8.450	1	8.450	1.053	.320
	Level 2 vs. Level 3	4.050	1	4.050	1.446	.247
	Level 3 vs. Level 4	33.800	1	33.800	5.564	.031
Error(TRIAL)	Level 1 vs. Level 2	128.400	16	8.025		
	Level 2 vs. Level 3	44.800	16	2.800		
	Level 3 vs. Level 4	97.200	16	6.075		

Tests of Between-Subjects Effects

Measure: ERRORS

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig
Intercept	7286.653	1	7286.653	347.604	.000
Drug	29.403	1	29.403	1.403	.254
Sex	512.578	1	512.578	24.452	.000
Drug * sex	.528	1	.528	.025	.876
Error	335.400	16	20.963		

Estimated Marginal Means

1. Drug

Measure: ERRORS

			95% Confidence Interval	
Drug	Mean	Std. Error	Lower Bound	Upper Bound
Present	17.875	1.448	14.806	20.944
Absent	20.300	1.448	17.231	23.369

TABLE 9.4 (*Continued*)

Three-Factor Mixed Design Output

2. Sex

Measure: ERRORS

			95% Confidence Interval	
Sex	**Mean**	**Std. Error**	**Lower Bound**	**Upper Bound**
Male	14.025	1.448	10.956	17.094
Female	24.150	1.448	21.081	27.219

3. Drug * Sex

Measure: ERRORS

				95% Confidence Interval	
Drug	**Sex**	**Mean**	**Std. Error**	**Lower Bound**	**Upper Bound**
Present	Male	12.650	2.048	8.309	16.991
	Female	23.100	2.048	18.759	27.441
Absent	Male	15.400	2.048	11.059	19.741
	Female	25.200	2.048	20.859	29.541

4. TRIAL

Estimates

Measure: ERRORS

			95% Confidence Interval	
TRIAL	**Mean**	**Std. Error**	**Lower Bound**	**Upper Bound**
1	24.500	1.293	21.808	27.292
2	21.100	1.093	18.784	23.416
3	17.550	1.062	15.298	19.802
4	13.150	.917	11.207	15.093

Pairwise Comparisons

Measure: ERRORS

					95% Confidence Interval for Difference[a]	
(I) TRIAL	**(J) TRIAL**	**Mean Difference (I-J)**	**Std. Error**	**Sig.[a]**	**Lower Bound**	**Upper Bound**
1	2	3.450*	.633	.000	1.544	5.356
	3	7.000*	.594	.000	5.214	8.786
	4	11.400*	.868	.000	8.788	14.012
2	1	−3.450*	.633	.000	−5.356	−1.544
	3	3.550*	.374	.000	2.424	4.676
	4	7.950*	.787	.000	5.584	10.316

(*Continued*)

TABLE 9.4 (*Continued*)

Three-Factor Mixed Design Output

Pairwise Comparisons

Measure: ERRORS

(I) TRIAL	(J) TRIAL	Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
3	1	−7.000*	.594	.000	−8.786	−5.214
	2	−3.550*	.374	.000	−4.676	−2.424
	4	4.400*	.551	.000	2.742	6.058
4	1	−11.400*	.868	.000	−14.012	−8.788
	2	−7.950*	.787	.000	−10.316	−5.584
	3	−4.400*	.551	.000	−6.058	−2.742

Based on estimated marginal means.

* The mean difference is significant at the .05 level.

[b] Adjustment for multiple comparisons: Bonferroni.

Multivariate Tests

	Value	F	Hypothesis df	Error df	Sig.
Pillai's trace	.927	59.694[a]	3.000	14.000	.000
Wilks' lambda	.073	59.694[a]	3.000	14.000	.000
Hotelling's trace	12.792	59.694[a]	3.000	14.000	.000
Roy's largest root	12.792	59.694[a]	3.000	14.000	.000

Each F tests the multivariate effect of TRIAL. These tests are based on the linearly independent pairwise comparisons among the estimated marginal means.

[a] Exact statistic

5. Drug * TRIAL

Measure: ERRORS

Drug	TRIAL	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Present	1	21.700	1.829	17.823	25.577
	2	19.300	1.545	16.024	22.576
	3	17.200	1.502	14.015	20.385
	4	13.300	1.296	10.552	16.048
Absent	1	27.400	1.829	23.523	31.277
	2	22.900	1.545	19.624	26.176
	3	17.900	1.502	14.715	21.085
	4	13.000	1.296	10.252	15.748

TABLE 9.4 (*Continued*)

Three-Factor Mixed Design Output

6. Sex * TRIAL

Measure: ERRORS

				95% Confidence Interval	
Sex	**TRIAL**	**Mean**	**Std. Error**	**Lower Bound**	**Upper Bound**
Male	1	20.800	1.829	16.923	24.677
	2	15.300	1.545	12.024	18.576
	3	11.800	1.502	8.615	14.985
	4	8.200	1.296	5.452	10.948
Female	1	28.300	1.829	24.423	32.177
	2	26.900	1.545	23.624	30.176
	3	23.300	1.502	20.115	26.485
	4	18.100	1.296	15.352	20.848

7. Drug * sex * TRIAL

Measure: ERRORS

					95% Confidence Interval	
Drug	**Sex**	**TRIAL**	**Mean**	**Std. Error**	**Lower Bound**	**Upper Bound**
Present	Male	1	17.400	2.587	11.917	22.883
		2	13.600	2.185	8.968	18.232
		3	12.000	2.125	7.496	16.504
		4	7.600	1.833	3.714	11.486
	Female	1	26.000	2.587	20.517	31.483
		2	25.000	2.185	20.368	29.632
		3	22.400	2.125	17.896	26.904
		4	19.000	1.833	15.114	22.886
Absent	Male	1	24.200	2.587	18.717	29.683
		2	17.000	2.185	12.368	21.632
		3	11.600	2.125	7.096	16.104
		4	8.800	1.833	4.914	12.686
	Female	1	30.600	2.587	25.117	36.083
		2	28.800	2.185	24.168	33.432
		3	24.200	2.125	19.696	28.704
		4	17.200	1.833	13.314	21.086

(*Continued*)

TABLE 9.4 (*Continued*)

Three-Factor Mixed Design Output

General Linear Model

Within-Subjects Factors

Measure: ERRORS

TRIAL	Dependent Variable
1	trial1
2	trial2
3	trial3
4	trial4

Between-Subjects Factors

		Value Lable	N
GROUP	1.00	DRUG PRESENT-MALE	5
	2.00	DRUG ABSENT-MALE	5
	3.00	DRUG PRESENT-FEMALE	5
	4.00	DRUG ABSENT-FEMALE	5

Multivariate Tests[c]

Effect		Value	F	Hypothesis df	Error df	Sig.
TRIAL	Pillai's Trace	.927	59.694[a]	3.000	14.000	.000
	Wilks' Lambda	.073	59.694[a]	3.000	14.000	.000
	Hotelling's Trace	12.792	59.694[a]	3.000	14.000	.000
	Roy's Largest Root	12.792	59.694[a]	3.000	14.000	.000
TRIAL * GROUP	Pillai's Trace	1.267	3.901	9.000	48.000	.001
	Wilks' Lambda	.138	4.785	9.000	34.223	.000
	Hotelling's Trace	3.659	5.149	9.000	38.000	.000
	Roy's Largest Root	2.924	15.595[b]	3.000	16.000	.000

[a] Exact statistic

[b] The statistic is an upper bound on F that yields a lower bound on the significance level.

[c] Design: Intercept+GROUP
Within-Subjects Design: TRIAL

TABLE 9.4 (*Continued*)

Three-Factor Mixed Design Output

Mauchly's Test of Sphericity[b]

Measure: ERRORS

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon[a] Greenhouse-Geisser	Huynh-Feldt	Lower-bound
TRIAL	.298	17.816	5	.003	.674	.917	.333

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables is proportional to an identify matrix.

[a] May be used to adjust the degrees of freedom for the averaged tests of significance. Corrected tests are displayed in the Tests of Within-Subjects Effects table.

[b] Design: Intercept+GROUP
Within-Subjects Design: TRIAL

Tests of Within-Subjects Effects

Measure: ERRORS

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
TRIAL	Sphericity Assumed	1430.137	3	476.712	111.295	.000
	Greenhouse-Geisser	1430.137	2.022	707.140	111.295	.000
	Huynh-Feldt	1430.137	2.751	519.912	111.295	.000
	Lower-Bound	1430.137	1.000	1430.137	111.295	.000
TRIAL * GROUP	Sphericity Assumed	189.013	9	21.001	4.903	.000
	Greenhouse-Geisser	189.013	6.067	31.153	4.903	.001
	Huynh-Feldt	189.013	8.252	22.905	4.903	.000
	Lower-Bound	189.013	3.000	63.004	4.903	.013
Error(TRIAL)	Sphericity Assumed	205.600	48	4.283		
	Greenhouse-Geisser	205.600	32.359	6.354		
	Huynh-Feldt	205.600	44.012	4.671		
	Lower-Bound	205.600	16.000	12.850		

TABLE 9.4 (*Continued*)

Three-Factor Mixed Design Output

Tests of Within-Subjects Contrasts

Measure: ERRORS

Source	TRIAL	Type III Sum of Squares	df	Mean Square	F	Sig.
TRIAL	Level 1 vs. Level 2	238.050	1	238.050	29.664	.000
	Level 2 vs. Level 3	252.050	1	252.050	90.018	.000
	Level 3 vs. Level 4	387.200	1	387.200	63.737	.000
TRIAL * GROUP	Level 1 vs. Level 2	114.550	3	38.183	4.758	.015
	Level 2 vs. Level 3	46.150	3	15.383	5.494	.009
	Level 3 vs. Level 4	51.600	3	17.200	2.831	.071
Error(TRIAL)	Level 1 vs. Level 2	128.400	16	8.025		
	Level 2 vs. Level 3	44.800	16	2.800		
	Level 3 vs. Level 4	97.200	16	6.075		

Tests of Between-Subjects Effects

Measure: ERRORS

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	7286.653	1	7286.653	347.604	.000
GROUP	542.509	3	180.836	8.627	.001
Error	335.400	16	20.963		

Estimated Marginal Means

GROUP * TRIAL

Measure: ERRORS

GROUP	TRIAL	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
DRUG PRESENT–MALE	1	17.400	2.587	11.917	22.883
	2	13.600	2.185	8.968	18.232
	3	12.000	2.125	7.496	16.504
	4	7.600	1.833	3.714	11.486
DRUG ABSENT–MALE	1	24.200	2.587	18.717	29.683
	2	17.000	2.185	12.368	21.632
	3	11.600	2.125	7.096	16.104
	4	8.800	1.833	4.914	12.686

TABLE 9.4 (*Continued*)

Three-Factor Mixed Design Output

Estimated Marginal Means

GROUP * TRIAL

Measure: ERRORS

				95% Confidence Interval	
GROUP	**TRIAL**	**Mean**	**Std. Error**	**Lower Bound**	**Upper Bound**
DRUG PRESENT–FEMALE	1	26.000	2.587	20.517	31.483
	2	25.000	2.185	20.368	29.632
	3	22.400	2.125	17.896	26.904
	4	19.000	1.833	15.114	22.886
DRUG ABSENT–FEMALE	1	30.600	2.587	25.117	36.083
	2	28.800	2.185	24.168	33.432
	3	24.200	2.125	19.696	28.704
	4	17.200	1.833	13.314	21.086

9.6.5 Results and Interpretation

9.6.5.1 Within-Subjects Effects

As **Mauchly's Test of Sphericity** is statistically significant ($p < .01$), the assumption of sphericity is violated, and the multivariate statistics should be interpreted. The **Multivariate Tests** test the within-subjects effects of the **TRIAL**, **TRIAL*DRUG**, **TRIAL*SEX**, and **TRIAL*DRUG*SEX** interactions.

9.6.5.1.1 TRIAL Main Effect

In the **Multivariate Tests** table, the main effect for the within-subjects variable **TRIAL** is presented first, and is significant, on the basis of all four multivariate tests of significance (Pillai's, Wilks', Hotelling's, Roy's), $F(3,14) = 59.69$, $p < .001$. From the cell means presented in the **Estimated Marginal Means** table of **TRIAL**, the results indicate that the subjects made progressively fewer errors across the four trials, averaged across the two **DRUG** groups and the two **SEX** groups (Trial1 = 24.55, Trial2 = 21.10, Trial3 = 17.55, Trial4 = 13.15).

The **Tests of Within-Subjects Contrasts** present the contrasts between the number of errors made across the four trials. The first contrast, LEVEL 1 vs. LEVEL 2, compares the number of errors made in Trial1 ($M = 24.55$) with those made in Trial2 ($M = 21.10$) and is statistically significant, $F(1,16) = 29.66$, $p < .001$. This suggests that the subjects made significantly fewer errors in Trial2 relative to Trial1. The second contrast, LEVEL 2 vs. LEVEL 3, compares the number of errors made in Trial2 ($M = 21.10$) with the number of errors made in Trial3 ($M = 17.55$) and is statistically significant, $F(1,16) = 90.02$, $p < .001$. This indicates that the subjects made significantly fewer errors in Trial3 than in Trial2. The third contrast, LEVEL 3 vs. LEVEL 4, compares the number of errors made in

Trial3 ($M = 17.55$) with the number of errors made in Trial4 ($M = 13.15$) and is statistically significant, $F(1,16) = 63.74$, $p < .001$. This indicates that the subjects made significantly fewer errors in Trial4 than in Trial3.

While the **Tests of Within-Subjects Contrasts** test for differences between *adjacent levels* within the within-subjects variable **TRIAL**, the **Pairwise Comparisons** (with Bonferroni correction) test for differences between all four levels. The comparison results presented in the **Pairwise Comparisons** table (under **Estimated Marginal Means**) indicate that (1) the number of errors made in Trial1 is significantly different (higher) from Trial2, Trial3, and Trial4 ($p < .001$), (2) the number of errors made in Trial2 is significantly different (higher) from Trial3 and Trial4 ($p < .001$), and (3) the number of errors made in Trial3 is significantly different (higher) from Trial4 ($p < .001$).

*9.6.5.1.2 TRIAL*DRUG*

For the **TRIAL*DRUG** interaction, all four multivariate tests of significance (Pillai's, Hotelling's, Wilks', Roy's) indicate that this interaction is statistically significant, $F(3,14) = 8.26$, $p < .01$, suggesting that the number of errors made across the four trials is dependent upon the presence or absence of the drug. In order to aid interpretation of the interaction effect, it is useful to analyze the graph shown in Figure 9.4. Note that the graph was plotted from the mean number of errors presented in the **Estimated Marginal Means** table of **DRUG*TRIAL** interaction.

Figure 9.4 shows that in both drug conditions (present/absent), the number of errors made decreased from Trial1 to Trial4. Nevertheless, the rate of

Profile Plots

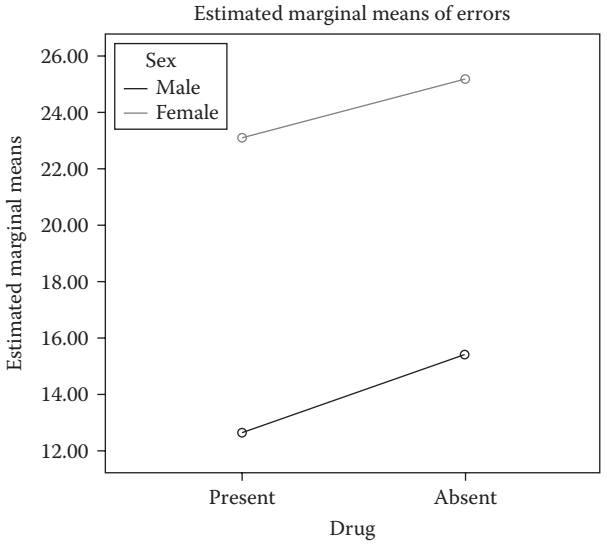


FIGURE 9.3
Drug × Sex interaction

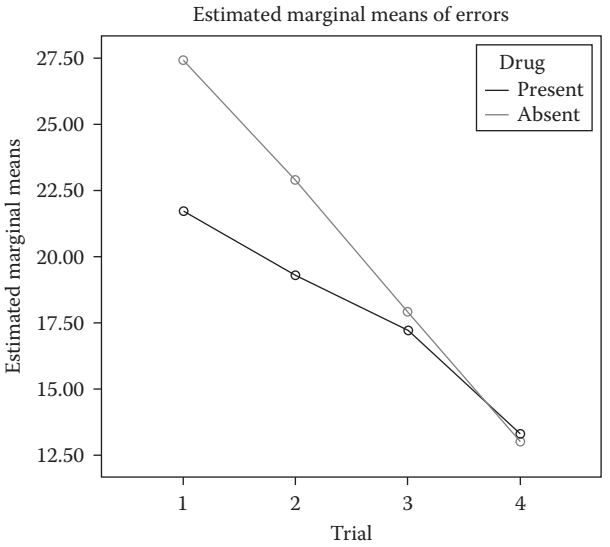


FIGURE 9.4
Trial × Drug interaction.

decrease is different, with subjects showing a greater rate of decrease in the drug-absent condition across the four tests.

The **Tests of Within-Subjects Contrasts** present the contrasts for the number of errors made across the four trials for the two **DRUG** groups. The first contrast, LEVEL 1 vs. LEVEL 2, compares the difference in the number of errors made between Trial1 and Trial2 for the two DRUG groups and is not significant, $F(1,16) = 2.75$, $p > .05$. In conjunction with Figure 9.4, the results indicate that the decrease in the number of errors made between Trial1 and Trial2 is similar for the two **DRUG** groups.

Mean Difference (Trial1 vs. Trial2)	
Drug-present	2.40 (21.70–19.30)
Drug-absent	4.50 (27.40–22.90)

The second contrast, LEVEL 2 vs. LEVEL 3, compares the difference in the number of errors made between Trial2 and Trial3 for the two DRUG groups, and is significant, $F(1,16) = 15.02$, $p < .01$. In conjunction with Figure 9.4, the results indicate that the reduction in the number of errors made between Trial2 and Trial3 is different for the two DRUG groups; the decrease is greater for the drug-absent group.

Mean Difference (Trial2 vs. Trial3)	
Drug-present	2.10 (19.30–17.20)
Drug-absent	5.00 (22.90–17.90)

The third contrast, LEVEL 3 vs. LEVEL 4, compares the difference in the number of errors made between Trial3 and Trial4 for the two DRUG groups, and is not significant, $F(1,16) = 0.82$, $p > .05$. In conjunction with Figure 9.4, the results indicate that the reduction in the number of errors made between Trial3 and Trial4 is similar for the two DRUG groups.

Mean Difference (Trial3 vs. Trial4)	
Drug-present	3.90 (17.20–13.30)
Drug-absent	4.90 (17.90–13.00)

*9.6.5.1.3 TRIAL*SEX*

For the **TRIAL*SEX** interaction, all four multivariate tests of significance (Pillai's, Hotelling's, Wilks', Roy's) indicate that this interaction is statistically significant, $F(3,14) = 4.45$, $p < .05$, suggesting that the number of errors made across the four trials is dependent upon subjects' gender. In order to aid interpretation of the interaction effect, it is useful to analyze the graph presented as Figure 9.5. Please note that the graph was plotted from the mean number of errors presented in the **Estimated Marginal Means** table of the **SEX*TRIAL** interaction.

Figure 9.5 shows that for both male and female subjects, the number of errors made decreased from Trial1 to Trial4. Nevertheless, the rate of decrease is different, with males showing a greater rate of decrease from Trial1 to Trial2.

The **Tests of Within-Subjects Contrasts** present the contrasts for the number of errors made across the four trials for the male and female groups. The first contrast, LEVEL 1 vs. LEVEL 2, is significant, $F(1,16) = 10.47$, $p < .001$. In conjunction with Figure 9.5, the results indicate that the decrease in the number of errors made between Trial1 and Trial2 is different for the male and female groups; the decrease is greater for the male group.

Mean Difference (Trial1 vs. Trial2)	
Male	5.50 (20.80–15.30)
Female	1.40 (28.30–26.90)

The second contrast, LEVEL 2 vs. LEVEL 3, is not significant, $F(1,16) = 0.02$, $p > .05$. In conjunction with Figure 9.5, the results indicate that the reduction in the number of errors made between Trial2 and Trial3 is similar for the male and female groups.

Mean Difference (Trial2 vs. Trial3)	
Male	3.50 (15.30–11.80)
Female	3.60 (26.90–23.30)

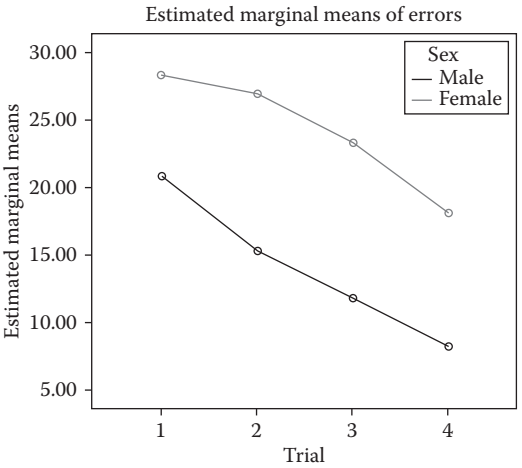


FIGURE 9.5
Trial × Sex interaction.

The third contrast, LEVEL 3 vs. LEVEL 4, compares the number of errors made in Trial3 with the number of errors made in Trial4, and is not significant, $F(1,16) = 2.11$, $p > .05$. In conjunction with Figure 9.5, the results indicate that the reduction in the number of errors made between Trial3 and Trial4 is similar for both the male and female groups.

Mean Difference (Trial3 vs. Trial4)	
Male	3.60 (11.80–8.20)
Female	5.20 (23.30–18.10)

*9.6.5.1.4 TRIAL*DRUG*SEX*

For the three-way **TRIAL*DRUG*SEX** interaction (presented as **TRIAL*GROUP** in the SPSS output), all four multivariate tests (Pillai's, Hotelling's, Wilks', Roy's) indicate that this interaction is statistically significant ($p < .01$), suggesting that the number of errors made across the four trials is dependent on the **DRUG*SEX** interaction. In order to aid interpretation of the interaction effect, it is useful to analyze the graph presented as Figure 9.6. Please note that the graph was plotted from the mean number of errors presented in the **Estimated Marginal Means** table of the **GROUP*TRIAL** interaction.

Figure 9.6 shows that female subjects made more errors under the drug-absent condition than under the drug-present condition across the first three trials. However, for the last trial, the effect is opposite with more errors made under the drug-present condition than under the drug-absent condition. For male subjects, the response pattern is similar to that of the female subjects across the first two trials (i.e., more errors were made under the drug-absent condition than under the drug-present condition). However, for

Profile Plots

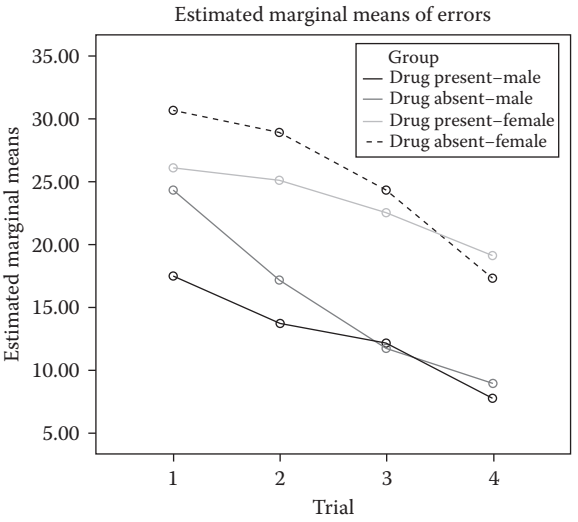


FIGURE 9.6
Trial × Drug × Sex interaction.

Mean Difference (Trial1 vs. Trial2)	
Drug-present male	3.80 (17.40–13.60)
Drug-absent male	7.20 (24.20–17.00)
Drug-present female	1.00 (26.00–25.00)
Drug-absent female	1.80 (30.60–28.80)
Mean Difference (Trial2 vs. Trial3)	
Drug-present male	1.60 (13.60–12.00)
Drug-absent male	5.40 (17.00–11.60)
Drug-present female	2.60 (25.00–22.40)
Drug-absent female	4.60 (28.80–24.20)

the last two trials, there is very little difference in the number of errors made under these two conditions.

For this three-way interaction, the **Tests of Within-Subjects Contrasts** showed that the first contrast, LEVEL 1 vs. LEVEL 2, and the second contrast, LEVEL 2 vs. LEVEL 3, are significant, $F(3,16) = 4.76$, $p < .05$ and $F(3,16) = 5.49$, $p < .01$, respectively. In conjunction with Figure 9.6, the results indicate that the mean differences in the number of errors made between Trial1 and Trial2, and between Trial2 and Trial3, are different for the four groups (**drug-present male, drug-present female, drug-absent male, drug-absent female).** For the male subjects, the decrease in the number of errors made from Trial1 to Trial2 is greater under the drug-absent condition ($M = 7.20$) than under the drug-present condition ($M = 3.80$). For the female subjects, the decrease in the

Mean Difference (Trial3 vs. Trial4)	
Drug-present male	4.40 (12.00–7.60)
Drug-absent male	2.80 (11.60–8.80)
Drug-present female	3.40 (22.40–19.00)
Drug-absent female	7.00 (24.20–17.20)

number of errors made from Trial1 to Trial2 is similar under both the drug-absent ($M = 1.80$) and the drug-present conditions ($M = 1.00$). The results show that between Trial2 and Trial3, the reduction in the number of errors made is greater under the drug-absent condition than under the drug-present condition for both male and females. However, the decrease in errors between the drug-absent and drug-present conditions is greater for males than for females.

The results indicate that the last contrast, LEVEL 3 vs. LEVEL 4, is not significant, $F(3,16) = 2.83$, $p > .05$. In conjunction with Figure 9.6, the results indicate that the reduction in the number of errors made between Trial3 and Trial4 is similar for the four groups of drug-present male, drug-present female, drug-absent male, and drug-absent female.

9.6.5.2 Between-Groups Effects

The final analysis presented is the **Tests for Between-Subjects Effects**. This is equivalent to a two-way ANOVA. The results indicate that only subjects' **SEX** exerted a significant effect on the number of errors made, $F(1,16) = 24.45$, $p < .001$. From the **Estimated Marginal Means** table of **SEX**, it can be seen that, on average (i.e., collapsing across the **DRUG** and **TRIAL** conditions), females made more errors than males (females: $M = 24.15$, males: $M = 14.02$). Neither the **DRUG** variable, $F(1,16) = 1.40$, $p > .05$, nor the **DRUG*SEX** interaction, $F(1,16) = 0.03$, $p > .05$, had a significant effect on the number of errors made. In conjunction with Figure 9.3, the results indicate that, irrespective of the presence or absence of the drug, females made more errors than males averaged across the four trials.

10

Correlation

10.1 Aim

Correlation is primarily concerned with investigating whether a relationship exists and with determining its magnitude and direction. When two variables vary together, such as loneliness and depression, they are said to be correlated. Accordingly, correlational studies attempt to find the extent to which two or more variables are related. Typically, in a correlational study, no variables are manipulated as in an experiment—the researcher measures naturally occurring events, behaviors, or personality characteristics and then determines if the measured scores covary. The simplest correlational study involves getting a pair of observations or measures on two different variables from a number of individuals. The paired measures are then statistically analyzed to determine if any relationship exists between them. For example, behavioral scientists have explored the relationship between variables such as anxiety level and self-esteem, attendance at classes in school and course grades, university performance and career success, and body weight and self-esteem.

In order to show quantitatively the extent to which two variables are related, it is necessary to calculate a correlation coefficient. There are many types of correlation coefficients, and the decision of which one to employ with a specific set of data depends on the following factors:

- The level of measurement on which each variable is measured
- The nature of the underlying distribution (continuous or discrete)
- The characteristics of the distribution of the scores (linear or nonlinear)

This chapter presents two correlation coefficients: the *Pearson product moment correlation coefficient* (r), employed with interval- or ratio-scaled variables, and the *Spearman rank order correlation coefficient* (*r*rho), employed with ordered or ranked data. It is important to note that, irrespective of which

correlational technique the researcher uses, they have the following characteristics in common:

1. Two sets of measurements are obtained on the same individuals or on pairs of individuals who are matched on some basis.
2. The values of the correlation coefficients vary between +1.00 and –1.00. Both of these extremes represent perfect relationships between the variables, and 0.00 represents the absence of a relationship.
3. A *positive relationship* means that individuals obtaining high scores on one variable tend to obtain high scores on a second variable. The converse is also true, that is, individuals scoring low on one variable tend to score low on a second variable.
4. A *negative relationship* means that individuals scoring low on one variable tend to score high on a second variable. Conversely, individuals scoring high on one variable tend to score low on a second variable.

10.2 Requirements

- For each subject in the study, there must be *related pairs of scores*. That is, if a subject has a score on variable X, then the same subject must also receive a score on variable Y.
- The variables should be measured at least at the *ordinal level*.

10.3 Assumptions

- **Linearity**—The relationship between the two variables must be *linear*, that is, the relationship can be most accurately represented by a straight line.
- **Homoscedasticity**—The variability of scores along the Y variable should remain constant at all values of the X variable.

10.4 Example 1: Pearson Product Moment Correlation Coefficient

Assume that a researcher wishes to ascertain whether there is a relationship between grade point average (GPA) and the scores on a reading-comprehension (READ) test of 15 first-year students. The researcher recorded the pair of scores below, together with their rankings.

Student	Read	Read_Rank	GPA	GPA_Rank
s1	38	13	2.1	13
s2	54	3	2.9	6
s3	43	10	3.0	5
s4	45	8	2.3	12
s5	50	4	2.6	7.5
s6	61	1	3.7	1
s7	57	2	3.2	4
s8	25	15	1.3	15
s9	36	14	1.8	14
s10	39	11.5	2.5	9.5
s11	48	5.5	3.4	2
s12	46	7	2.6	7.5
s13	44	9	2.4	11
s14	39	11.5	2.5	9.5
s15	48	5.5	3.3	3

10.4.1 Data Entry Format

The data set has been saved under the name **CORR.SAV.**

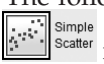
Variables	Column(s)	Code
READ	1	Reading score
READ_RANK	2	Ranking
GPA	3	Grade point average
GPA_RANK	4	Ranking

10.4.2 Testing Assumptions

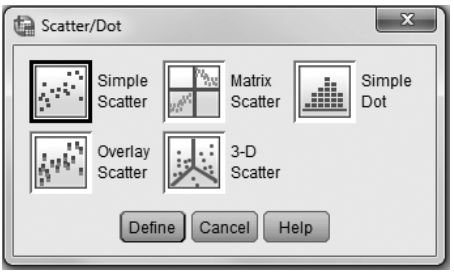
Linearity and **Homoscedasticity**

Scatterplots will be used to test the assumptions of linearity and homoscedasticity.

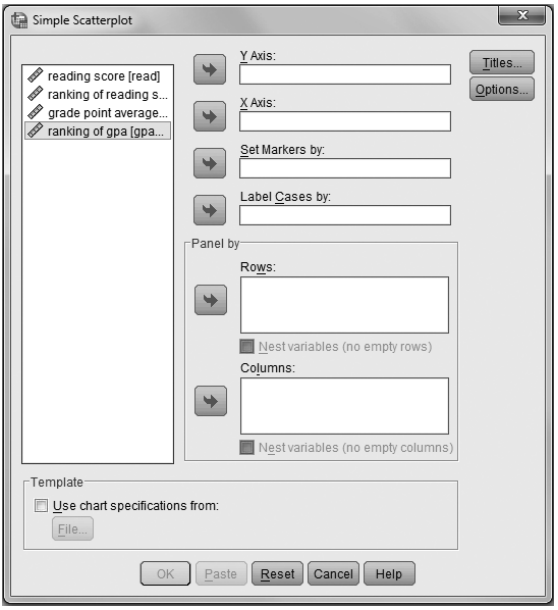
10.4.2.1 Windows Method

1. From the menu bar, click **Graphs**, then **Legacy Dialogs**, and then **Scatter/Dot…**. The following **Scatter/Dot** window will open. Click (highlight) the 

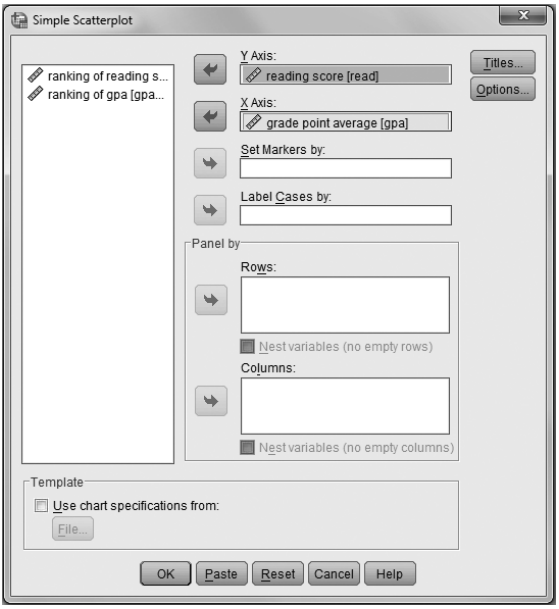
icon.



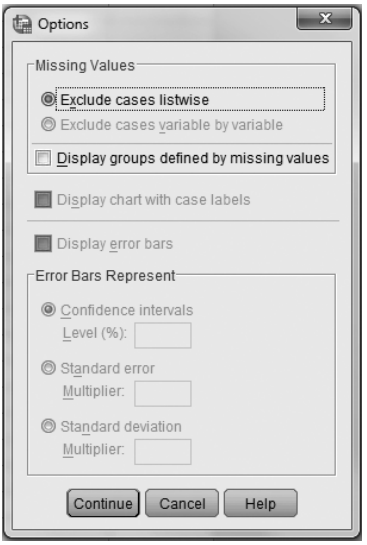
2. Click Define to open the **Simple Scatterplot** window below.



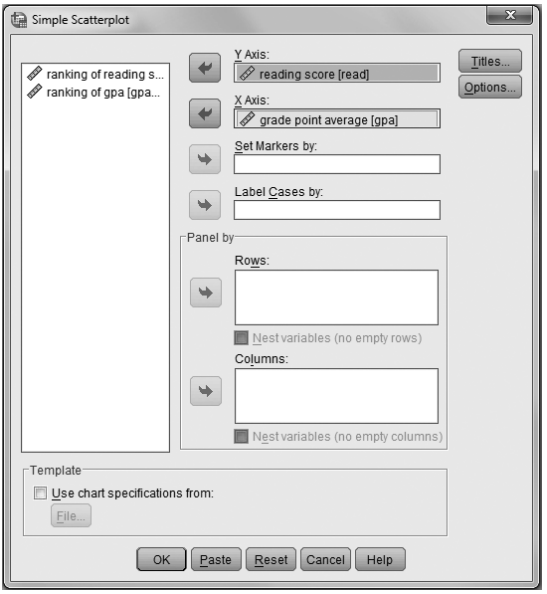
3. Transfer the **READ** variable to the **Y Axis:** field by clicking (highlight) the variable and then clicking . Transfer the **GPA** variable to the **X Axis:** field by clicking (highlight) the variable and then clicking .



4. Click Options... to open the **Options** window below. Under **Missing Values**, ensure that the **Exclude cases listwise** field is checked. By default, for scatterplot, SPSS employs the **listwise** method for handling missing data. (For a discussion of the differences in the pairwise and the listwise methods of handling missing data, please see Section 10.3.6.)



5. Click Continue to return to the **Simple Scatterplot** window.



6. When the **Simple scatterplot** window opens, click OK to complete the analysis. See Figure 10.1 for the results.

10.4.2.2 SPSS Syntax Method

```
GRAPH
/SCATTERPLOT(BIVAR) = GPA WITH READ
/MISSING = LISTWISE.
```

10.4.2.3 Scatterplot

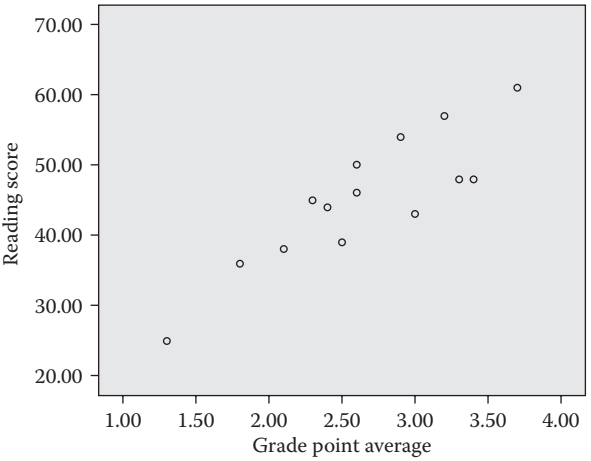


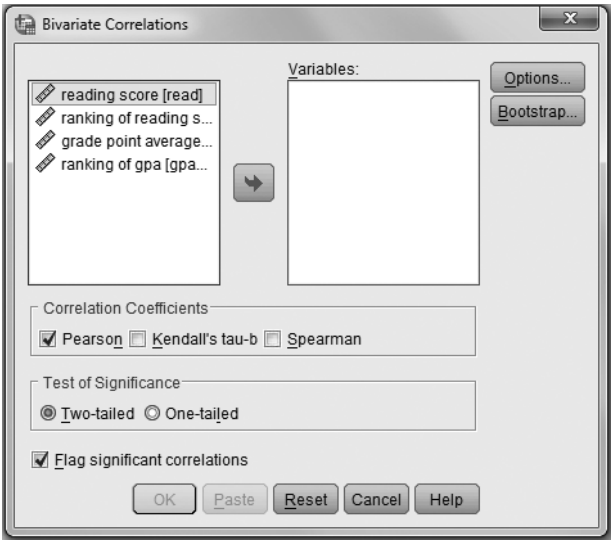
FIGURE 10.1
Scatterplot between READ and GPA.

10.4.2.4 Interpretation

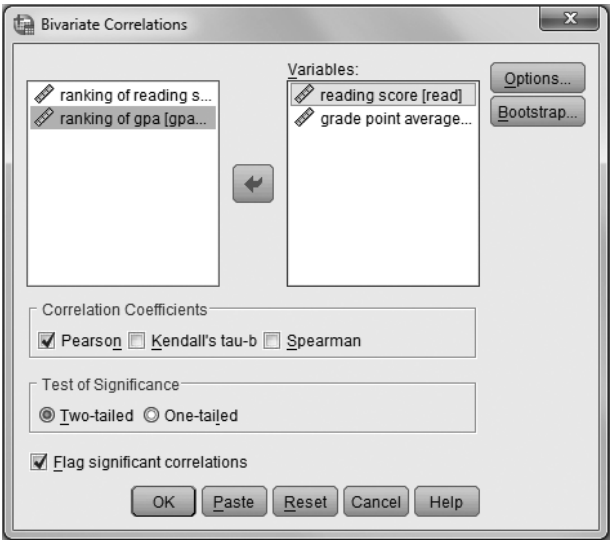
As can be seen from Figure 10.1, there is a linear relationship between the variables of reading score and grade point average, such that as reading score increases, so does grade point average. The figure also shows that the homoscedasticity assumption is met, because the variability of the READ score remains relatively constant from one GPA score to the next. Heteroscedasticity is usually shown by a cluster of points that is wider as the values for the Y variable (READ) get larger.

10.4.3 Windows Method: Pearson Product Moment Correlation

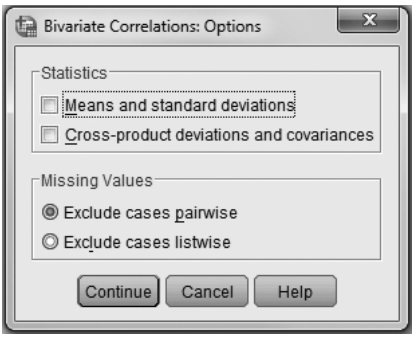
1. From the menu bar, click **Analyze**, then **Correlate,** and then **Bivariate…**. The following **Bivariate Correlations** window will open.



2. Transfer the **READ** and **GPA** variables to the **Variables:** field by clicking (highlight) them and then clicking ▶. By default, SPSS will employ the **Pearson correlation analysis**, and a **two-tailed test of significance** (both fields are checked).

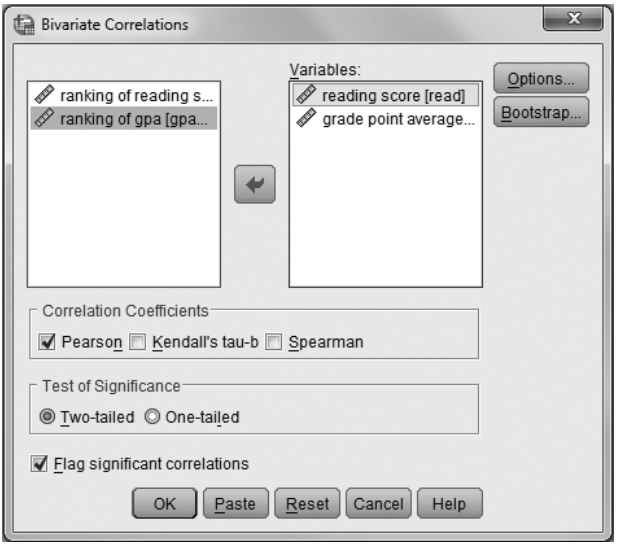


3. Click Options... to open the **Bivariate Correlation: Options** window.



By default, SPSS employs the **Exclude cases pairwise** method for handling missing data (this field is checked). This method treats the calculation of each pair of variables as a separate problem, using all cases with complete data for the pair. With this method, only cases with missing data on a specific pair of variables will be excluded from the computation. As such, the correlation of different pairs of variables may be on the basis of different numbers of subjects. As an option, the **Exclude cases listwise** method can be used to include only those cases with complete data. With this method, any case with missing data on any pair of variables will be excluded from the computation. As such, the sample size for the correlation between any pair of variables can be reduced further with the listwise method.

4. Click Continue to return to the **Bivariate Correlations** window.



5. When the **Bivariate Correlations** window opens, click OK to complete the analysis. See Table 10.1 for the results.

10.4.4 SPSS Syntax Method: Pearson Product Moment Correlation

```
CORRELATIONS READ WITH GPA
/MISSING = PAIRWISE.
```

10.4.5 SPSS Output

TABLE 10.1

Pearson Product Moment Correlation

Correlations

		Reading Score	Grade Point Average
reading score	Pearson Correlation	1	.867**
	Sig. (2-tailed)		.000
	N	15	15
grade point average	Pearson Correlation	.867**	1
	Sig. (2-tailed)	.000	
	N	15	15

** Correlation is significant at the 0.01 level (2-tailed).

10.4.6 Results and Interpretation

The correlation between reading scores and grade point average is positive and statistically significant ($r = 0.867$, $p < .001$). This means that as the students' reading scores increase, so do their grade point averages. Please note that this interpretation in no way implies *causality*—that increases in reading scores caused increases in GPA scores. The significant relationship merely indicates that the two variables *covary*.

10.5 Testing Statistical Significance between Two Correlation Coefficients Obtained *from Two Samples*

Suppose that a researcher is interested in finding out whether the correlations obtained between grade point average (GPA) and the scores on a reading-comprehension test (READ) for 15 first-year students from each of two universities are significantly different (University A: $\boldsymbol{r_A = 0.87}$, $N_A = 15$; University B: $\boldsymbol{r_B = 0.46}$, $N_B = 15$). The Z test will be used to accomplish this. The test involves calculating the Z-obtained value (Z_{obt}) and comparing this against the values ±1.96. If the $Z_{obt} \geq \pm 1.96$, then the two coefficients are statistically significant ($p < .05$).

1. Covert the r values into Z scores using the following table below.

r	z_r	r	z_r	r	z_r	r	z_r	r	z_r
.000	.000	.200	.203	.400	.424	.600	.693	.800	1.099
.005	.005	.205	.208	.405	.430	.605	.701	.805	1.113
.010	.010	.210	.213	.410	.436	.610	.709	.810	1.127
.015	.015	.215	.218	.415	.442	.615	.717	.815	1.142
.020	.020	.220	.224	.420	.448	.620	.725	.820	1.157
.025	.025	.225	.229	.425	.454	.625	.733	.825	1.172
.030	.030	.230	.234	.430	.460	.630	.741	.830	1.188
.035	.035	.235	.239	.435	.466	.636	.750	.835	1.204
.040	.040	.240	.245	.440	.472	.640	.758	.840	1.221
.045	.045	.245	.250	.445	.478	.645	.767	.845	1.238
.050	.050	.250	.255	.450	.485	.650	.775	.850	1.256
.055	.055	.255	.261	.455	.491	.655	.784	.855	1.274
.060	.060	.260	.266	.460	.497	.660	.793	.860	1.293
.065	.065	.265	.271	.465	.504	.665	.802	.865	1.313
.070	.070	.270	.277	.470	.510	.670	.811	.870	1.333
.075	.075	.275	.282	.475	.517	.675	.820	.875	1.354
.080	.080	.280	.288	.460	.523	.680	.829	.880	1.376
.085	.085	.285	.293	.485	.530	.685	.838	.885	1.398
.090	.090	.290	.299	.490	.536	.690	.848	.890	1.422
.095	.095	.295	.304	.495	.543	.695	.858	.895	1.447
.100	.100	.300	.310	.500	.549	.700	.867	.900	1.472
.105	.105	.305	.315	.505	.556	.705	.877	.905	1.499
.110	.110	.310	.321	.510	.563	.710	.887	.910	1.528
.115	.116	.315	.326	.515	.570	.715	.897	.915	1.557
.120	.121	.320	.332	.520	.576	.720	.908	.920	1.589
.125	.126	.325	.337	.525	.583	.725	.918	.925	1.623
.130	.131	.330	.343	.530	.590	.730	.929	.930	1.658
.135	.136	.335	.348	.535	.597	.735	.940	.935	1.697
.140	.141	.340	.354	.540	.604	.740	.950	.940	1.738
.145	.146	.345	.360	.545	.611	.745	.962	.945	1.783
.150	.151	.350	.365	.550	.618	.750	.973	.950	1.832
.155	.156	.355	.371	.555	.626	.755	.984	.955	1.886
.160	.161	.360	.377	.560	.633	.760	.996	.960	1.946
.165	.167	.365	.383	.565	.640	.765	1.008	.965	2.014
.170	.172	.370	.388	.570	.648	.770	1.020	.970	2.092
.175	.177	.375	.394	.575	.655	.775	1.033	.975	2.185
.180	.182	.380	.400	.580	.662	.780	1.045	.980	2.298
.185	.187	.385	.406	.585	.670	.785	1.058	.985	2.443
.190	.192	.390	.412	.590	.678	.790	1.071	.990	2.647
.195	.198	.395	.418	.595	.685	.795	1.085	.995	2.994

Source: McCall (1990); originally from Edwards, A. L. (1967). *Statistical methods* (2nd edition). Holt, Rinehart and Winston.

Thus, for r_A = **0.87**, the corresponding ***Z* score** is **1.333**, and for r_B = **0.46**, the corresponding ***Z* score** is **0.497**.

2. Plug these values into the following equation.

$$Z_{obt} = \frac{Z_A - Z_B}{\sqrt{(1/[N_A - 3]) + (1/[N_B - 3])}}$$

$$Z_{obt} = \frac{1.33 - 0.497}{\sqrt{(1/[15 - 3]) + (1/[15 - 3])}}$$

$$Z_{obt} = \frac{0.833}{\sqrt{(1/12) + (1/12)}}$$

$$Z_{obt} = \frac{0.833}{\sqrt{0.0833 + 0.0833}}$$

$$Z_{obt} = \frac{0.833}{\sqrt{0.167}}$$

$$Z_{obt} = \frac{0.833}{0.409} = 2.04$$

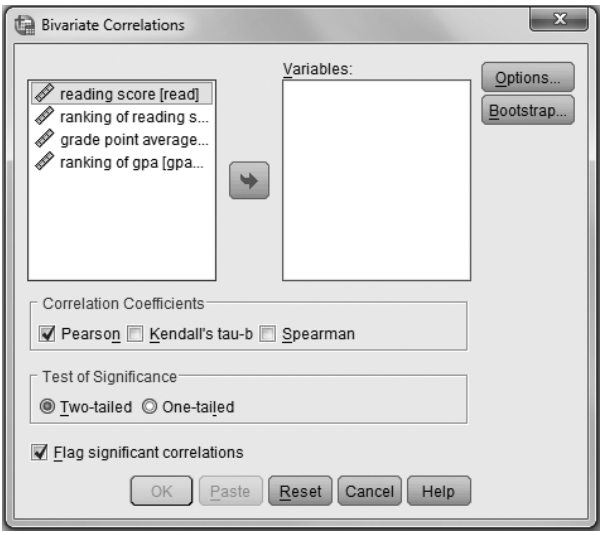
3. The calculated Z_{obt} value (2.04) is greater than +1.96. Therefore, it can be concluded that there is a statistically significant difference in the strength of the correlation between GPA and READ scores for the students at University A and University B.

10.6 Example 2: Spearman Rank Order Correlation Coefficient

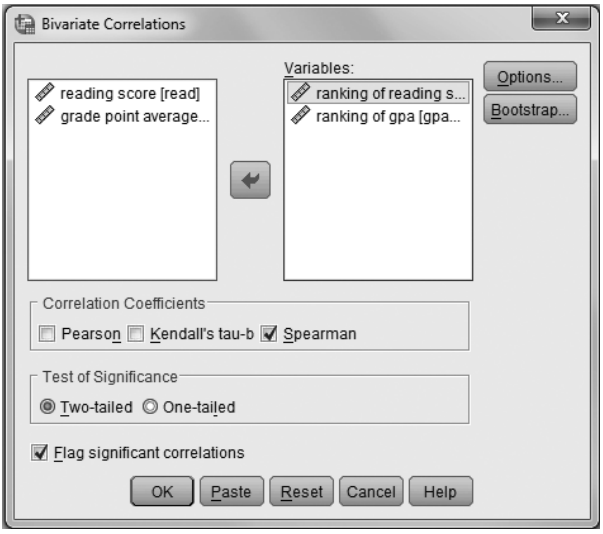
For this example, the same data set (**CORR.SAV**) will be used. However, the rank order of the two variables (READ_RANK, GPA_RANK) will be used instead of their actual values as recorded. Thus, the computation for this coefficient is not sensitive to asymmetrical distributions or to the presence of outliers.

10.6.1 Windows Method

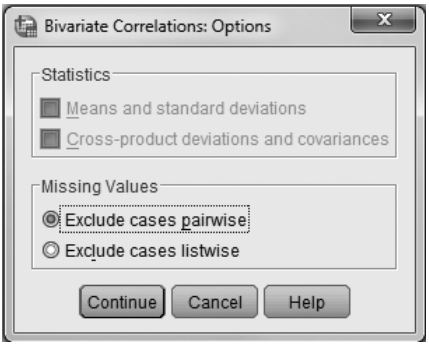
1. From the menu bar, click **Analyze**, then **Correlate,** and then **Bivariate…**. The following **Bivariate Correlations** window will open.



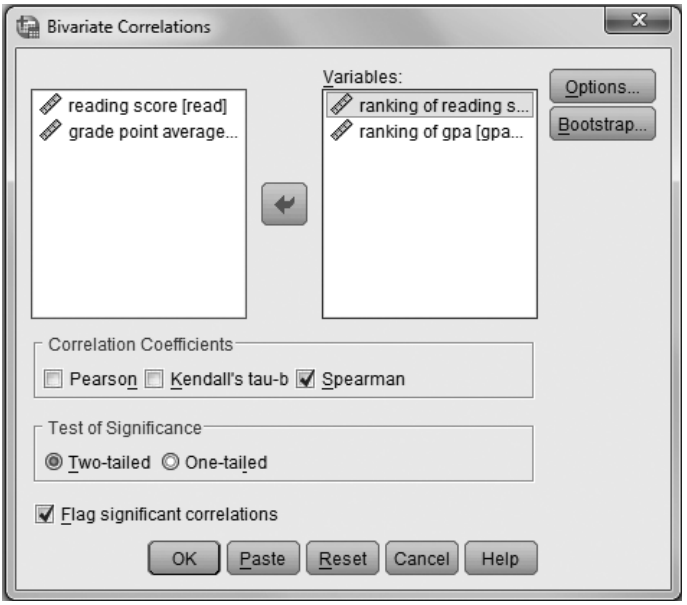
2. Transfer the **READ_RANK** and **GPA_RANK** variables to the **Variables:** field by clicking (highlight) them and then clicking . By default, SPSS will employ the **Pearson correlation analysis** (this field is checked). Uncheck the **Pearson** field and check the **Spearman** field.



3. Click Options... to open the **Bivariate Correlation: Options** window. Ensure that the **Exclude cases pairwise** field is checked.



4. Click Continue to return to the **Bivariate Correlations** window.



5. When the **Bivariate Correlations** window opens, click OK to complete the analysis. See Table 10.2 for the results.

10.6.2 SPSS Syntax Method

```
NONPAR CORR READ_RANK WITH GPA_RANK.
```

10.6.3 SPSS Output

TABLE 10.2

Spearman Rank Order Correlation

Correlations

			Ranking of Reading Scores	Ranking of gpa
Spearman's rho	ranking of reading scores	Correlation Coefficient	1.000	.826**
		Sig. (2-tailed)	.	.000
		N	15	15
	ranking of gpa	Correlation Coefficient	.826**	1.000
		Sig. (2-tailed)	.000	.
		N	15	15

** Correlation is significant at the 0.01 level (2-tailed).

10.6.4 Results and Interpretation

The obtained Spearman rank-order coefficient (rho = 0.826, $p < .001$) is highly similar in magnitude and direction to that in the Pearson correlation table (Table 10.1). Thus, similar to the Pearson coefficient, the Spearman coefficient indicates that as the students' ranked reading scores increase, so do their ranked grade point average scores.

11

Linear Regression

11.1 Aim

Regression and correlation are closely related. Both techniques involve the relationship between two variables, and they both use the same set of paired scores taken from the same subjects. However, while correlation is concerned with the magnitude and direction of the relationship, regression focuses on applying the relationship for prediction. In terms of prediction, if two variables were correlated perfectly, then knowing the value of one score permits a perfect prediction of the score on the second variable. Generally, whenever two variables are significantly correlated, the researcher may use the score on one variable to predict the score on the second.

There are many reasons why researchers want to predict one variable from another. For example, knowing a person's IQ, what can we say about this person's prospects of successfully completing a university course? Knowing a person's prior voting record, can we make any informed guesses concerning his vote in the coming election? Knowing his mathematics aptitude score, can we estimate the quality of his performance in a course in statistics? These questions involve predictions from one variable to another, and psychologists, educators, biologists, sociologists, and economists are constantly being called upon to perform this function.

11.2 Requirements

- For each subject in the study, there must be *related pairs of scores*. That is, if a subject has a score on variable *X*, then the same subject must also have a score on variable *Y*.
- The variables should be measured at least at the *ordinal level*.

11.3 Assumptions

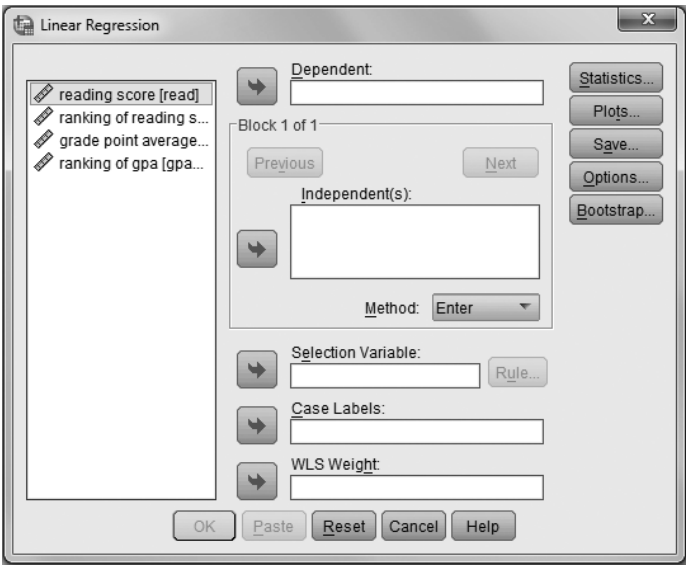
- **Linearity**—The relationship between the two variables must be *linear*, that is, the relationship can be more accurately represented by a straight line.
- **Homoscedasticity**—The variability of scores on the *Y* variable should remain constant at all values of the *X* variable.

11.4 Example: Linear Regression

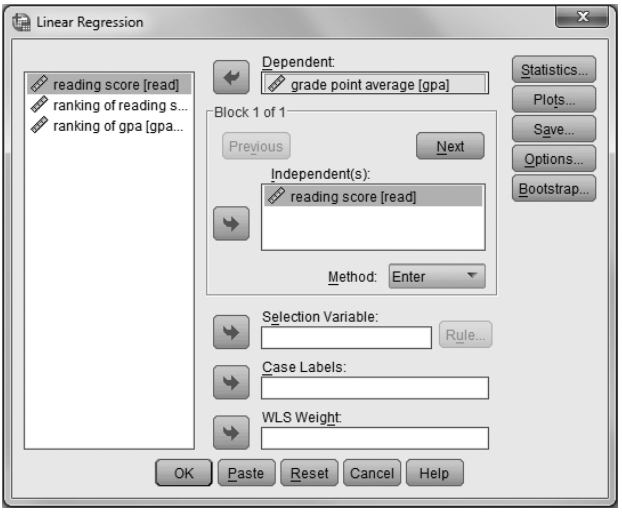
This example employs the data set **CORR.SAV** from the previous chapter (see Section 10.4 in Chapter 10). In this example, we wish to (1) find the prediction equation that allows us to best predict students' grade point average scores (**GPA**) from their reading scores (**READ**), (2) determine the strength of this relationship, and (3) test the null hypothesis that **READ** and **GPA** scores are unrelated.

11.4.1 Windows Method

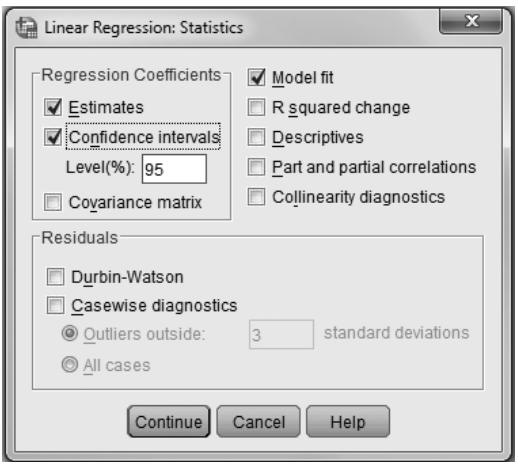
1. From the menu bar, click **Analyze**, then **Regression**, and then **Linear**. The following **Linear Regression** window will open.



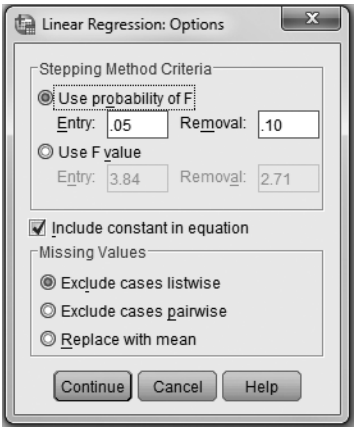
2. Click (highlight) the **GPA** variable and then click ▶ to transfer this variable to the **Dependent:** field. Next, click (highlight) the **READ** variable and then click ▶ to transfer this variable to the **Independent(s):** field. In the **Method:** field, select **Enter** from the drop-down list as the method of entry for the independent (predictor) variable into the prediction equation.



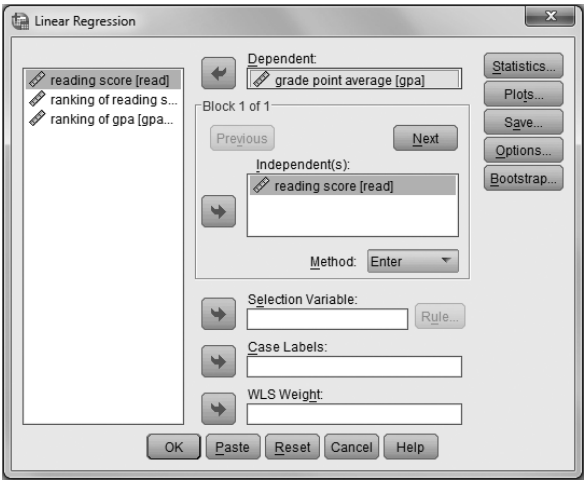
3. Click Statistics... to open the **Linear Regression: Statistics** window. Check the fields to obtain the statistics required. For this case, check the fields for **Estimates, Confidence intervals,** and **Model fit**. Click Continue when finished.



4. When the **Linear Regression** window opens, click Options... to open the **Linear Regression: Options** window below. Ensure that both the **Use probability of F** and the **Include constant in equation** fields are checked. Click Continue to return to the **Linear Regression** window.



5. When the **Linear Regression** window opens, click OK to complete the analysis. See Table 11.1 for the results.



11.4.2 SPSS Syntax Method

```
REGRESSION VARIABLES = (COLLECT)
/MISSING LISTWISE
/STATISTICS = DEFAULTS CI
/DEPENDENT = GPA
/METHOD = ENTER READ.
```

11.4.3 SPSS Output

TABLE 11.1

Linear Regression Analysis Output Regression

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	Reading score[a]	.	Enter

[a] All requested variables entered.
[b] Dependent Variable: grade point average.

Model Summary

Model	*R*	*R* Square	Adjusted *R* Square	Std. Error of the Estimate
1	.867[a]	.752	.733	.32848

[a] Predictors: (Constant), reading score.

ANOVA[b]

Model		Sum of Squares	d*f*	Mean Square	*F*	Sig.
1	Regression	4.253	1	4.253	.39.418	.000[a]
	Residual	1.403	13	.108		
	Total	5.656	14			

[a] Predictors: (Constant), reading score.
[b] Dependent Variable: grade point average.

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			95% Confidence Interval for B	
Model		B	Std. Error	Beta	*t*	Sig.	Lower Bound	Upper Bound
1	(Constant)	−.111	.446		−.248	.808	−1.075	.853
	reading score	.061	.010	.867	6.278	.000	.040	.082

[a] Dependent Variable: grade point average.

11.4.4 Results and Interpretation

11.4.4.1 Prediction Equation

The prediction equation is $Y' = A + BX$ where Y' is the predicted dependent variable, A is the constant, B is the unstandardized regression coefficient, and X is the value of the predictor variable.

The relevant information for constructing a least-squares regression (prediction) equation is presented in the **Coefficients** table (see Table 11.1).

In order to predict the students' grade point average scores (**GPA**) from their reading scores (**READ**), use the values presented in the **Unstandardized Coefficients** column. Using the **Constant** and **B** (unstandardized coefficient) values, the prediction equation would be:

$$\text{Predicted GPA} = -0.111 + (.061 \times \text{READ})$$

Thus, for a student who has a reading score of 56, his/her predicted GPA score will be:

$$\text{Predicted GPA} = -0.111 + (.061 \times 56) = \mathbf{3.31}$$

However, in the **Model Summary** table, the **Standard Error of the Estimate** is 0.32848. This means that at the 95% confidence interval, the predicted GPA score of 3.31 lies between the scores of **2.66** (3.31 − (1.96 × 0.32848)) and **3.95** (3.31 + (1.96 × 0.32848)).

11.4.4.2 Evaluating the Strength of the Prediction Equation

A measure of the strength of the computed equation is ***R*-square**, sometimes called the **coefficient of determination**. *R*-square is simply the square of the multiple correlation coefficient listed under ***R*** in the **Model Summary** table, and represents the proportion of variance accounted for in the dependent variable (GPA) by the predictor variable (READ). In a simple regression such as this, where there is only one predictor variable, the multiple *R* is equivalent to the simple *R* (Pearson product-moment correlation). For this example, the multiple correlation coefficient is 0.867, and the *R*-square is 0.752. Therefore, for this sample, the predictor variable of READ has explained 75.2% of the variance in the dependent variable of GPA.

The **ANOVA** table presents results from the test of the null hypothesis that *R*-square is zero. An *R*-square of zero indicates no linear relationship between the predictor and the dependent variable. The **ANOVA** table shows that the computed *F* statistic is 39.42, with an observed significance level of less than 0.001. Thus, the hypothesis that there is no linear relationship between the predictor and the dependent variable is rejected.

11.4.4.3 Identifying an Independent Relationship

The **Coefficients** table presents the **standardized Beta** coefficient between the predictor variable READ and the dependent variable GPA. The Beta coefficient is shown to be positive and statistically significant at the 0.001 level. Thus, the higher the students' reading scores, the higher their GPA scores, (Beta = 0.87), $t = 6.28$, $p < .001$. Note that the standardized Beta coefficient of 0.87 is identical to the multiple *R* coefficient. This is because there is only one predictor variable.

12

Factor Analysis

12.1 Aim

The major purpose of factor analysis is the orderly simplification of a large number of inter-correlated measures to a few representative constructs or factors. Suppose that a researcher wants to identify the major dimensions underlying a number of personality tests. He begins by administering the personality tests to a large sample of people ($N = 1000$), with each test supposedly measuring a specific aspect of a person's personality (e.g., ethnocentrism, authoritarianism, locus of control). Assume that there are 30 such tests, each consisting of 10 test items. What the researcher will end up with is a mass of numbers (i.e., $1000 \times 30 \times 10 = 300{,}000$ scores) that will say very little about the dimensions underlying these personality tests. On average, some of the scores will be high, some will be low, and some intermediate, but interpretation of these scores will be extremely difficult if not impossible. This is where factor analysis comes in. It allows the researcher to "reduce" this mass of numbers to a few representative factors that can then be used for subsequent analysis.

Factor analysis is based on the assumption that all variables correlate to some degree. Consequently, those variables that share similar underlying dimensions should be highly correlated, and those variables that measure dissimilar dimensions should yield low correlations. Using the above example, if the researcher inter-correlates the scores obtained from the 30 personality tests, then those tests that measure the same underlying personality dimension should yield high correlation coefficients, while those tests that measure different personality dimensions should yield low correlation coefficients. These high/low correlation coefficients will become apparent in the correlation matrix because they form clusters indicating that variables "hang" together. For example, measures of ethnocentrism, authoritarianism, and aggression may be highly inter-correlated indicating that they form an identifiable personality dimension. The primary role of factor analysis is to identify these clusters of high inter-correlations as independent factors.

There are three basic steps to factor analysis:

1. Computation of the correlation matrix for all variables
2. Extraction of initial factors
3. Rotation of the extracted factors to a terminal solution

12.1.1 Computation of the Correlation Matrix

As factor analysis is based on correlations between measured variables, a correlation matrix containing the inter-correlation coefficients for the variables must be computed. The variables should be measured at least at the ordinal level, although two-category nominal variables (coded 1–2) can be used. If all variables are nominal variables, then specialized forms of factor analysis, such as Boolean factor analysis (BMPD Statistical Software, Inc., 1992), are more appropriate.

12.1.2 Extraction of Initial Factors

At this phase, the number of common factors needed to adequately describe the data is determined. To do this, the researcher must decide on (1) the method of extraction, and (2) the number of factors selected to represent the underlying structure of the data.

12.1.2.1 Method of Extraction

There are two basic methods for obtaining factor solutions. They are **principal components** analysis and common **factor analysis**. (*Note:* SPSS provides six methods of extraction under the common factor analysis model; these are principal-axis factoring, unweighted least-squares, generalized least-squares, maximum-likelihood, alpha factoring, and image factoring.) The choice between these two basic methods of factor extraction lies with the objective of the researcher. If the purpose is no more than to "reduce data" in order to obtain the minimum number of factors needed to represent the original set of data, then principal components analysis is appropriate. The researcher works from the assumption that the factors extracted need not have any theoretical validity. Conversely, when the main objective is to identify theoretically meaningful underlying dimensions, the common factor analysis method is the appropriate model. Given the more restrictive assumptions underlying common factor analysis, the principal component method has attracted more widespread use.

12.1.2.2 Determining the Number of Factors to Be Extracted

There are two conventional criteria for determining the number of initial unrotated factors to be extracted. These are the **eigenvalues** criterion and the **scree test** criterion. While these two criteria are easy to use as they are

based on a "rules of thumb" technique that is intuitively easy to understand, they have nonetheless attracted a number of criticisms. First, the eigenvalues ≥ 1.00 rule typically overestimates, and sometimes underestimates, the number of components extracted, and therefore does not always result in factors that are reliable (Zwick and Velicer, 1986). Second, the scree test rule is based on a visual plot of the eigenvalues against the number of factors in their order of extraction. It involves eyeballing the plot for sharp demarcations between the eigenvalues for major and trivial factors. In practice though, such sharp demarcations do not always exist or there may be more than one demarcation point (O'Connor, 2000). Not surprisingly, the reliability of scree plot interpretation is low, even among experts (Crawford and Koopman, 1979; Streiner, 1998).

In addition to these two rules-of-thumb criteria, there are two statistically based procedures that are considered by statisticians to be superior and typically yield optimum solutions to the number of factors problem (Wood, Tataryn, and Gorsuch, 1996; Zwick and Velicer, 1982, 1986). These are **parallel analysis** and **Velicer's minimum average partial (MAP) test**.

The following presents brief descriptions of these four criteria.

12.1.2.2.1 Eigenvalues

Only factors with eigenvalues of 1 or greater are considered to be significant; all factors with eigenvalues less than 1 are ignored. An eigenvalue is a *ratio* between the common (shared) variance and the specific (unique) variance explained by a specific factor extracted. The rationale for using the eigenvalue criterion is that the amount of common variance explained by an extracted factor should be at least equal to the variance explained by a single variable (unique variance), if that factor is to be retained for interpretation. An eigenvalue greater than 1 indicates that more common variance than unique variance is explained by that factor.

12.1.2.2.2 Scree Test

This test is used to identify the optimum number of factors that can be extracted before the amount of unique variance begins to dominate the common variance structure (Hair, Anderson, Tatham, and Black, 1995). The scree test is derived by plotting the eigenvalues (on the Y axis) against the number of factors in their order of extraction (on the X axis). The initial factors extracted are large factors (with high eigenvalues), followed by smaller factors. Diagrammatically, the plot will indicate a steep slope between the large factors and the gradual trailing off of the rest of the factors. The point at which the curve first begins to straighten out (point of demarcation) is taken to indicate the maximum number of factors to extract. That is, those factors above this point of inflection are deemed meaningful and those below are not. As a general rule, the scree test results in at least one and sometimes two or three more factors being considered significant than does the eigenvalue criterion (Cattel, 1966).

12.1.2.2.3 Velicer's Minimum Average Partial (MAP) Test

This test involves a complete principal component analysis followed by the examination of a series of matrices of partial correlations (O'Connor, 2000).

The test then calculates the average *R*-square (systematic variation) of the correlation matrix (computed from the variables in the data set) ***after*** the factors (identified from principal components analysis) have been partialled out of the matrix. The number of factors corresponding to the ***smallest average R-square*** of the correlation matrix then represents the number of factors to be retained. The rationale for this is that the lower the *R*-square in the correlation matrix (the residual variance), the stronger the factors extracted.

12.1.2.2.4 Parallel Analysis

This procedure involves extracting eigenvalues from random data sets that parallel the actual data set with regard to the number of cases and variables (O'Connor, 2000). For example, if the original data set contains 704 cases for each of 18 variables, the analysis will generate random data matrices of the same size as the actual data set (i.e., 704 × 18), and then eigenvalues will be computed for the correlation matrix from the original data and from each of the random data sets. The eigenvalues derived from the actual data are then compared to the eigenvalues derived from the random data sets. Factors are retained as long as the *i*th eigenvalue from the actual data is ***greater*** than the *i*th eigenvalue from the random data.

12.1.3 Rotation of Extracted Factors

Factors produced in the initial extraction phase are often difficult to interpret. This is because the procedure in this phase ignores the possibility that variables identified to load on or represent factors may already have high loadings (correlations) with previous factors extracted. This may result in significant cross-loadings in which many factors are correlated with many variables. This makes interpretation of each factor difficult, because different factors are represented by the same variables. The rotation phase serves to "sharpen" the factors by identifying those variables that load on one factor and not on another. The ultimate result of the rotation phase is to achieve a simpler, theoretically more meaningful factor pattern.

12.1.4 Rotation Methods

There are two main classes of factor rotation method: **orthogonal** and **oblique**. Orthogonal rotation assumes that the extracted factors are independent, and the rotation process maintains the reference axes of the factors at 90°. Oblique rotation allows for correlated factors instead of maintaining independence between the rotated factors. The oblique rotation process does not require that the reference axes be maintained at 90°. Of the two rotation methods, oblique rotation is more flexible because the factor axes need not be

orthogonal. Moreover, at the theoretical level, it is more realistic to assume that influences in nature are correlated. By allowing for correlated factors, oblique rotation often represents the clustering of variables more accurately.

There are three major methods of orthogonal rotation: **varimax**, **quartimax**, and **equimax**. Of the three approaches, varimax has achieved the most widespread use as it seems to give the clearest separation of factors. It does this by producing the maximum possible simplification of the columns (factors) within the factor matrix. In contrast, both the quartimax and equimax approaches have not proven very successful in producing simpler structures, and have not gained widespread acceptance. While the orthogonal approach to rotation has several choices offered by SPSS, the oblique approach is limited to one method: **oblimin**.

12.1.5 Orthogonal (Varimax) versus Nonorthogonal (Oblique) Rotation

In choosing between orthogonal and oblique rotation, there is no compelling analytical reason to favor one method over the other. Indeed, there are no hard and fast rules to guide the researcher in selecting a particular orthogonal or oblique rotational technique. However, convention suggests that the following guidelines may help in the selection process. If the goal of the research is no more than to "reduce the data" to more manageable proportions, regardless of how meaningful the resulting factors may be, and if there is reason to assume that the factors are uncorrelated, then orthogonal rotation should be used. Conversely, if the goal of the research is to discover theoretically meaningful factors, and if there are theoretical reasons to assume that the factors will be correlated, then oblique rotation is appropriate.

Sometimes the researcher may not know whether or not the extracted factors might be correlated. In such a case, the researcher should try an oblique solution first. This suggestion is based on the assumption that, realistically, very few variables in a particular research project will be uncorrelated. If the correlations between the factors turn out to be very low (e.g., <0.20), the researcher could redo the analysis with an orthogonal rotation method.

12.1.6 Number of Factor Analysis Runs

It should be noted that when factor analysis is used for research (either for the purpose of data reduction or to identify theoretically meaningful dimensions), a minimum of two runs will normally be required. In the first run, the researcher allows factor analysis to extract factors for rotation. All factors with eigenvalues of 1 or greater will be subjected to varimax rotation by default within SPSS. Nevertheless, even after rotation, not all extracted rotated factors will be meaningful. For example, some small factors may be represented by very few items, and there may still be significant cross-loading of items across several factors. At this point, the researcher must determine which factors are substantively meaningful (either theoretically or intuitively), and retain only these for further rotation. It is not uncommon

for a data set to be subjected to a series of factor analysis and rotation before the obtained factors can be considered "clean" and interpretable.

12.1.7 Interpreting Factors

In interpreting factors, the size of the factor loadings (correlation coefficients between the variables and the factors they represent) will help in the interpretation. As a general principle, variables with large loadings indicate that they are representative of the factor, while small loadings suggest that they are not. In deciding what is large or small, a rule of thumb suggests factor loadings greater than ±0.33 are considered to meet the minimal level of practical significance. The reason for using the ±0.33 criterion is that, if the value is squared, the squared value represents the amount of the variable's total variance accounted for by the factor. Therefore, a factor loading of 0.33 denotes that approximately 10% of the variable's total variance is accounted for by the factor. The grouping of variables with high factor loadings should suggest what the underlying dimension is for that factor.

12.2 Checklist of Requirements

- Variables for factor analysis should be measured at least at the ordinal level.
- If the researcher has some prior knowledge about the factor structure, then several variables (five or more) should be included to represent each proposed factor.
- The sample size should be 100 or larger. A basic rule of thumb is to have at least five times as many cases as variables entered into the factor analysis. The more acceptable range would be a ten-to-one ratio.

12.3 Assumptions

The assumptions underlying factor analysis can be classified as **statistical** and **conceptual**.

12.3.1 Key Statistical Assumptions

- **Normality and linearity**—Departures from normality and linearity can diminish the observed correlation between measured variables, and thus degrade the factor solution.
- **Sufficient significant correlations in data matrix**—The researcher must ensure that the data matrix has sufficient correlations to justify

the application of factor analysis. If visual inspection reveals no substantial number of correlations of 0.33 or greater, then factor analysis is probably inappropriate.

12.3.2 Key Conceptual Assumptions

- **Selection of variables**—Variables should be selected to reflect the underlying dimensions that are hypothesized to exist in the set of selected variables. This is because factor analysis has no means to determine the appropriateness of the selected variables other than the correlations among the variables.
- **Homogeneity**—The sample must be homogeneous with respect to the underlying factor structure. If the sample consists of two or more distinct groups (e.g., males and females), separate factor analysis should be performed.

12.4 Factor Analysis: Example 1

A study was designed to investigate the way people perceive and respond to the issue of euthanasia (Ho, 1998). Twelve statements were written to reflect the sub-categorical distinctions of (1) active-voluntary euthanasia, (2) active-nonvoluntary euthanasia, (3) passive-voluntary euthanasia, and (4) passive-nonvoluntary euthanasia. Each statement was rated on a five-point scale, with high scores indicating strong support for that type of euthanasia. A total of 357 subjects provided responses to these 12 statements. Factor analysis (with principal components extraction) was employed to investigate whether these 12 statements represent identifiable factors, that is, types of euthanasia. The 12 statements (together with their SPSS variable name) written to reflect the four sub-categorical distinctions of euthanasia are listed below.

Active-Voluntary Euthanasia

- E1*—Doctors have the right to administer medication that will painlessly end the life of a terminally ill person, if he/she requests it.
- E5—Terminally ill patients have the right to decide about their own lives and deaths.
- E9—Denying the request of terminally ill patients to "die with dignity" is unfair and cruel.

Active-Nonvoluntary Euthanasia

- E3—It is all right for doctors who attend childbirth to administer medication that will painlessly end the lives of grossly deformed babies.

- E7—Doctors have the right to administer medication that will painlessly end the life of a patient who has been diagnosed to be in a permanent "vegetative" state, if members of the patient's family request it.
- E11—When a patient is suffering a debilitating terminal disease such as Alzheimer's, it is all right for the physician to administer medication that will painlessly end the patient's life, if members of the patient's family request it.

Passive-Voluntary Euthanasia

- E2—Terminally ill patients have the right to decide that life-sustaining drugs or mechanism be withheld or withdrawn, to hasten their death.
- E6—The golden rule requires that we respect the request of terminally ill patients who have asked that life-sustaining drugs or mechanisms be withheld or withdrawn, to hasten their death.
- E10—Doctors have the right to withhold or withdraw life-sustaining drugs or mechanisms for a terminally ill person, if he or she requests it.

Passive-Nonvoluntary Euthanasia

- E4—Doctors have the right to remove life support from a patient who is in a state of constant unconsciousness with no prognosis for recovery, if members of the patient's family request it.
- E8—The withdrawal of medical treatment to allow death to occur should be permissible in cases where a patient, although biologically alive, has been diagnosed as clinically brain-dead.
- E12—It is all right for doctors who attend childbirth not to resuscitate grossly deformed babies.

(*Name and order of items in questionnaire)

12.4.1 Data Entry Format

The data set is named **EUTHAN_1.SAV**.

Variables	Column(s)	Code
e1 to e12	1–12	1 = strongly disagree, 5 = strongly agree

12.4.2 Testing Assumptions

12.4.2.1 Normality

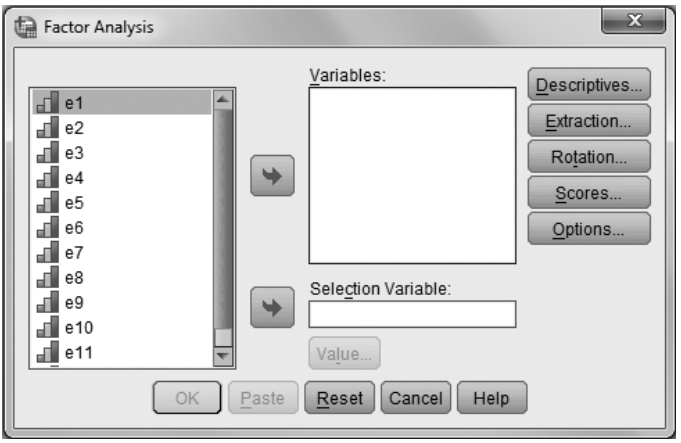
Normality can be tested using the normal Q–Q plot, the detrended Q–Q plot, and the z test for skewness. Follow steps 1 through 5 outlined in Chapter 4, Section 4.4.2.2.1 to achieve this.

12.4.2.2 Sufficient Significant Correlations in Data Matrix

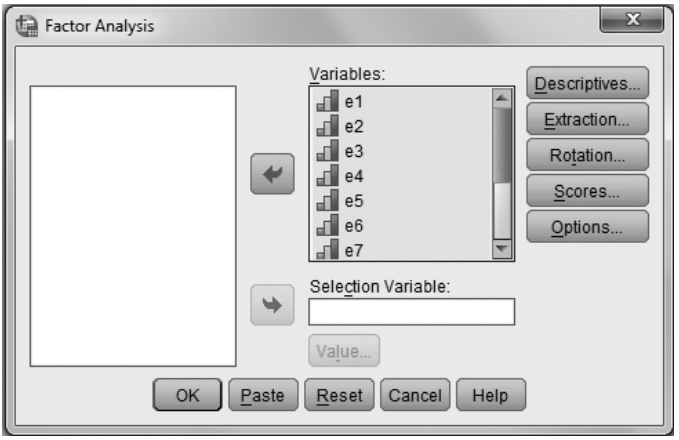
Bartlett's Test of Sphericity can be used to test for the adequacy of the correlation matrix, that is, the correlation matrix has significant correlations among at least some of the variables. This test will be conducted as part of the factor analysis.

12.4.3 Windows Method: Factor Analysis (First Run)

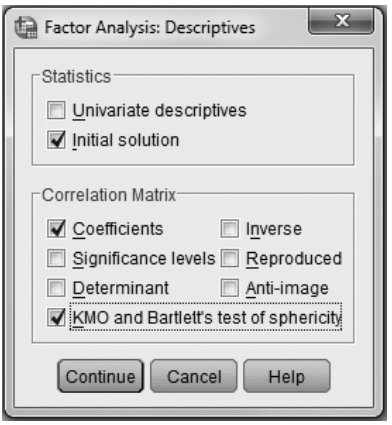
1. From the menu bar, click **Analyze**, then **Dimension Reduction**, and then **Factor**. The following **Factor Analysis** window will open.



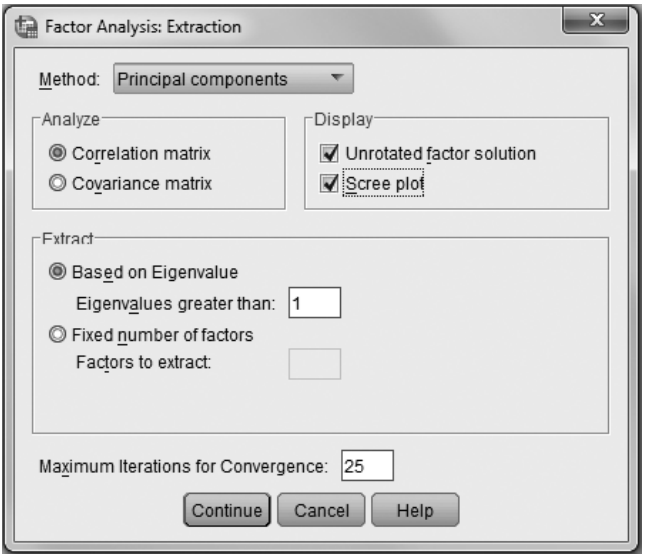
2. Transfer the 12 variables **e1** to **e12** to the **Variables:** field by clicking these variables (highlight) and then clicking .



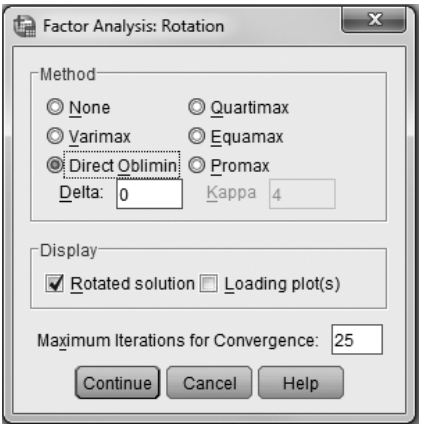
3. In order to (i) obtain a correlation matrix for the 12 variables, and (ii) to test that the correlation matrix has sufficient correlations to justify the application of factor analysis, click Descriptives... . The following **Factor Analysis: Descriptives** window will open. Check the **Coefficients** and **KMO and Bartlett's test of sphericity** fields, and then click Continue.



4. When the **Factor Analysis** window opens, click Extraction... . This will open the **Factor Analysis: Extraction** window below. In the **Method:** drop-down list, choose **Principal components** as the extraction method. Check the **Based on Eigenvalue** field. In the **Eigenvalues greater than:** field, accept the default value of **1**. To obtain a scree plot of the number of factors extracted, check the **Scree plot** field. Click Continue.



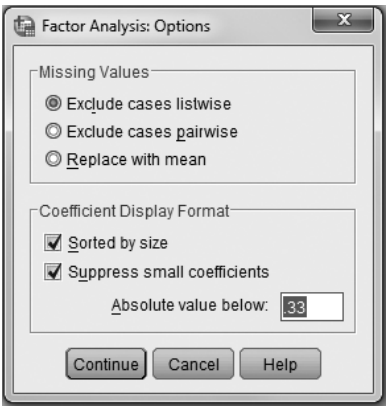
5. When the **Factor Analysis** window opens, click Rotation... . This will open the **Factor Analysis: Rotation** window below. In order to subject the extracted factors to **oblique** rotation, check the **Direct Oblimin** field. Click Continue.



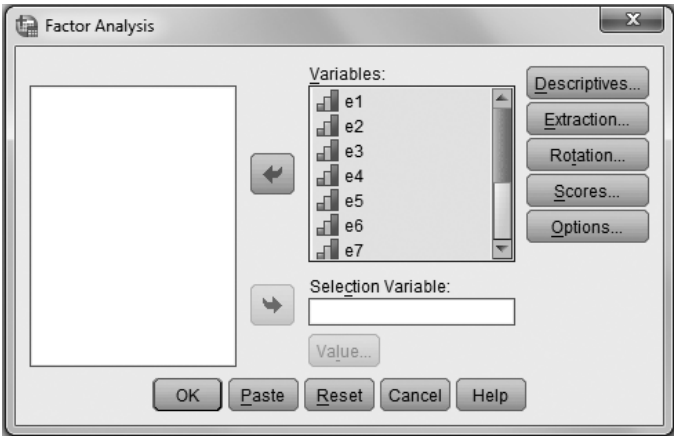
6. When the **Factor Analysis** window opens, click Options... . This will open the following **Factor Analysis: Options** window.

   If the data set has **missing values**, the researcher can choose one of the three methods offered to deal with the missing values: (1) **Exclude cases listwise**—any case (subject) with a missing value for any of the variables in the factor analysis will be excluded from the analysis; this method is the most restrictive as the presence of missing values can reduce the sample size substantially. (2) **Exclude cases pairwise**—any variable in the factor analysis that has a missing value will be excluded from the analysis; this method is less restrictive as only variables (with missing values), rather than cases are excluded. (3) **Replace with mean**—all missing values are replaced with mean values; this method is less restrictive as all variables in the factor analysis will be included in the analysis.

   Under **Coefficient Display Format**, check the **Sorted by size** field**.** This procedure will present the factor loadings (correlation coefficients) in a descending order of magnitude format in the output. Check the **Suppress small coefficients:** field, and in the **Absolute value below:** field, type the coefficient **0.33**. This procedure will suppress the presentation of any factor loadings with values less than 0.33 in the output (an item with a factor loading of 0.33 or higher indicates that approximately 10% or more of the variance in that item is accounted for by its common factor). Click Continue.



7. When the **Factor Analysis** window opens, click OK to complete the analysis. See Table 12.1 for the results.



12.4.4 SPSS Syntax Method: Factor Analysis (First Run)

```
FACTOR VARIABLES = E1 TO E12
/FORMAT = SORT BLANK(.33)
/PRINT = INITIAL EXTRACTION ROTATION CORRELATION KMO
/PLOT = EIGEN
/EXTRACTION = PC
/ROTATION = OBLIMIN.
```

(Note: The method of factor extraction used is **principal component** analysis.)

12.4.5 SPSS Output

TABLE 12.1

Factor Analysis Output

Correlation Matrix

		e1	e2	e3	e4	e5	e6	e7	e8	e9	e10	e11	e12
Correlation	e1	1.000	.593	.352	.268	.648	.534	.458	.296	.672	.394	.438	.197
	e2	.593	1.000	.263	.380	.600	.725	.324	.428	.540	.532	.335	.251
	e3	.352	.263	1.000	.269	.329	.272	.370	.377	.413	.223	.424	.607
	e4	.268	.380	.269	1.000	.254	.369	.572	.476	.299	.235	.467	.285
	e5	.648	.600	.329	.254	1.000	.615	.347	.244	.608	.363	.320	.203
	e6	.534	.725	.272	.369	.615	1.000	.363	.356	.532	.466	.299	.246
	e7	.458	.324	.370	.572	.347	.363	1.000	.492	.415	.168	.597	.222
	e8	.296	.428	.377	.476	.244	.356	.492	1.000	.313	.358	.364	.352
	e9	.672	.540	.413	.299	.608	.532	.415	.313	1.000	.348	.379	.253
	e10	.394	.532	.223	.235	.363	.466	.168	.358	.348	1.000	.285	.192
	e11	.438	.335	.424	.467	.320	.299	.597	.364	.379	.285	1.000	.346
	e12	.197	.251	.607	.285	.203	.246	.222	.352	.253	.192	.346	1.000

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		.865
Bartlett's Test of Sphericity	Approx. Chi-Square	2035.306
	d*f*	66
	Sig.	.000

Communalities

	Initial	Extraction
e1	1.000	.667
e2	1.000	.723
e3	1.000	.795
e4	1.000	.694
e5	1.000	.688
e6	1.000	.683
e7	1.000	.748
e8	1.000	.518
e9	1.000	.623
e10	1.000	.400
e11	1.000	.586
e12	1.000	.797

Extraction Method: Principal Component Analysis.

(*Continued*)

TABLE 12.1 (*Continued*)

Factor Analysis Output

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings[a]
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total
1	5.323	44.360	44.360	5.323	44.360	44.360	4.658
2	1.536	12.796	57.157	1.536	12.796	57.157	3.745
3	1.063	8.861	66.017	1.063	8.861	66.017	2.594
4	.941	7.843	73.860				
5	.635	5.290	79.150				
6	.550	4.580	83.729				
7	.413	3.442	87.171				
8	.372	3.098	90.269				
9	.339	2.827	93.096				
10	.306	2.553	95.649				
11	.298	2.486	98.134				
12	.224	1.866	100.000				

Extraction Method: Principal Component Analysis.

[a] When components are correlated, sums of squared loadings cannot be added to obtain a total variance.

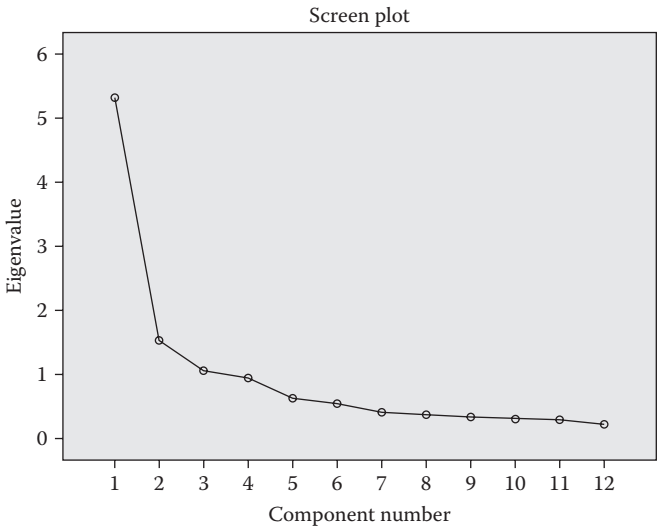


TABLE 12.1 (*Continued*)

Factor Analysis Output

Component Matrix[a]

	Component		
	1	**2**	**3**
e2	.771	−.357	
e1	.760		
e6	.747	−.353	
e9	.746		
e5	.719	−.401	
e7	.668	.360	−.416
e11	.650	.371	
e8	.619	.331	
e4	.599	.375	−.441
e3	.591	.409	.527
e10	.570		
e12	.484	.472	.583

Extraction Method: Principal Component Analysis.

[a] 3 components extracted.

Pattern Matrix[a]

	Component		
	1	**2**	**3**
e5	.870		
e2	.841		
e6	.822		
e1	.785		
e9	.728		
e10	.639		
e4		.884	
e7		.876	
e11		.652	
e8		.611	
e12			.906
e3			.840

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization.

[a] Rotation converged in 6 iterations.

(*Continued*)

TABLE 12.1 (*Continued*)

Factor Analysis Output

Structure Matrix

	Component		
	1	**2**	**3**
e2	.847	.450	
e5	.825	.334	
e6	.823	.432	
e1	.815	.448	
e9	.780	.426	.385
e10	.632		
e7	.435	.863	
e4	.365	.827	
e11	.428	.745	.455
e8	.417	.702	.418
e12		.343	.892
e3	.394	.422	.886

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization.

Component Correlation Matrix

Component	**1**	**2**	**3**
1	1.000	.495	.349
2	.495	1.000	.399
3	.349	.399	1.000

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization.

12.4.6 Results and Interpretation

12.4.6.1 Correlation Matrix

Examination of the **Correlation Matrix** (see Table 12.1) reveals fairly high correlations between the 12 variables written to measure support for specific types of euthanasia. For example, the inter-correlations between the variables e1, e5, and e9 (active-voluntary euthanasia) are greater than 0.33. Similarly, the inter-correlations between e3, e7, and e11 (active-involuntary euthanasia) are also greater than 0.33. Given the number of high inter-correlations between the euthanasia-specific variables, the hypothesized factor model appears to be appropriate.

Bartlett's Test of Sphericity can be used to test for the adequacy of the correlation matrix, that is, the correlation matrix has significant correlations

among at least some of the variables. If the variables are independent, the observed correlation matrix is expected to have small off-diagonal coefficients. Bartlett's Test of Sphericity tests the hypothesis that the correlation matrix is an identity matrix, that is, all the diagonal terms are 1 and all off-diagonal terms are 0. If the test value is large and the significance level is small (<0.05), the hypothesis that the variables are independent can be rejected. In the present analysis, Bartlett's Test of Sphericity yielded a value of 2035.31 and an associated degree of significance smaller than 0.001. Thus, the hypothesis that the correlation matrix is an identity matrix is rejected.

12.4.6.2 Factor Analysis Output

The **Communalities** section presents the **communality** of each variable (i.e., the proportion of variance in each variable accounted for by the common factors). In using the principal component method of factor extraction, it is possible to compute as many factors as there are variables. When all factors are included in the solution, all of the variance of each variable is accounted for by the common factors. Thus, the proportion of variance accounted for by the common factors, or the communality of a variable, is 1 for all the variables.

The **Total Variance Explained** section presents the number of common factors extracted, the eigenvalues associated with these factors, the percentage of total variance accounted for by each factor, and the cumulative percentage of total variance accounted for by the factors. While 12 factors have been extracted, it is obvious that not all 12 factors will be useful in representing the list of 12 variables. In determining how many factors to extract to represent the data, it is helpful to examine the eigenvalues associated with the factors. Using the criterion of retaining only factors with eigenvalues of 1 or greater, the first three factors will be retained for rotation (SPSS default procedure). These three factors account for 44.36%, 12.80%, and 8.86% of the total variance, respectively. That is, 66.02% of the total variance is attributable to these three factors. The remaining nine factors together account for only about 34% of the variance. Therefore, a model with three factors may be adequate to represent the data. From the scree plot, it again appears that a three-factor model should be sufficient to represent the data set.

12.4.6.3 Determining the Number of Factors Using Velicer's Minimum Average Partial (MAP) Test and Parallel Analysis

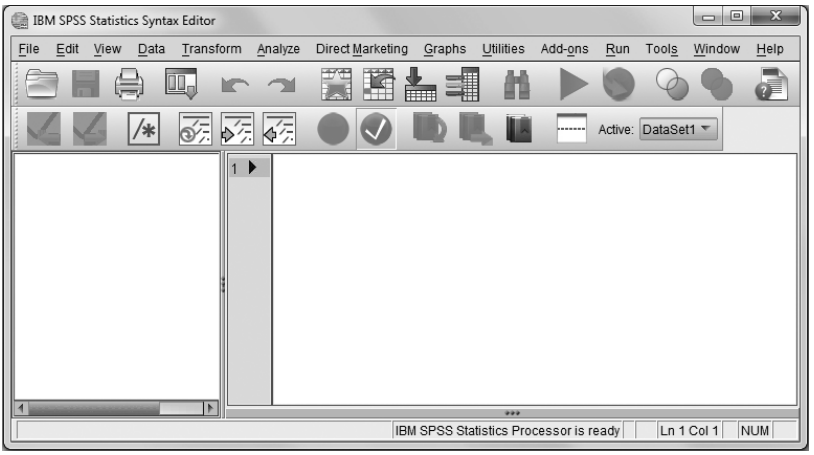
While three factors have been extracted on the basis of the conventional "**eigenvalues ≥ 1.00**" rule, it would be useful to cross-check the number of factors extracted by using the statistically based procedures of **Velicer's MAP test** and **parallel analysis.** While the SPSS statistical software package does not have the intrinsic capacity to conduct these analyses, it does allow users to create customized programs. O'Connor (2000) developed programs

for running both Velicer's MAP test and parallel analysis and these can be downloaded from the following Internet addresses:

- http://flash.lakeheadu.ca/~boconno2/nfactors.html
- http://www.crcpress.com/e_products/downloads/download.asp?cat_no=C6021

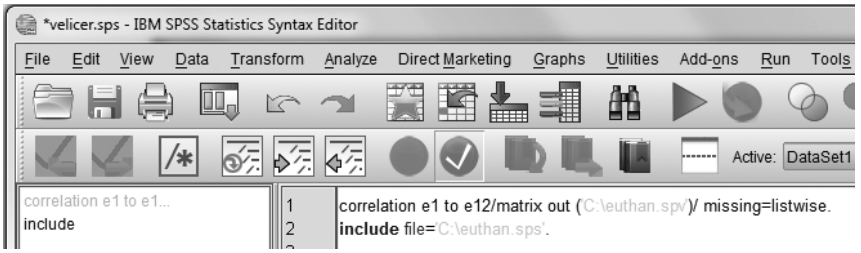
12.4.6.4 Velicer's Minimum Average Partial (MAP) Test

1. Download the MAP program from the above website. Save the program (in the C:\ directory) as a syntax a file called C:**euthan.sps**;
2. Open the data set **EUTHAN_1.SAV**.
3. From the menu bar, click **File**, then **New**, and then **Syntax**. The following syntax window will open.



4. Write the following syntax in the syntax window.

```
correlation e1 to e12/matrix out
('c:\euthan.spv')/missing = listwise.
include file = 'c:\euthan.sps'.
```



5. To conduct the MAP test, highlight the syntax and click ▶.

12.4.6.4.1 MAP Output

	Correlations[a]											
	e1	**e2**	**e3**	**e4**	**e5**	**e6**	**e7**	**e8**	**e9**	**e10**	**e11**	**e12**
e1 Pearson Correlation	1	.593	.352	.268	.648	.534	.458	.296	.672	.394	.438	.197
Sig. (2-tailed)		.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
e2 Pearson Correlation	.593	1	.263	.380	.600	.725	.324	.428	.540	.532	.335	.251
Sig. (2-tailed)	.000		.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
e3 Pearson Correlation	.352	.263	1	.269	.329	.272	.370	.377	.413	.223	.424	.607
Sig. (2-tailed)	.000	.000		.000	.000	.000	.000	.000	.000	.000	.000	.000
e4 Pearson Correlation	.268	.380	.269	1	.254	.369	.572	.476	.299	.235	.467	.285
Sig. (2-tailed)	.000	.000	.000		.000	.000	.000	.000	.000	.000	.000	.000
e5 Pearson Correlation	.648	.600	.329	.254	1	.615	.347	.244	.608	.363	.320	.203
Sig. (2-tailed)	.000	.000	.000	.000		.000	.000	.000	.000	.000	.000	.000
e6 Pearson Correlation	.534	.725	.272	.369	.615	1	.363	.356	.532	.466	.299	.246
Sig. (2-tailed)	.000	.000	.000	.000	.000		.000	.000	.000	.000	.000	.000
e7 Pearson Correlation	.458	.324	.370	.572	.347	.363	1	.492	.415	.168	.597	.222
Sig. (2-tailed)	.000	.000	.000	.000	.000	.000		.000	.000	.001	.000	.000
e8 Pearson Correlation	.296	.428	.377	.476	.244	.356	.492	1	.313	.358	.364	.352
Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000		.000	.000	.000	.000
e9 Pearson Correlation	.672	.540	.413	.299	.608	.532	.415	.313	1	.348	.379	.253
Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000		.000	.000	.000
e10 Pearson Correlation	.394	.532	.223	.235	.363	.466	.168	.358	.348	1	.285	.192
Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.001	.000	.000		.000	.000
e11 Pearson Correlation	.438	.335	.424	.467	.320	.299	.597	.364	.379	.285	1	.346
Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000		.000
e12 Pearson Correlation	.197	.251	.607	.285	.203	.246	.222	.352	.253	.192	.346	1
Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	

[a] Listwise N = 357

```
Run MATRIX procedure:
MGET created matrix CR.
The matrix has 12 rows and 12 columns.
The matrix was read from the record(s) of row type CORR.
```

```
Velicer's Minimum Average Partial (MAP) Test:
Eigenvalues
5.3232
1.5356
1.0633
 .9411
 .6348
 .5496
 .4130
 .3718
 .3392
 .3064
 .2983
 .2239

Average Partial Correlations
```

	Squared	Power4
.0000	.1672	.0413
1.0000	.0475	.0046
2.0000	**.0424**	.0053
3.0000	.0534	.0136
4.0000	.0551	.0142
5.0000	.0776	.0230
6.0000	.1144	.0433
7.0000	.1511	.0633
8.0000	.2102	.1048
9.0000	.3000	.1837
10.0000	.4884	.4001
11.0000	1.0000	1.0000

```
The smallest average squared partial correlation is .0424
The Number of Components According to the Original (1976) MAP
Test is 2

 — — — END MATRIX— — -
```

12.4.6.4.2 Interpretation

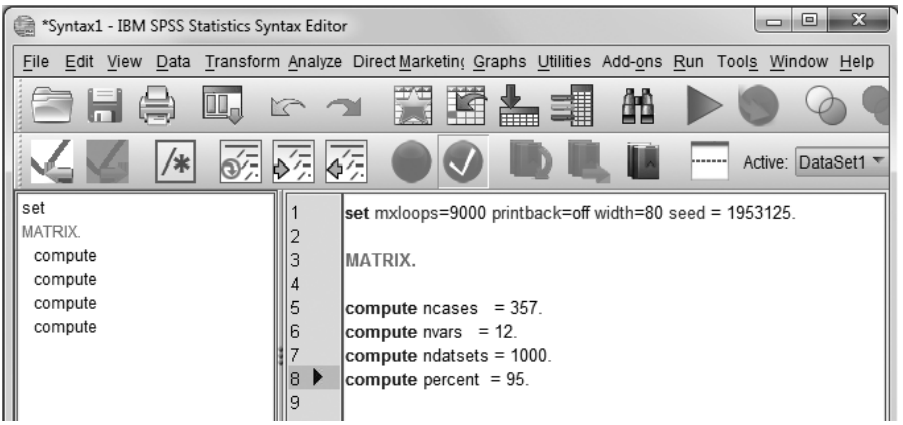
The first section presents the **eigenvalues** generated from a principal component analysis on the 12 variables. These eigenvalues are identical to those generated from the original principal components analysis (see Table 12.1) and show three factors with eigenvalues >1.00.

The second section presented under the heading **Average Partial Correlations** presents the average *R*-square of the correlation matrix after the identified factors have been partialled out. That is, the average R-square is an index of the amount of residual variance after the identified factors have been partialled/extracted out.

What the researcher is looking for is the ***smallest*** "average squared correlation." Looking at the "average squared correlation" results, one can ascertain that the smallest average squared correlation is .0424 and the number of factors corresponding to this value is 2. This indicates that two factors should be retained for interpretation and subsequent rotation.

12.4.6.5 Parallel Analysis

1. Download the parallel analysis program from the above website (http://www.crcpress.com/e_products/downloads/download.asp?cat_no=C6021). Save the program (in the C:\ directory) as a syntax file called **parallel.sps**.
2. Open the data set **EUTHAN_1.SAV**.
3. From the menu bar, click **Open** and then **Syntax** to locate the saved **parallel.sps** file.
4. Open the **parallel.sps** file. Specify the number of cases (n = 357), variables (n = 12), data sets (1000), and the desired percentile (95) at the start of the program. (Note: The length of the program precludes its full inclusion here.)



5. To conduct the parallel analysis, highlight the syntax and click ▶.

12.4.6.5.1 Parallel Analysis Output

```
Run MATRIX procedure:
Specifications for this Run:
Ncases 357
Nvars 12
Ndatsets 1000
Percent 95
```

```
Random Data Eigenvalues
```

Root	Means	Prcntyle
1.000000	1.311082	1.389250
2.000000	1.226860	1.283428
3.000000	1.162877	1.207275
4.000000	1.109321	1.148755
5.000000	1.060342	1.097873
6.000000	1.014231	1.050053
7.000000	.969679	1.003713
8.000000	.925006	.957185
9.000000	.879538	.914133
10.000000	.833172	.871259
11.000000	.784259	.825299
12.000000	.723632	.772224

```
— — — END MATRIX— — -
```

12.4.6.5.2 Interpretation

The eigenvalues are computed from the random data sets. The **Root** column presents the 12 factors. The **Means** column presents the mean eigenvalue for each factor. The **Prcntyle** column presents the eigenvalues that correspond to the 95th percentile of the distribution of the random data eigenvalues. The researcher compares the eigenvalues from the actual data with the above eigenvalues computed from the random data sets. The eigenvalues computed from the actual data can be obtained from the MAP test results above, or by simply running a principal component analysis on the 12 euthanasia variables. To assist with interpretation, the eigenvalues from the actual data set are presented here.

```
Eigenvalues
5.3232
1.5356
1.0633
 .9411
 .6348
 .5496
 .4130
 .3718
 .3392
 .3064
 .2983
 .2239
```

Comparing the eigenvalues from the actual data set with the eigenvalues derived from the random data sets, it can be seen that only the first two factors' eigenvalues (5.3232, 1.5356) are ***larger*** than the corresponding first two 95th percentile (and mean) random data eigenvalues (95th percentile: 1.389250,

1.283428; Means: 1.311082, 1.226860). Therefore, like the MAP test results, two factors should be retained for interpretation and subsequent rotation.

The **Component Matrix** (see Table 12.1) represents the unrotated component analysis factor matrix, and presents the correlations that relate the variables to the three factors using the conventional "eigenvalues ≥ 1.00 rule." These coefficients, called factor loadings, indicate how closely the variables are related to each factor. However, as the factors are unrotated (the factors were extracted on the basis of the proportion of total variance explained), significant cross-loadings have occurred. For example, variables e2, e6, e5, e11, and e8 have loaded highly on Factor 1 and Factor 2; variables e7, e4, e3, and e12 have loaded highly on Factor 1, Factor 2, and Factor 3. These high cross-loadings make interpretation of the factors difficult and theoretically less meaningful.

The **Rotated Component Matrix** presents the three factors after **oblimin** (nonorthogonal) rotation. To subject the three factors to **varimax** (orthogonal) rotation, (1) check the **Varimax** field in the **Factor Analysis: Rotation** window, or (2) substitute the word **OBLIMIN** with **VARIMAX** in the ROTATION sub-command in the SPSS Syntax Method in Section 12.4.4. The VARIMAX rotation output is presented in Table 12.2.

As there is no overwhelming theoretical reason to employ one rotation method over the other, the decision to interpret either the **varimax** rotated matrix or the **oblimin** matrix depends on the magnitude of the factor correlations presented in the **Component Correlation Matrix** (Table 12.2).

TABLE 12.2

Varimax Rotation Output

Rotated Component Matrix[a]			
	Component		
	1	**2**	**3**
e5	.812		
e2	.811		
e6	.790		
e1	.769		
e9	.723		
e10	.610		
e7		.828	
e4		.814	
e11		.667	
e8		.627	
e12			.869
e3			.831

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
[a] Rotation converged in 4 iterations.

Examination of the factor correlations indicates that the three factors are reasonably correlated (all coefficients are larger than 0.33), which suggests that the oblimin (nonorthogonal) matrix should be interpreted. In interpreting the oblimin-rotated matrix, the decision must be made to interpret either the **pattern matrix** or the **structure matrix**. The structure matrix presents the correlations between variables and factors, but these may be confounded by correlations between the factors. The pattern matrix shows the uncontaminated correlations between variables and factors and is generally used for interpreting factors.

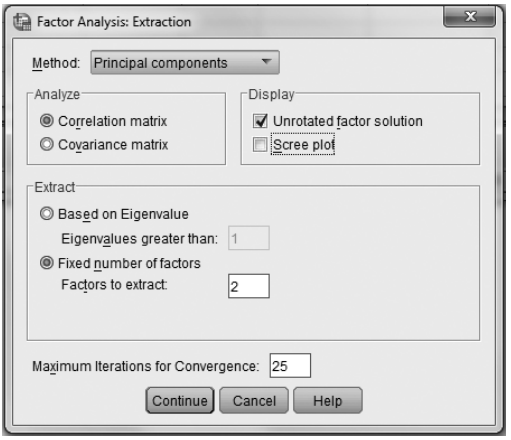
The **rotated pattern matrix** presents the three factors after oblique rotation. In order to identify what these factors represent, it would be necessary to consider what items loaded on each of the three factors. The clustering of the items in each factor and their wording offer the best clue as to the significance of that factor. For example, six items loaded on Factor 1. An inspection of these items (see Section 12.4) clearly indicates that these items reflect the "voluntary" dimension of euthanasia (e.g., terminally ill patients have the right to decide about their own lives and deaths; denying the request of terminally ill patients to "die with dignity" is unfair and cruel). Factor 2 contains four items that clearly reflect the "nonvoluntary" dimension of euthanasia (e.g., doctors have the right to dispense medication that will painlessly end the life of a patient who has been diagnosed to be in a permanent "vegetative" state, if members of the patient's family request it; the withdrawal of medical treatment to allow death to occur should be permissible in cases where a patient, although biologically alive, has been diagnosed as clinically brain-dead). Factor 3 contains two items that also reflect the "nonvoluntary" dimension of euthanasia but related to newborn (e.g., it is all right for doctors who attend childbirth to administer medication that will painlessly end the lives of grossly deformed infants; it is all right for doctors who attend childbirth not to resuscitate grossly deformed babies). These findings suggest that a two-factor solution focusing on the voluntary–nonvoluntary dimension (i.e., the decision to support/not support euthanasia appears to be made on the basis of the presence or absence of the wish of the patient to die) offers the optimal representation of the 12-item euthanasia data set. The determination of a two-factor solution is in line with the findings from both **Velicer's MAP test** and **parallel analysis**. Given that the factor structure appears to be represented by two dimensions of euthanasia—**voluntary** and **nonvoluntary**—it was decided to rerun factor analysis, stipulating the extraction of only two factors.

12.4.7 Windows Method (Second Run)

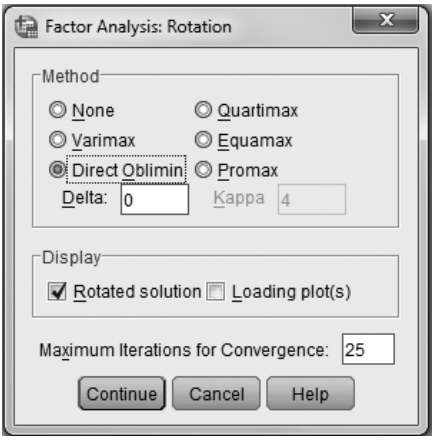
See Section 12.4.3 and repeat step 1 through step 3.

4. Click Extraction... to open the following **Factor Analysis: Extraction** window. In the **Method:** drop-down list, choose **Principal components** as the extraction method. In order to extract only two factors

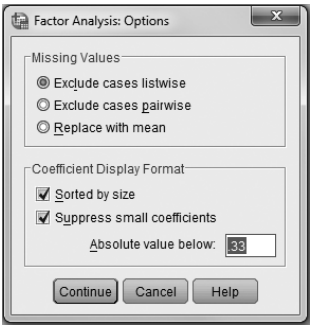
from the correlation matrix, check the **Fixed number of factors:** field, and then type **2** in the **Factors to extract**: field. This procedure will override the default extraction of all factors with eigenvalues greater than 1. Click Continue.



5. When the **Factor Analysis** window opens, click Rotation... . This will open the **Factor Analysis: Rotation** window below. In order to subject the extracted factors to **oblique** rotation, check the **Direct Oblimin** cell. Click Continue.



6. When the **Factor Analysis** window opens, click Options... . This will open the following **Factor Analysis: Options** window. Under **Coefficient Display Format**, check the **Sorted by size** field. Check the **Suppress small coefficients:** field, and in the **Absolute value below:** field, type the coefficient **0.33**. Click Continue.



7. When the **Factor Analysis** window opens, click OK to complete the analysis. See Table 12.3 for the results.

12.4.8 SPSS Syntax Method (Second Run)

```
FACTOR VARIABLES = E1 TO E12
/FORMAT = SORT BLANK(.33)
/CRITERIA = FACTOR(2)
/EXTRACTION = PC
/ROTATION = OBLIMIN.
```

12.4.9 SPSS Output

TABLE 12.3

Two-Factor Structure (Selected) Output

Total Variance Explained

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings[a]
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total
1	5.323	44.360	44.360	5.323	44.360	44.360	4.653
2	1.536	12.796	57.157	1.536	12.796	57.157	4.119
3	1.063	8.861	66.017				
4	.941	7.843	73.860				
5	.635	5.290	79.150				
6	.550	4.580	83.729				
7	.413	3.442	87.171				
8	.372	3.098	90.269				
9	.339	2.827	93.096				
10	.306	2.553	95.649				
11	.298	2.486	98.134				
12	.224	1.866	100.000				

Extraction Method: Principal Component Analysis.

[a] When components are correlated, sums of squared loadings cannot be added to obtain a total variance.

TABLE 12.3 (*Continued*)

Two-Factor Structure (Selected) Output

Component Matrix[a]

	Component	
	1	**2**
e2	.771	−.357
e1	.760	
e6	.747	−.353
e9	.746	
e5	.719	−.401
e7	.668	.360
e11	.650	.371
e8	.619	.331
e4	.599	.375
e3	.591	.409
e10	.570	
e12	.484	.472

Extraction Method: Principal Component Analysis.
[a] 2 components extracted.

Pattern Matrix[a]

	Component	
	1	**2**
e5	.858	
e2	.850	
e6	.829	
e1	.786	
e9	.714	
e10	.633	
e12		.736
e3		.722
e11		.710
e7		.707
e4		.689
e8		.652

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization.
[a] Rotation converged in 5 iterations.

12.4.10 Results and Interpretation

The results presented in the **Total Variance Explained** section (see Table 12.3) are identical to those obtained in the first run (Table 12.1). This is not surprising as the same extraction method (principal components analysis) was

applied to the same 12 items. Thus, the same three factors were extracted accounting for a combined 66.02% of the total variance.

The **rotated pattern matrix** presents only two rotated factors as stipulated in both the SPSS Windows and syntax file methods. The rotated factor structure is "clean" in that there is no cross-loading of items. Factor 1 contains the same six items that were extracted to represent Factor 1 in the first run, and clearly reflect the "voluntary" dimension of euthanasia. Factor 2 is also represented by six items—the combination of four items that loaded on Factor 2 and two items that loaded on Factor 3—from the first run. Examination of these six items clearly shows that they reflect the "nonvoluntary" dimension of euthanasia. This two-factor model represents the combination of the three original factors, and appears to reflect adequately the underlying factor structure of the 12-item euthanasia scale.

12.5 Factor Analysis: Example 2

A survey was designed to identify the motives for the maintenance of smoking behavior and its possible cessation (Ho, 1989). Twenty-five statements were written to represent these motives. Each statement was rated on a five-point scale with higher scores indicating strong agreement with that motive as a reason for smoking. A total of 91 smokers provided responses to these 25 statements. Factor analysis (with principal component extraction, followed by varimax rotation) was used to investigate the factor structure of this 25-item smoking inventory. The 25 statements written to reflect smoking motives are listed in Table 12.4.

TABLE 12.4

Reasons for Smoking

1.	I find smoking enjoyable
2.	When I feel stressed, tense, or nervous, I light up a cigarette
3.	Smoking lowers my appetite and therefore keeps my weight down
4.	Lighting up a cigarette is a habit for me
5.	I smoke cigarettes to alleviate boredom
6.	It gives me something to do with my hands
7.	I feel good when I am smoking
8.	I enjoy lighting up after pleasurable experiences, e.g., after a good meal
9.	Smoking relaxes me
10.	Smoking gives me a lift
11.	Smoking allows me to be "part of a crowd"

TABLE 12.4 (*Continued*)

Reasons for Smoking	
12.	I smoke because I am addicted to cigarettes
13.	I smoke because members of my family smoke
14.	Smoking is a means of socializing
15.	Smoking helps me to concentrate when I am working
16.	I smoke because most of my friends smoke
17.	I smoke because it makes me feel confident
18.	Smoking makes me feel sophisticated and glamorous
19.	I smoke as an act of defiance
20.	I smoke because I find it difficult to quit
21.	I enjoy the taste of cigarettes
22.	I find smoking pleasurable
23.	I smoke to annoy non-smokers
24.	The health statistics regarding smoking cigarettes and health problems don't bother me, as they are highly exaggerated anyway
25.	I am willing to live with my health problems that my smoking may cause me

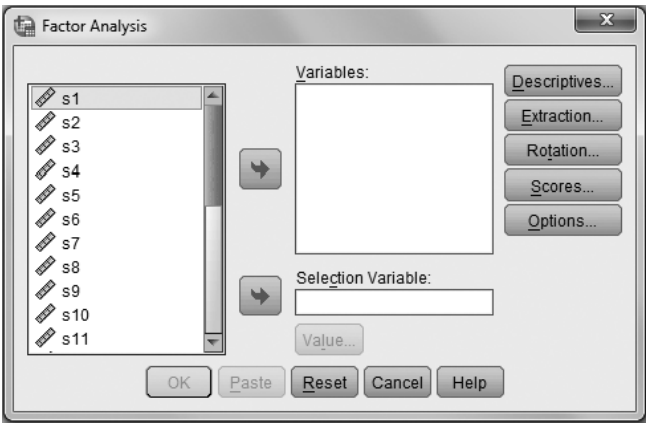
12.5.1 Data Entry Format

The data set has been saved under the name **SMOKE.SAV**.

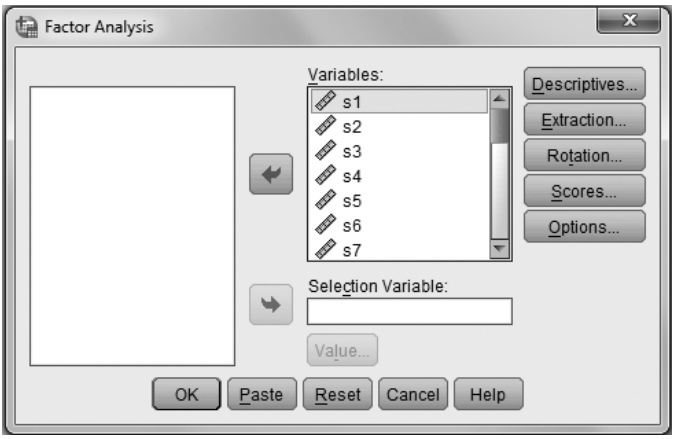
Variables	Columns	Code
s1–s25	1–25	1 = strongly agree, 5 = strongly disagree

12.5.2 Windows Method: Factor Analysis (First Run)

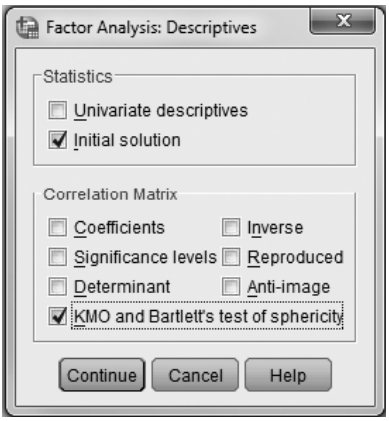
1. From the menu bar, click **Analyze**, then **Dimension Reduction**, and then **Factor**. The following **Factor Analysis** window will open.



2. Transfer the 25 variables **s1** to **s25** to the **Variables:** field by clicking these variables (highlight) and then clicking ▶.

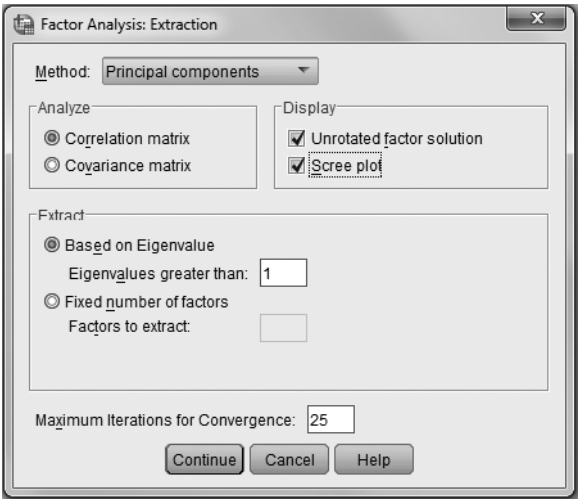


3. In order to test that the correlation matrix, generated from the 25 variables, has sufficient correlations to justify the application of factor analysis, click Descriptives... . The following **Factor Analysis: Descriptives** window will open. Check the **KMO and Bartlett's test of sphericity** field, and then click Continue.

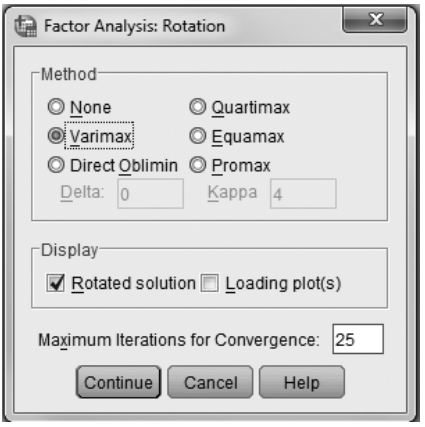


4. When the **Factor Analysis** window opens, click Extraction... . This will open the following **Factor Analysis: Extraction** window. In the **Method:** drop-down list, choose **Principal components** as the extraction method. Ensure that the **Correlation matrix:** field is checked. Check the **Based on Eigenvalue** field. In the **Eigenvalues greater than:** field, accept the default value of **1**. Leave the **Factors to extract:** field blank (i.e., allow principal components analysis to extract as many factors as

there are with eigenvalues > 1). To obtain a scree plot of the number of factors extracted, check the scree plot field. Click Continue.

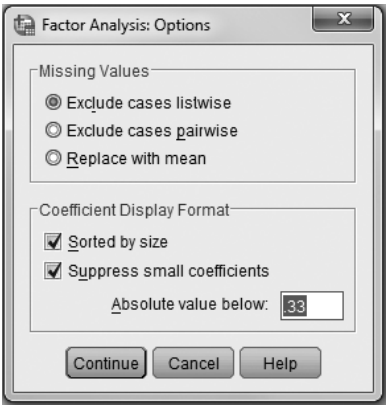


5. When the **Factor Analysis** window opens, click Rotation... . This will open the **Factor Analysis: Rotation** window below. In order to subject the extracted factors to varimax rotation, check the **Varimax** field. Click Continue.

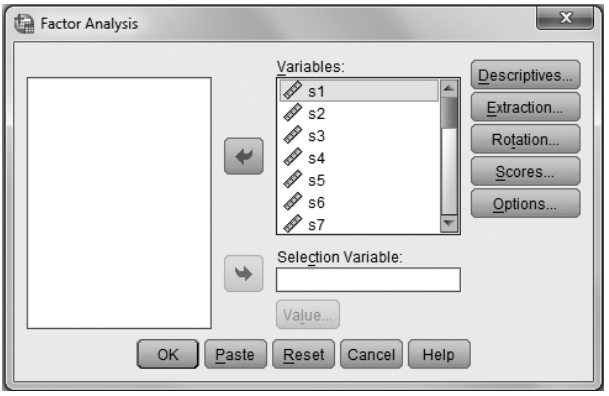


6. When the **Factor Analysis** window opens, click Options... . This will open the following **Factor Analysis: Options** window below. If the data set has **missing values**, the researcher can choose one of the three methods offered to deal with the missing values: (1) **Exclude cases listwise**, (2) **Exclude cases pairwise**, and (3) **Replace with mean** (see Section 12.4.3).

Under **Coefficient Display Format**, check the **Sorted by size** cell**.** This procedure will present the factor loadings (correlation coefficients) in a descending order of magnitude format in the output. Check the **Suppress small coefficients** cell, and then type the coefficient of **0.33** in the **Absolute value below**: field. This procedure will suppress the presentation of any factor loadings with values less than 0.33 in the out-put. Click Continue.



7. When the **Factor Analysis** window opens, click OK to complete the analysis. See Table 12.5 for the results.



12.5.3 SPSS Syntax Method: Factor Analysis (First Run)

```
FACTOR VARIABLES = S1 TO S25
/FORMAT = SORT BLANK(.33)
/PRINT = INITIAL EXTRACTION ROTATION KMO
```

```
/PLOT = EIGEN
/EXTRACTION = PC
/ROTATION = VARIMAX.
```

12.5.4 SPSS Output

TABLE 12.5

Factor Analysis Output

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy		.687
Bartlett's Test of Sphericity	Approx. Chi-Square	876.290
	df	300
	Sig.	.000

Communalities		
	Initial	**Extraction**
S1	1.000	.747
S2	1.000	.636
S3	1.000	.399
S4	1.000	.535
S5	1.000	.751
S6	1.000	.794
S7	1.000	.698
S8	1.000	.687
S9	1.000	.658
S10	1.000	.723
S11	1.000	.722
S12	1.000	.727
S13	1.000	.610
S14	1.000	.703
S15	1.000	.593
S16	1.000	.795
S17	1.000	.698
S18	1.000	.740
S19	1.000	.648
S20	1.000	.695
S21	1.000	.578
S22	1.000	.722
S23	1.000	.558
S24	1.000	.716
S25	1.000	.403

Extraction Method: Principal Component Analysis.

(*Continued*)

TABLE 12.5

Factor Analysis Output (*Continued*)

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5.315	21.258	21.258	5.315	21.258	21.258	4.199	16.797	16.797
2	2.891	11.565	32.823	2.891	11.565	32.823	2.550	10.201	26.998
3	2.636	10.544	43.367	2.636	10.544	43.367	2.169	8.675	35.672
4	1.666	6.662	50.029	1.666	6.662	50.029	2.165	8.660	44.332
5	1.512	6.049	56.078	1.512	6.049	56.078	2.022	8.089	52.421
6	1.324	5.296	61.375	1.324	5.296	61.375	1.979	7.916	60.338
7	1.190	4.760	66.134	1.190	4.760	66.134	1.449	5.797	66.134
8	.937	3.748	69.882						
9	.897	3.590	73.471						
10	.806	3.224	76.696						
11	.732	2.930	79.625						
12	.689	2.755	82.381						
13	.571	2.286	84.666						
14	.561	2.244	86.910						
15	.514	2.057	88.968						
16	.451	1.805	90.773						
17	.370	1.479	92.252						
18	.354	1.414	93.667						
19	.322	1.287	94.954						
20	.307	1.228	96.182						
21	.258	1.034	97.216						
22	.232	.929	98.145						
23	.180	.722	98.866						
24	.166	.666	99.532						
25	.117	.468	100.000						

Extraction Method: Principal Component Analysis.

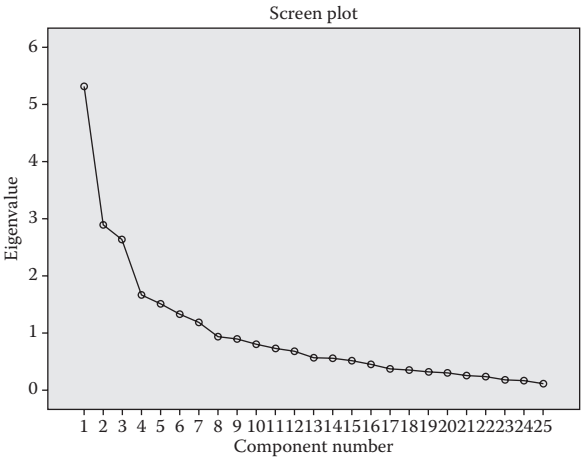


TABLE 12.5

Factor Analysis Output (*Continued*)

Component Matrix[a]

	Component						
	1	**2**	**3**	**4**	**5**	**6**	**7**
S18	.772						
S17	.769						
S16	.764						
S11	.751						
S14	.680						
S7	.676	.442					
S10	.606					−.339	
S13	.576			.357		.347	
S5	.493	.355		.409			−.371
S3	.388		−.337				
S12		.734			.357		
S20		.607					.501
S4		.584		.391			
S2		.547			−.459		
S24	.431	−.448		.377		.357	
S1			.844				
S22			.809				
S21			.623				
S23	.404			.520			
S6	.418	.363		.507			
S19	.361				.552		
S9	.355		.336	−.360	−.440		
S8		.345	.476			.511	
S25							.395
S15					.350	.367	−.384

Extraction Method: Principal Component Analysis.
[a] 7 components extracted.

Rotated Component Matrix[a]

	Component						
	1	**2**	**3**	**4**	**5**	**6**	**7**
S14	.802						
S11	.780						
S18	.770						
S17	.748						
S16	.744		.401				
S10	.592	.441					
S19	.485				−.477		.390
S1		842					
S22		.830					
S21		.691					
S3		−.357					
S24			.791				

(*Continued*)

TABLE 12.5 (*Continued*)

Factor Analysis Output

Rotated Component Matrix[a]

	Component						
	1	**2**	**3**	**4**	**5**	**6**	**7**
S13			.680				
S23			.643				
S20				.809			
S12				.788			
S4				.594		.347	
S2					.737		
S9					.734		
S7	.460				.502	.357	
S5						.812	
S6						.808	
S15							.724
S25							−.536
S8		.383		.346			.493

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
[a] Rotation converged in 9 iterations.

Component Transformation Matrix

Component	**1**	**2**	**3**	**4**	**5**	**6**	**7**
1	.830	.099	.379	.066	.218	.303	.116
2	−.173	−.055	−.316	.727	.427	.343	.195
3	−.144	.945	.038	−.032	.255	−.139	.006
4	−.371	.112	.581	.182	−.383	.552	−.167
5	.188	.199	−.142	.346	−.683	−.190	.532
6	−.250	−.206	.603	.078	.297	−.427	.506
7	.158	−.017	.185	.554	−.063	−.498	−.618

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

12.5.5 Results and Interpretation

12.5.5.1 Correlation Matrix

Bartlett's Test of Sphericity (see Table 12.5) tests the adequacy of the correlation matrix, and yielded a value of 876.29 and an associated level of significance smaller than 0.001. Thus, the hypothesis that the correlation matrix is an identity matrix can be rejected, that is, the correlation matrix has significant correlations among at least some of the variables.

12.5.5.2 Factor Analysis Output

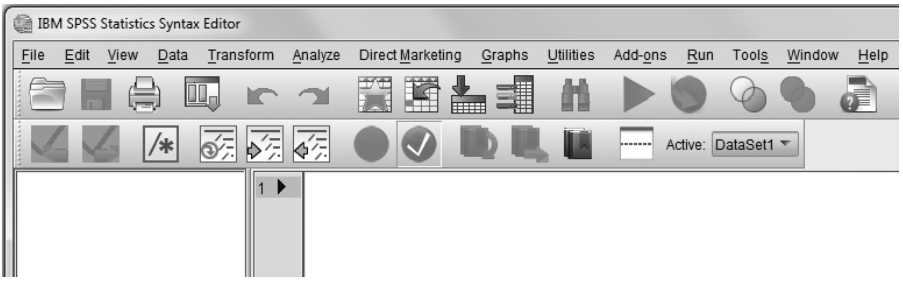
The **Total Variance Explained** section presents the number of common factors extracted, the eigenvalues associated with these factors, the percentage of total variance accounted for by each factor, and the cumulative percentage of total variance accounted for by the factors. Using the criterion of retaining only factors with eigenvalues of 1 or greater, seven factors were retained for rotation. These seven factors accounted for 21.26%, 11.56%, 10.54%, 6.66%, 6.05%, 5.30%, and 4.76% of the total variance, respectively, for a total of 66.13%. The scree plot, however, suggests a four-factor solution.

12.5.5.3 Determining the Number of Factors Using Velicer's Minimum Average Partial (MAP) Test and Parallel Analysis

While seven factors have been extracted on the basis of the conventional "**eigenvalues ≥ 1.00**" rule, it would be useful to cross-check the number of factors extracted by using the statistically based procedures of Velicer's MAP test and parallel analysis**.** As mentioned earlier, both platforms can be downloaded from the following website: http://www.crcpress.com/ e_products/ downloads/download.asp?cat_no=C6021.

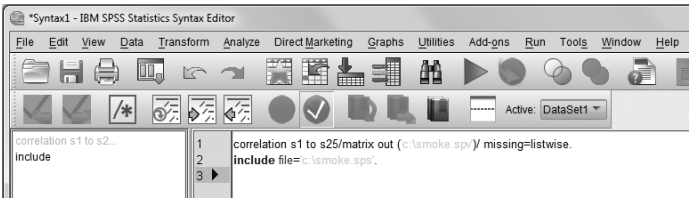
12.5.5.4 Velicer's Minimum Average Partial (MAP) Test

1. Download the MAP program from the above website. Save the program (in the C:\ directory) as a syntax file called **smoke.sps**.
2. Open the data set **SMOKE.SAV**.
3. From the menu bar, click **File**, then **New**, and then **Syntax**. The following syntax window will open.



4. Write the following syntax in the syntax window.

```
correlation s1 to s25/matrix out ('c:\smoke.spv')/
missing = listwise.
include file = 'c:\smoke.sps'.
```



5. To conduct the MAP test, highlight the syntax and click ▶.

12.5.5.4.1 MAP Output

```
Run MATRIX procedure:
MGET created matrix CR.
The matrix has 25 rows and 25 columns.
The matrix was read from the record(s) of row type CORR.

Velicer's Minimum Average Partial (MAP) Test:

Eigenvalues
5.3146
2.8912
2.6360
1.6655
1.5123
1.3241
1.1899
 .9369
 .8974
 .8060
 .7324
 .6888
 .5715
 .5610
 .5144
 .4513
 .3698
 .3536
 .3218
 .3071
 .2585
 .2322
 .1804
 .1664
 .1170
```

Average Partial Correlations

	Squared	Power4
.0000	.0531	.0101
1.0000	.0317	.0033
2.0000	.0271	.0025
3.0000	**.0244**	.0017
4.0000	.0260	.0018
5.0000	.0272	.0020
6.0000	.0268	.0021
7.0000	.0281	.0023
8.0000	.0330	.0035
9.0000	.0379	.0050
10.0000	.0422	.0078
11.0000	.0488	.0092
12.0000	.0568	.0114
13.0000	.0648	.0140
14.0000	.0729	.0154
15.0000	.0820	.0175
16.0000	.0912	.0214
17.0000	.1063	.0272
18.0000	.1266	.0418
19.0000	.1543	.0571
20.0000	.1883	.0762
21.0000	.2444	.1213
22.0000	.3303	.1999
23.0000	.4956	.3655
24.0000	1.0000	1.0000

```
The smallest average squared partial correlation is .0244
The Number of Components According to the Original (1976) MAP
Test is 3
— — — END MATRIX— — -
```

12.5.5.4.2 Interpretation

The first section presents the **eigenvalues** generated from a principal components analysis on the 25 variables. These eigenvalues are identical to those generated from the original principal components analysis (see Table 12.5) and show seven factors with eigenvalues >1.00.

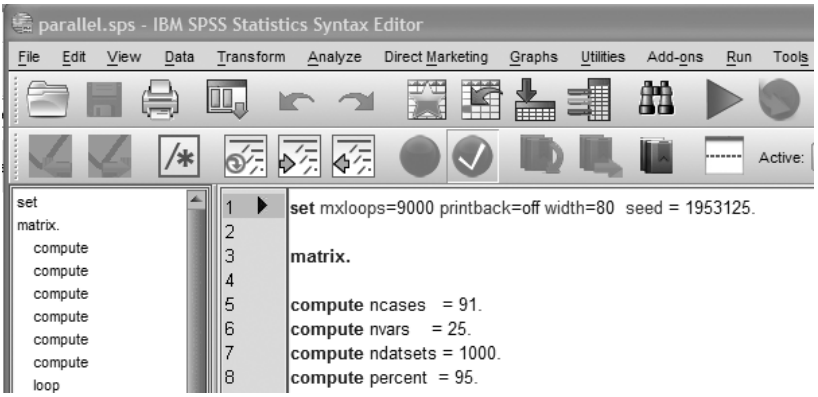
The second section presented under the heading **Average Partial Correlations** presents the average *R*-square of the correlation matrix after the identified factors have been partialled out. That is, the average *R*-square is an index of the amount of residual variance after the identified factors have been partialled/extracted out.

What the researcher is looking for is the ***smallest*** "average squared correlation." Looking at the "average squared correlation" results, one

can ascertain that the smallest average squared correlation is .0244 and the number of factors corresponding to this value is 3. This indicates that three factors should be retained for interpretation and subsequent rotation.

12.5.5.5 Parallel Analysis

1. Download the parallel analysis program from the above website. Save the program (in the C:\ directory) as a syntax file called **parallel.sps**.
2. Open the data set **SMOKE.SAV**.
3. From the menu bar, click **Open** and then **Syntax** to locate the saved **parallel.sps** file.
4. Open the **parallel.sps** file. Specify the number of cases ($n = 91$), variables ($n = 25$), data sets (1000), and the desired percentile (95) at the start of the program. (Note: The length of the program precludes its full inclusion here.)



5. To conduct the parallel analysis, highlight the syntax and click ▶.

12.5.5.5.1 Parallel Analysis Output

```
Run MATRIX procedure:
Specifications for this Run:
Ncases 91
Nvars 25
Ndatsets 1000
Percent 95
Random Data Eigenvalues
```

Root	Means	Prcntyle
1.000000	2.112722	2.296901
2.000000	1.918647	2.049225
3.000000	1.769619	1.878844
4.000000	1.647535	1.738643
5.000000	1.541522	1.628224
6.000000	1.445186	1.520942
7.000000	1.355663	1.432298
8.000000	1.272248	1.345288
9.000000	1.192969	1.259345
10.000000	1.121249	1.185343
11.000000	1.050054	1.112958
12.000000	.980545	1.036869
13.000000	.916364	.974461
14.000000	.854483	.909637
15.000000	.794440	.850977
16.000000	.736530	.789656
17.000000	.683086	.736174
18.000000	.627876	.682287
19.000000	.575534	.622100
20.000000	.525745	.574712
21.000000	.476222	.524795
22.000000	.427651	.472650
23.000000	.378697	.426005
24.000000	.326276	.373178
25.000000	.269139	.320770

```
— — — END MATRIX— — -
```

12.5.5.5.2 Interpretation

The eigenvalues are computed from the random data sets. The **Root** column presents the 25 factors. The **Means** column presents the mean eigenvalue for each factor. The **Prcntyle** column presents the eigenvalues that correspond to the 95th percentile of the distribution of the random data eigenvalues. The researcher compares the eigenvalues from the actual data with the above eigenvalues computed from the random data sets. The eigenvalues computed from the actual data can be obtained from the MAP test results above, or by simply running a principal component analysis on the 25 smoking variables. To assist with interpretation, the eigenvalues from the actual data set are presented here.

```
Eigenvalues
5.3146
2.8912
2.6360
```

```
1.6655
1.5123
1.3241
1.1899
 .9369
 .8974
 .8060
 .7324
 .6888
 .5715
 .5610
 .5144
 .4513
 .3698
 .3536
 .3218
 .3071
 .2585
 .2322
 .1804
 .1664
 .1170
```

Comparing the eigenvalues from the actual data set with the eigenvalues derived from the random data sets, it can be seen that only the first three factors' eigenvalues (5.3146, 2.8912, 2.6360) are ***larger*** than the corresponding first three 95th percentile random data eigenvalues (2.296901, 2.049225, 1.878844). Therefore, like the MAP test results, three factors should be retained for interpretation and subsequent rotation.

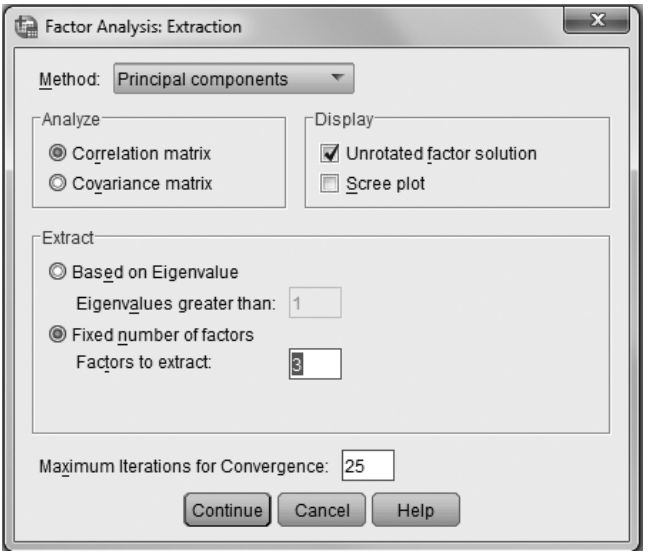
The **Rotated Component Matrix** presents the seven factors after varimax rotation. In order to identify what these factors present, it would be necessary to consider what items loaded on each of the seven factors. The clustering of the items in each factor and their wording offer the best clue as to the significance of that factor. For example, eight items loaded on Factor 1. An inspection of these points (see Table 12.4) clearly indicates that the majority of these items reflect a social motive for smoking (e.g., smoking is a means of socializing; smoking allows me to be part of a crowd; I smoke because most of my friends smoke). Factor 2 contains five items that clearly reflect the **pleasure** that a smoker gains from smoking (e.g., I find smoking enjoyable; I find smoking pleasurable; I love the taste of cigarettes). Factors 4–6 contain items that appear to reflect two related motives—**addiction** and **habit** (e.g., I smoke because I find it hard to quit; I smoke because I am addicted to cigarettes; lighting up is a habit for me; smoking gives me something to do with my hands). The two remaining factors, Factor 3 and Factor 7, contain items that do not "hang" together conceptually, and as such, are not easily interpretable. In fact, some of the items that load on these two factors appear to overlap in meaning with

other factors. For example, item s13 (I smoke because members of my family smoke) in Factor 3 appears to reflect a social motive, and thus overlaps in meaning with Factor 1. Similarly, item s8 (I enjoy lighting up after pleasurable experiences) in Factor 7 appears to overlap in meaning with Factor 2 (pleasure motive). The commonality in meaning of some of these factors suggests that a number of factors can be combined. The combination of factors is purely a subjective decision, aimed at reducing the number of extracted factors to a smaller, more manageable, and ultimately, more meaningful set of factors. Given that the present factor structure appears to be represented by three dimensions of smoking motives—**social**, **pleasure**, and **addiction/habit**—it was decided to rerun factor analysis, stipulating the extraction of only three factors.

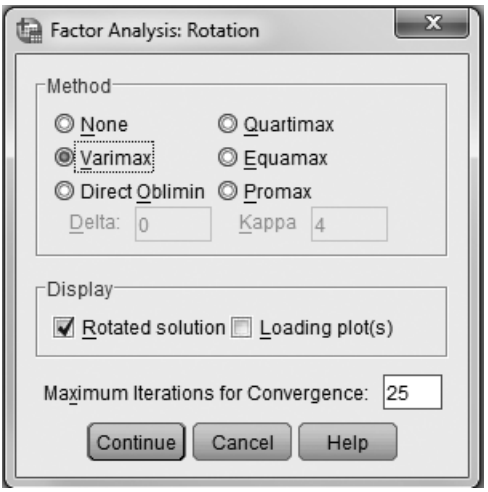
12.5.6 Windows Method (Second Run)

See Section 12.5.2 and repeat step 1 through step 3.

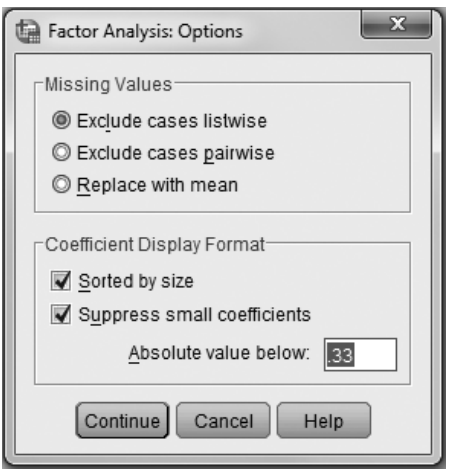
4. Click Extraction... to open the **Factor Analysis: Extraction** window below. In the **Method:** drop-down list, choose **Principal components** as the extraction method. In order to extract only three factors from the correlation matrix, check the **Fixed number of factors:** field, and then type **3** in the **Factors to extract**: field. This procedure will override the default extraction of all factors with eigenvalues greater than 1. Click Continue.



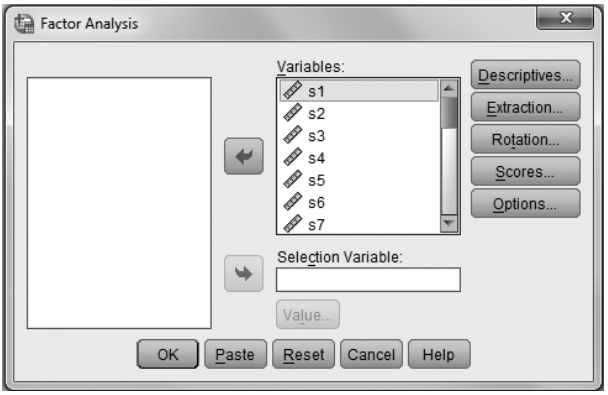
5. When the **Factor Analysis** window opens, click Rotation... . This will open the **Factor Analysis: Rotation** window below. In order to subject the extracted factors to varimax rotation, check the **Varimax** cell. Click Continue.



6. When the **Factor Analysis** window opens, click Options... . This will open the **Factor Analysis: Options** window below. Under **Coefficient Display Format**, check the **Sorted by size** field. Check the **Suppress small coefficients** cell, and then type the coefficient **0.33** in the **Absolute value below**: field. This procedure will suppress the presentation of any factor loadings with values less than 0.33 in the output. Click Continue.



7. When the **Factor Analysis** window opens, click OK to complete the analysis. See Table 12.6 for the results.



12.5.7 SPSS Syntax Method (Second Run)

```
FACTOR VARIABLES = S1 TO S25
/FORMAT = SORT BLANK(.33)
/CRITERIA = FACTOR(3)
/EXTRACTION = PC
/ROTATION = VARIMAX.
```

12.5.8 SPSS Output

TABLE 12.6

Three-Factor Structure Output

Total Variance Explained

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5.315	21.258	21.258	5.315	21.258	21.258	4.953	19.814	19.814
2	2.891	11.565	32.823	2.891	11.565	32.823	3.163	12.651	32.465
3	2.636	10.544	43.367	2.636	10.544	43.367	2.725	10.902	43.367
4	1.666	6.662	50.029						
5	1.512	6.049	56.078						
6	1.324	5.296	61.375						
7	1.190	4.760	66.134						
8	.937	3.748	69.882						
9	.897	3.590	73.471						
10	.806	3.224	76.696						
11	.732	2.930	79.625						
12	.689	2.755	82.381						
13	.571	2.286	84.666						

(*Continued*)

TABLE 12.6 (*Continued*)

Three-Factor Structure Output

Total Variance Explained

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
14	.561	2.244	86.910						
15	.514	2.057	88.968						
16	.451	1.805	90.773						
17	.370	1.479	92.252						
18	.354	1.414	93.667						
19	.322	1.287	94.954						
20	.307	1.228	96.182						
21	.258	1.034	97.216						
22	.232	.929	98.145						
23	.180	.722	98.866						
24	.166	.666	99.532						
25	.117	.468	100.000						

Extraction Method: Principal Component Analysis.

Component Matrix[a]

	Component		
	1	2	3
S18	.772		
S17	.769		
S16	.764		
S11	.751		
S14	.680		
S7	.676	.442	
S10	.606		
S13	.576		
S5	.493	.355	
S6	.418	.363	
S23	.404		
S3	.388		−.337
S19	.361		
S9	.355		.336
S15			
S12		.734	
S20		.607	
S4		.584	
S2		.547	
S24	.431	−.448	

TABLE 12.6 (*Continued*)

Three-Factor Structure Output

Component Matrix[a]

	Component		
	1	**2**	**3**
S1			.844
S22			.809
S21			.623
S8		.345	.476
S25			

Extraction Method: Principal Component Analysis.
[a] 3 components extracted.

Rotated Component Matrix[a]

	Component		
	1	**2**	**3**
S11	.796		
S16	.790		
S18	.783		
S17	.750		
S14	.721		
S13	.567		
S24	.525		
S23	.486		
S19	.463		
S10	.439	.370	.421
S12		.685	
S7	.478	.643	
S20		.607	
S4		.561	
S2		.555	
S5	.350	.502	
S6		.491	
S3		.421	
S15			
S1			.850
S22			.827
S21			.629
S8		.356	.487
S9			.392
S25			

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
[a] Rotation converged in 4 iterations.

12.5.9 Results and Interpretation

The results presented in the **Total Variance Explained** section (see Table 12.6) are identical to those obtained in the first run (Table 12.5). This is not surprising as the same extraction method (principal components analysis) was applied to the same 25 items. Thus, the same seven factors were extracted accounting for a combined 66.13% of the total variance.

The **Rotated Component Matrix** presents only three rotated factors as stipulated in both the SPSS Windows and syntax file methods. The rotated factor structure shows a number of cross-loaded items (s10, s7, s5, s8). Convention suggests three possible ways of handling significant cross-loadings.

1. If the matrix indicates many significant cross-loadings, this may suggest further commonality between the cross-loaded variables and the factors. The researcher may decide to rerun factor analysis again, stipulating a smaller number of factors to be extracted.
2. Examine the wording of the cross-loaded variables, and based on their face validity, assign them to the factors that they are most conceptually/logically representative of.
3. Delete all cross-loaded variables. This will result in "clean" factors and will make interpretation of the factors that much easier. This method works best when there are only few significant cross-loadings.

The four cross-loaded items were deleted before interpretation. Deletion of cross-loaded items serves to clarify the factors and makes their interpretation easier. Factor 1 contains nine items that clearly reflect the social motive for smoking, and was thus labeled **SOCIAL**. Factor 2 contains six items that reflect addiction and habit as motives for smoking, and was labeled **ADDICTION/HABIT**. Factor 3 contains four items that reflect the pleasure gained from smoking, and was labeled **PLEASURE**. The determination of a three-factor solution is in line with the findings from both Velicer's MAP test and parallel analysis. This three-factor model represents the combination of the seven original factors, and appears to reflect adequately the underlying factor structure of the 25-item smoking inventory.

13

Reliability

13.1 Aim

The reliability of a measuring instrument is defined as the ability of the instrument to measure consistently the phenomenon it is designed to assess. Reliability, therefore, refers to **test consistency**. The importance of reliability lies in the fact that it is a prerequisite for the validity of a test. Simply, for the validity of a measuring instrument to be sustained, it must be demonstrably reliable. Any measuring instrument that does not reflect some attribute consistently has little chance of being considered a valid measure of that attribute.

Several methods exist for determining the reliability of a measuring instrument. These methods may be divided into two broad categories: **external consistency** procedures, and **internal consistency** procedures.

13.1.1 External Consistency Procedures

External consistency procedures utilize cumulative test results against themselves as a means of verifying the reliability of the measure. Two major methods of ascertaining the reliability of a test by external consistency are:

1. **Test-retest**—Test results for a group of people are compared in two different time periods.
2. **Parallel forms of the same test**—Two sets of results from equivalent but different tests are compared.

13.1.2 Internal Consistency Procedures

Internal consistency refers to the extent to which the items in a test measure the same construct. Items that measure the same phenomenon should logically "cling/hang" together in some consistent manner. Examining the internal consistency of the test enables the researcher to determine which items are not consistent with the test in measuring the phenomenon under investigation.

The object is to remove the inconsistent items and improve the internal consistency of the test. An internally consistent test increases the chances of the test being reliable. Three major methods of ascertaining the reliability of a test by internal consistency are:

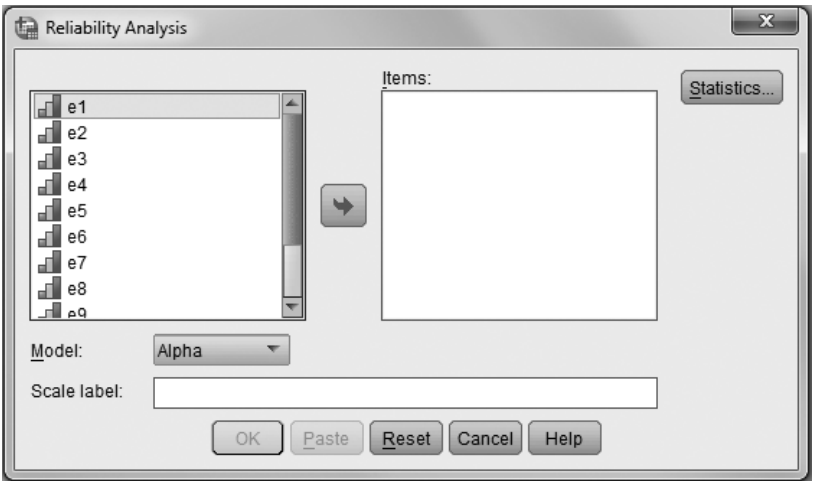
1. **Split-half technique**—This method correlates one half of the test items with the other half of them. The higher the correlation the more the measure is internally consistent. The **Spearman-Brown** formula is widely used in determining reliability by the split-half method.
2. **Cronbach's alpha**—This is a single correlation coefficient that is an estimate of the average of all the correlation coefficients of the items within a test. If alpha is high (0.80 or higher), then this suggests that all of the items are reliable and the entire test is internally consistent. If alpha is low, then at least one of the items is unreliable, and must be identified via item analysis procedure.
3. **Item analysis**—-This procedure represents a refinement of test reliability by identifying "problem" items in the test, i.e., those items that yield low correlations with the sum of the scores on the remaining items. Rejecting those items that are inconsistent with the rest (and retaining those items with the highest average inter-correlations) will increase the internal consistency of the measuring instrument. Item analysis is achieved through the **corrected item-total correlation** procedure.

13.2 Example: Reliability

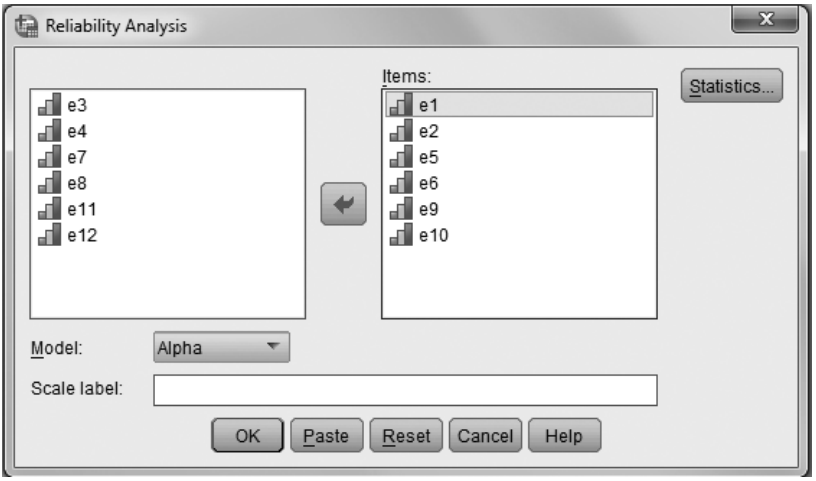
In Example 1 in Chapter 12 on factor analysis, two factors were identified to reflect two dimensions of euthanasia. These two factors (extracted from 12 variables) represent the attitudes toward voluntary and nonvoluntary euthanasia. Before these factors can be computed and applied in subsequent analyses, the internal consistency of each factor should be tested to ensure the reliability of the factors. For this example, only the reliability of the **voluntary euthanasia** factor will be tested. The data set to be analyzed is the same as the one used in Example 1 on factor analysis (**EUTHAN_1.SAV**).

13.2.1 Windows Method

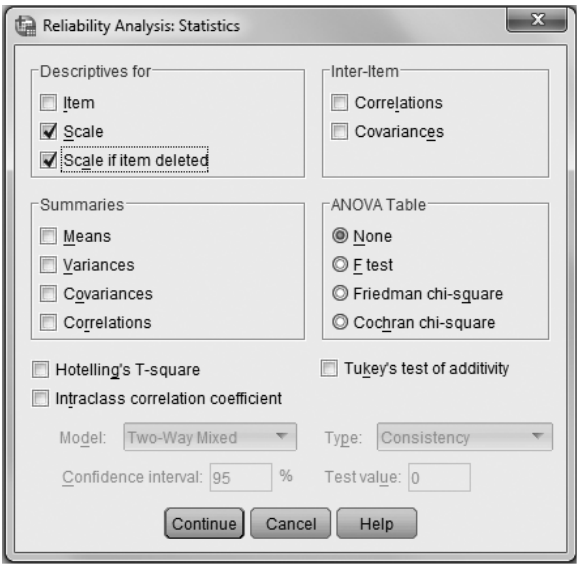
1. From the menu bar, click **Analyze**, then **Scale**, and then **Reliability Analysis**. The following **Reliability Analysis** window will open.



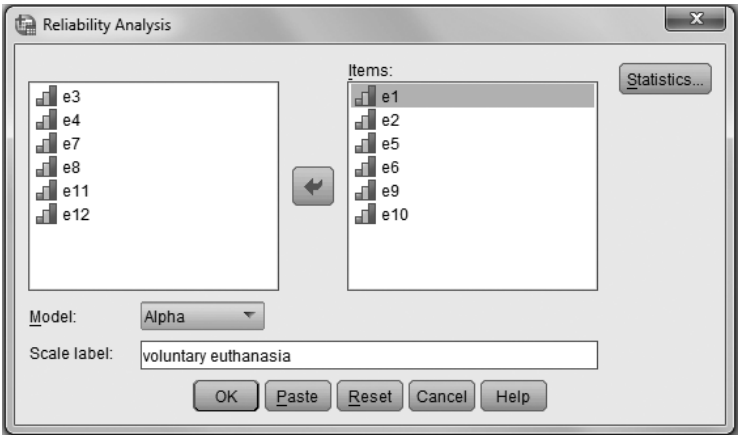
2. Since the **VOLUNTARY EUTHANASIA** factor incorporates the measurement items e1, e2, e5, e6, e9, and e10, transfer these six items to the **Items:** field by clicking these items (highlight) and then clicking . In the **Model:** field select **Alpha** from the drop-down list. **Alpha** (**Cronbach's alpha**) is the default for testing the overall internal consistency of a factor.



3. Click Statistics... to open the following **Reliability Analysis: Statistics** window. In the **Descriptives for** section, check the **Scale** and the **Scale if item deleted** fields. Click Continue.



4. When the **Reliability Analysis** window opens, type **voluntary euthanasia** in the **Scale label:** field. Click OK to complete the analysis. See Table 13.1 for the results.



13.2.2 SPSS Syntax Method

```
RELIABILITY VARIABLES = E1 TO E12
/SCALE(VOLUNTARY_EUTHANASIA) = E1 E2 E5 E6 E9 E10
/STATISTICS = SCALE
/SUMMARY = TOTAL.
```

13.2.3 SPSS Output

TABLE 13.1

Reliability Output

Reliability

Scale: Voluntary Euthanasia

Case Processing Summary

		N	%
Cases	Valid	357	100.0
	Excluded[a]	0	.0
	Total	357	100.0

Reliability Statistics

Cronbach's Alpha	*N* of Items
.868	6

Item-Total Statistics

	Scale Mean If Item Deleted	Scale Variance If Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha If Item Deleted
e1	21.02	14.098	.717	.838
e2	20.65	16.206	.758	.836
e5	20.66	15.659	.713	.839
e6	20.79	15.893	.715	.839
e9	20.95	15.059	.679	.844
e10	21.06	15.881	.500	.879

Scale Statistics

Mean	Variance	Standard Deviation	Number of Items
25.03	21.699	4.658	6

[a] Listwise deletion based on all variables in the procedure.

13.2.4 Results and Interpretation

Cronbach's alpha is 0.87, which indicates high overall internal consistency among the six items representing the voluntary euthanasia factor. The **corrected item-total correlation** shows the correlation (consistency) between

each item and the sum of the remaining items. In deciding which items to retain or delete, the 0.33 criterion can be used (an item-total correlation of 0.33 indicates that approximately 10% of the variance in the scale is accounted for by that item). Based on this criterion, all six items will be kept. Indeed, deleting any of the items (except item e10) will reduce the overall reliability of the scale, as indicated by the column **Cronbach's Alpha If Item Deleted**.

14

Multiple Regression

14.1 Aim

Multiple regression is a statistical technique through which one can analyze the relationship between a dependent or criterion variable and a set of independent or predictor variables. As a statistical tool, multiple regression is often used to accomplish three objectives.

1. To find the best prediction equation for a set of variables, that is, given X and Y (the predictors), what is Z (the criterion variable)?
2. To control for confounding factors in order to assess the contribution of a specific variable or set of variables, that is, identifying independent relationships.
3. To find structural relationships and provide explanations for seemingly complex multivariate relationships, such as is done in path analysis.

14.2 Multiple Regression Techniques

There are three major multiple regression techniques: **standard multiple regression**, **hierarchical regression**, and **statistical (stepwise) regression**. They differ in terms of how the overlapping variability owing to correlated independent variables is handled, and who determines the order of entry of independent variables into the equation (Tabachnick and Fidell, 2001).

14.2.1 Standard Multiple Regression

For this regression model, all the study's independent variables are entered into the regression equation at a time. Each independent variable is then assessed in terms of the unique amount of variance it accounts for. The

disadvantage of the standard regression model is that it is possible for an independent variable to be strongly related to the dependent variable, and still be considered an unimportant predictor, if its unique contribution in explaining the dependent variable is small.

14.2.2 Hierarchical Multiple Regression

This regression model is more flexible as it allows the researcher to specify the order of entry of the independent variables in the regression equation. Each independent variable is assessed at its own point of entry in terms of the additional explanatory power it contributes to the equation. The order of entry is usually dictated by logical or theoretical considerations. For example, on the basis of theoretical reasons, a researcher may determine that two specific independent variables (from a set of independent variables) will be the strongest predictors of the dependent variable. Hence, these two independent variables will be accorded priority of entry, and their total explanatory power (in terms of the total amount of variance explained) evaluated. Then the less important independent variables are entered and evaluated in terms of what they add to the explanation above and beyond that afforded by the first two independent variables. It is also possible to select the opposite tack in which less important independent variables are entered into the equation first, to cull away "nuisance" variance. Then the important set of independent variables is entered and evaluated in terms of what it adds to the explanation of the dependent variable.

14.2.3 Statistical (Stepwise) Regression

For this statistical regression model, the order of entry of predictor variables is based solely on statistical criteria. Variables that correlate most strongly with the dependent variable will be afforded priority of entry, with no reference to theoretical considerations. The disadvantage of this type of regression is that the statistical criteria used for determining priority of entry may be specific to the sample at hand. For another sample, the computed statistical criteria may be different, resulting in a different order of entry for the same variables. The statistical regression model is applied primarily in exploratory work, in which the researcher is unsure about the relative predictive power of the study's independent variables.

Statistical regression can be accomplished through one of following three methods: **forward selection**, **backward deletion**, and **stepwise regression**.

In forward selection, the variables are assessed against a set of statistical standards, and if they meet these criteria, they are afforded priority of entry on the basis of their relative correlations with the dependent variable. The variable that correlates most strongly with the dependent variable gets entered into the equation first, and once in the equation, it stays in the equation.

In **backward deletion**, the equation starts out with all the independent variables entered. Each variable is then evaluated one at a time, in terms of its contribution to the regression equation. Those variables that do not contribute significantly are deleted.

In **stepwise regression**, variables are evaluated for entry into the equation under both forward selection and backward deletion criteria. That is, variables are entered one at a time if they meet the statistical criteria, but they may also be deleted at any step where they no longer contribute significantly to the regression model.

14.3 Checklist of Requirements

- The size of the sample has a direct impact on the statistical power of the significance testing in multiple regression. **Power** in multiple regression refers to the probability of detecting as statistically significant a specific level of *R*-square, or a regression coefficient at a specified significance level and a specific sample size (Hair, Anderson, Tatham, and Black, 1995). Thus, for a desired level of power and with a specified number of independent variables, a certain sample size will be required to detect a significant *R*-square at a specified significance level (see Cohen and Cohen, 1983, for sample size calculations). As a rule of thumb, there should be at least 20 times more cases than independent variables. That is, if a study incorporates five independent variables, there should be at least 100 cases.
- The measurement of the variables can be either continuous (metric) or dichotomous (non-metric). When the dependent variable is dichotomous (coded 0–1), discriminant analysis is appropriate. When the independent variables are discrete, with more than two categories, they must be converted into a set of dichotomous variables by dummy variable coding. A dummy variable is a dichotomous variable that represents one category of a non-metric independent variable. Any non-metric variable with K categories can be represented as K-1 dummy variables, that is, one dummy variable for each degree of freedom. For example, an ethnicity variable may originally be coded as a discrete variable with 1 for Australian, 2 for Chinese, 3 for European, and 4 for "other. " The variable can be converted into a set of three dummy variables (Australian versus non-Australian, Chinese versus non-Chinese, European versus non-European), one for each degree of freedom. This new set of dummy variables can be inserted into the regression equation.

14.4 Assumptions

- **Linearity—**As regression analysis is based on the concept of correlation, the linearity of the relationship between dependent and independent variables is important. Linearity can easily be examined by residual plots.
- **Homoscedasticity**—The assumption of equal variances between pairs of variables can also be detected by residual plots.
- **Independence of error terms**—In regression, it is assumed that the predicted value is not related to any other prediction; that is, each predicted value is independent. Violation of this assumption can be detected by plotting the residuals against a sequence of cases. If the residuals are independent, the pattern should appear random. Violations will be indicated by a consistent pattern in the residuals. Violation of this assumption can also be detected by the Durbin-Watson statistic. If the Durbin-Watson d statistic is between the two critical values of $1.5 < d < 2.5$, it can assumed that there is no linear auto-correlation in the data.
- **Normality**—It is assumed that errors of prediction (differences between the obtained and predicted dependent variable scores) are normally distributed. Violation of this assumption can be detected by an examination of the residual plots.

14.5 Multicollinearity

Multicollinearity refers to the situation where the independent/predictor variables are highly correlated. When independent variables are multicollinear, there is "overlap" or sharing of predictive power. This may lead to the paradoxical effect whereby the regression model fits the data well, but none of the predictor variables has a significant impact in predicting the dependent variable. This is because when the predictor variables are highly correlated, they share essentially the same information. So, together, they may explain a great deal of the dependent variable, but individually may not contribute significantly to the model. Thus, the impact of multicollinearity is to reduce any individual independent variable's predictive power by the extent to which it is associated with the other independent variables. That is, none of the predictor variables may contribute uniquely and significantly to the prediction model after the others are included.

14.5.1 Checking for Multicollinearity

In SPSS, it is possible to request the display of **Tolerance** and **VIF** values for each predictor as a check for multicollinearity. A tolerance value can be described in this manner. From a set of three predictor variables, use X_1 as a dependent variable, and X_2 and X_3 as predictors. Compute the R^2 (the proportion of variance that X_2 and X_3 explain in X_1), and then take $1 - R^2$. Therefore, the tolerance value is an indication of the percent of variance in the predictor that cannot be accounted for by the other predictors. Hence, very small values indicate "overlap" or sharing of predictive power (i.e., the predictor is redundant). Values that are less than 0.10 may merit further investigation. The VIF, which stands for *variance inflation factor*, is computed as "1/tolerance," and it is suggested that predictor variables whose VIF values are greater than 10 may merit further investigation.

Most multiple regression programs have default values of tolerance that will not admit multicollinear variables. Another way to handle the problem of multicollinearity is to either retain only one variable to represent the multicollinear variables, or combine the highly correlated variables into a single composite variable.

14.6 Example 1: Prediction Equation and Identification of Independent Relationships (Forward Entry of Predictor Variables)

A study was designed to investigate how the defenses of **self-defense**, **provocation**, and **insanity** influenced jurors' verdict judgments in trials of battered women who killed their abusive spouses (Ho and Venus, 1995). Eight statements were written to reflect these three defense strategies. Each statement was rated on an eight-point scale, with high scores indicating strong support for that particular defense strategy. A total of 397 subjects provided responses to these eight statements.

Assume that the researcher is interested in predicting the level of responsibility (rated on an eight-point scale, with 1 = not at all responsible, to 8 = entirely responsible) attributed to the battered woman for her fatal action, from the scores obtained from the three defense strategies of self-defense, provocation, and insanity (items are rated on an eight-point scale, with 1 = strongly disagree, to 8 = strongly agree). Specifically, the researcher is interested in (1) predicting the degree of responsibility attributed by a subject who strongly believes that the battered woman's action was motivated by self-defense and provocation (a score of 8 on both scales) and not by an impaired mental state (a score of 1 on the insanity scale), and (2) identifying

the independent relationships between the three defense strategies and the dependent variable of responsibility attribution (coded RESPON). Multiple regression will be used to accomplish these two objectives.

The eight statements (together with their SPSS variable name) written to reflect the three defense strategies are listed below.

Provocation Defense

- **PROVO**—In killing her husband, the defendant's action reflects a sudden and temporary loss of self-control as a result of the provocative conduct of the deceased.
- **PASSION**—In killing her husband, the defendant acted in the heat of passion as a response to the deceased sudden provocation on that fateful day.

Self-Defense

- **PROTECT**—In killing her husband, the defendant was justified in using whatever force (including lethal force) to protect herself.
- **SAVE**—The defendant's lethal action was justified in that she acted to save herself from grievous bodily harm.
- **DEFEND**—In killing her husband, the defendant used such force as was necessary to defend herself.

Insanity Defense

- **MENTAL**—The action of the defendant is the action of a mentally impaired person.
- **INSANE**—In killing her husband, the defendant was either irrational or insane.
- **STABLE**—The action of the accused is not typical of the action of a mentally stable person.

14.6.1 Data Entry Format

Note: The survey questionnaire used in this study was designed as part of a larger study. It contains additional variables apart from the eight defense strategy variables. The data set is named **DOMES.SAV**.

14.6.2 Windows Method: Computation of Factors

The three defense strategies will be computed from the variables written to represent these three strategies. These three defense strategies are coded:

PROVOKE—provocation defense

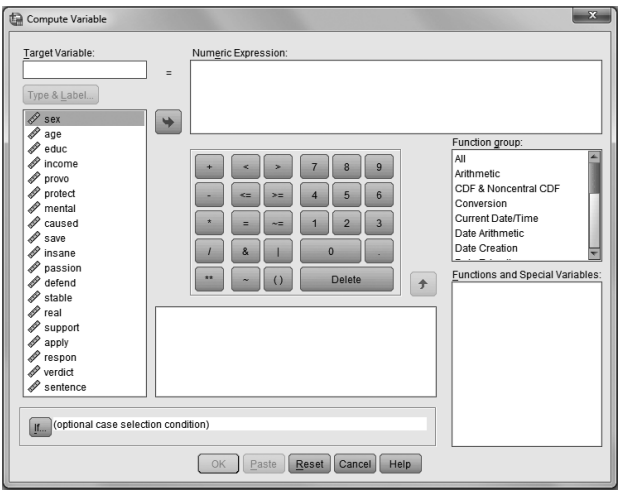
SELFDEF—self-defense defense

INSANITY—insanity defense

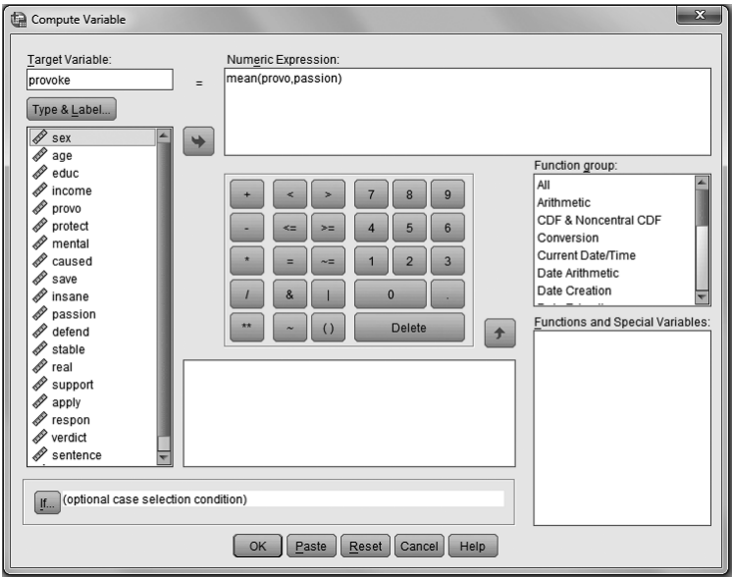
Variables	Column(s)	Code
Sex	1	1 = male, 2 = female
Age	2	in years
Educ	3	1 = primary, to 6 = tertiary
Income	4	1 = <$10,00 per year, 5 = ≥$40,000 per year
Provo	5	1 = strongly disagree, 8 = strongly agree
Protect	6	1 = strongly disagree, 8 = strongly agree
Mental	7	1 = strongly disagree, 8 = strongly agree
Caused	8	1 = strongly disagree, 8 = strongly agree
Save	9	1 = strongly disagree, 8 = strongly agree
Insane	10	1 = strongly disagree, 8 = strongly agree
Passion	11	1 = strongly disagree, 8 = strongly agree
Defend	12	1 = strongly disagree, 8 = strongly agree
Stable	13	1 = strongly disagree, 8 = strongly agree
Real	14	1 = totally unbelievable syndrome, 8 = totally believable syndrome
Support	15	1 = no support at all, 8 = a great deal of support
Apply	16	1 = does not apply to the defendant at all, 8 = totally applies to the defendant
Respon	17	1 = not at all responsible, 8 = entirely responsible
Verdict	18	1 = not guilty, 2 = guilty of manslaughter, 3 = guilty of murder
Sentence	19	1 = 0 years, 6 = life imprisonment

Step 1. Compute the three variables **PROVOKE, SELFDEF**, and **INSANITY**

1. In order to compute the three variables **PROVOKE, SELFDEF**, and **INSANITY**, click **Transform** and then **Compute** from the menu bar. The following **Compute Variable** window will open.



2. To compute the first variable **PROVOKE**, type **provoke** in the **Target Variable**: field. As this variable is computed as the mean of the two variables **PROVO** and **PASSION**, type **mean(provo, passion)** in the **Numeric Expression**: field. Click OK to compute the **PROVOKE** variable.



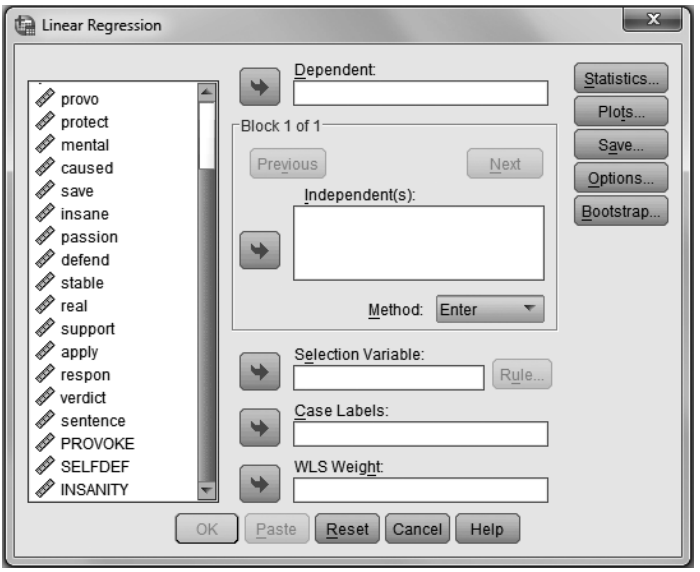
3. Repeat step 2 above to compute the **SELFDEF** and **INSANITY** variables of **SELFDEF** and **INSANITY**.

14.6.3 Testing Assumptions

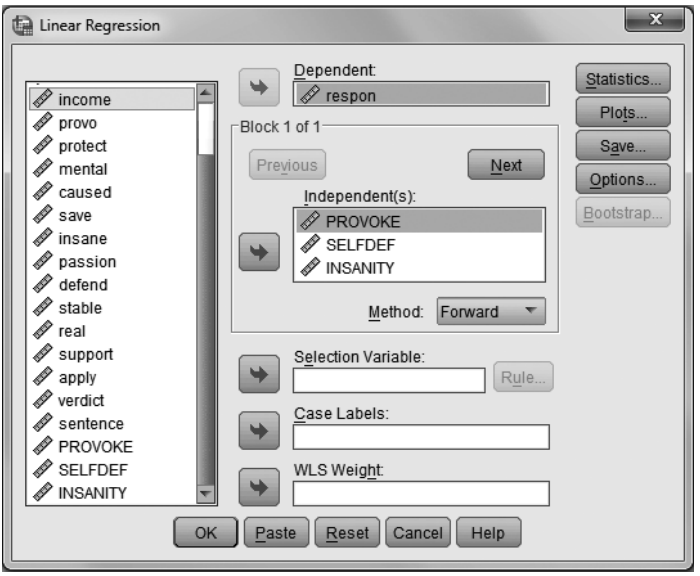
The assumptions of linearity, homoscedasticity, normality, and independence of error terms will be assessed through the regression analysis.

14.6.4 *Windows Method*: Multiple Regression—Predicting the Level of Responsibility from the Three Defense Strategies PROVOKE, SELFDEF, and INSANITY

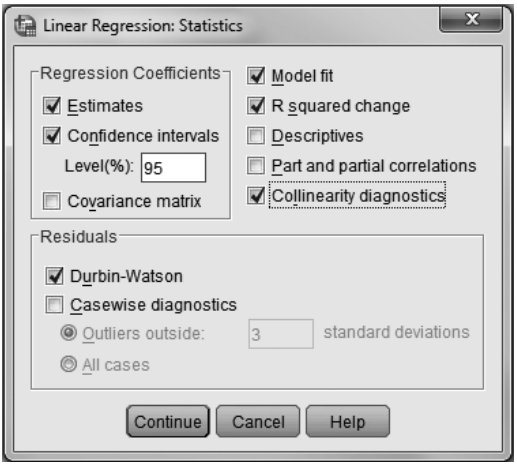
1. In order to predict the level of responsibility attributed to the battered woman's action from the three defense strategies **PROVOKE**, **SELFDEF**, and **INSANITY**, click **Analyze**, then **Regression**, and then **Linear…** from the menu bar. The following **Linear Regression** window will open. Notice that the list of variables now contains the newly created variables **PROVOKE**, **SELFDEF**, and **INSANITY**.



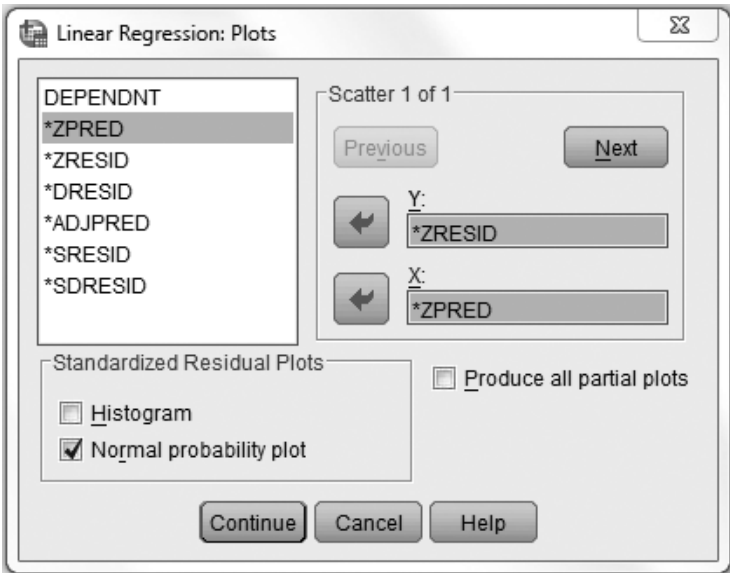
2. Click (highlight) the **RESPON** variable and then click to transfer this variable to the **Dependent:** field. Next, click (highlight) the three newly created variables **PROVOKE, SELFDEF**, and **INSANITY** and then click to transfer these variables to the **Independent(s):** field. In the **Method:** field, select **Forward** from the drop-down list as the method of entry for the three independent (predictor) variables into the prediction equation.



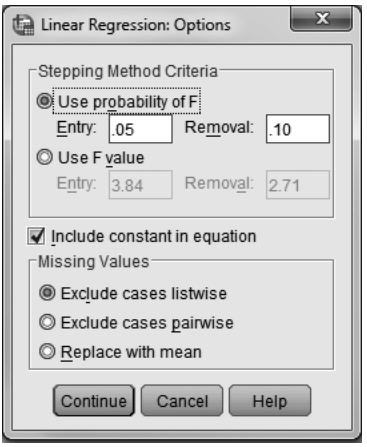
3. Click Statistics... to open the **Linear Regression: Statistics** window. Check the fields to obtain the statistics required. For this case, check the fields for **Estimates, Confidence intervals, Model fit, R squared change, Collinearity diagnostics,** and **Durbin-Watson**. Click Continue when finished.



4. When the **Linear Regression** window opens, click Plots... to open the **Linear Regression: Plots** window below. Click ***ZRESID** and then click ▶ to transfer this item to the **Y** field. Click ***ZPRED** and then click ▶ to transfer this item to the **X** field. Check the **Normal probability plot** field. Click Continue to return to the **Linear Regression** window.

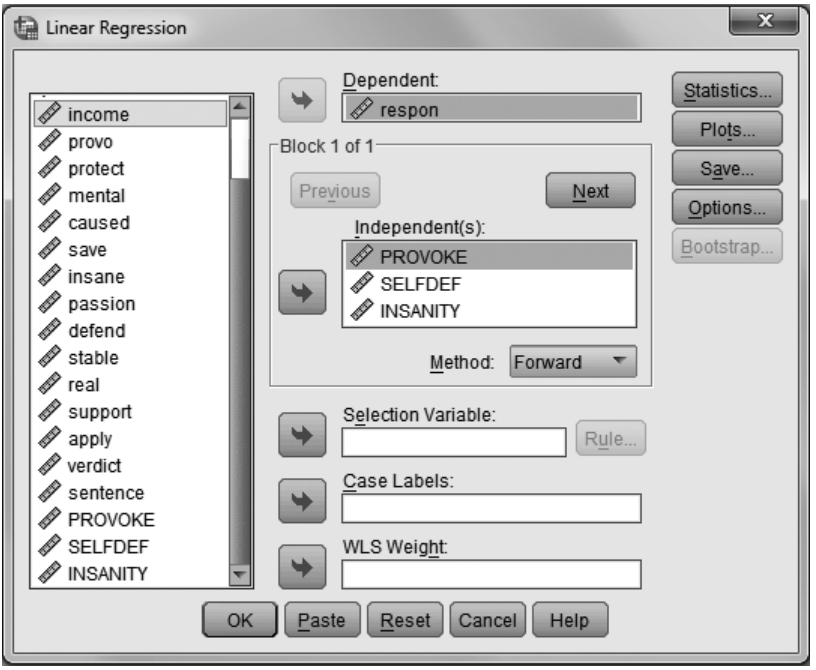


5. When the **Linear Regression** window opens, click Options... to open the **Linear Regression: Options** window below. Ensure that both the **Use probability of F** and the **Include constant in equation** fields are checked. Under **Missing Values**, check the **Exclude cases listwise** field.



Click Continue to return to the **Linear Regression** window.

6. When the **Linear Regression** window opens, click OK to complete the analysis. See Table 14.1 for the results.



14.6.5 SPSS Syntax Method

```
COMPUTE PROVOKE = MEAN(PROVO,PASSION).
COMPUTE SELFDEF = MEAN(PROTECT,SAVE,DEFEND).
COMPUTE INSANITY = MEAN(MENTAL,INSANE,STABLE).

REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS CI(95) R ANOVA COLLIN
  TOL CHANGE
/CRITERIA = PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT RESPON
/METHOD = FORWARD PROVOKE SELFDEF INSANITY
/SCATTERPLOT = (*ZRESID,*ZPRED)
/RESIDUALS NORMPROB(ZRESID).
```

14.6.6 SPSS Output

TABLE 14.1

Multiple Regression Analysis Output

Regression

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	SELFDEF	.	Forward (Criterion: Probability-of-F-to-enter <= 050)
2	INSANITY	.	Forward (Criterion: Probability-of-F-to-enter <= 050)
3	PROVOKE	.	Forward (Criterion: Probability-of-F-to-enter <= 050)

[a] Dependent Variable: respon

Model Summary[d]

					Change Statistics					
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change	Durbin-Watson
1	.417[a]	.174	.172	1.678	.174	81.920	1	389	.000	
2	.431[b]	.186	.182	1.668	.012	5.740	1	388	.017	
3	.446[c]	.198	.192	1.658	.012	6.032	1	387	.014	2.052

[a] Predictors: (Constant), SELFDEF
[b] Predictors: (Constant), SELFDEF, INSANITY
[c] Predictors: (Constant), SELFDEF, INSANITY, PROVOKE
[d] Dependent Variable: respon

TABLE 14.1 (*Continued*)
Multiple Regression Analysis Output

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	230.785	1	230.785	81.920	.000[b]
	Residual	1095.891	389	2.817		
	Total	1326.675	390			
2	Regression	246.761	2	123.380	44.329	.000[c]
	Residual	1079.915	388	2.783		
	Total	1326.675	390			
3	Regression	263.334	3	87.778	31.946	.000[d]
	Residual	1063.342	387	2.748		
	Total	1326.675	390			

[a] Dependent Variable: respon
[b] Predictors: (Constant), SELFDEF
[c] Predictors: (Constant), SELFDEF, INSANITY
[d] Predictors: (Constant), SELFDEF, INSANITY, PROVOKE

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for B		Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	5.970	.267		22.394	.000	5.446	6.495		
	SELFDEF	−.443	.049	−.417	−9.051	.000	−.539	−.346	1.000	1.000
2	(Constant)	5.441	.345		15.769	.000	4.763	6.119	.990	1.011
	SELFDEF	−.431	.049	−.406	−8.813	.000	−.527	−.335	.990	1.011
	INSANITY	.129	.054	.110	2.396	.017	.023	.234		
3	(Constant)	5.905	.391		15.086	.000	5.135	6.674	.956	1.046
	SELFDEF	−.408	.049	−.385	−8.262	.000	−.505	−.311	.953	1.050
	INSANITY	.155	.054	.132	2.839	.005	048	.262	.937	1.067
	PROVOKE	−.129	.053	−.115	−2.456	.014	−.232	−.026		

[a] Dependent Variable: respon

Excluded Variables[a]

						Collinearity Statistics		
Model		Beta In	t	Sig.	Partial Correlation	Tolerance	VIF	Minimum Tolerance
1	PROVOKE	−.090[b]	−1.928	.055	−.097	.973	1.028	.973
	INSANITY	.110[b]	2.396	.017	.121	.990	1.011	.990
2	PROVOKE	−.115[c]	−2.456	.014	−.124	.937	1.067	.937

[a] Dependent Variable: respon
[b] Predictors in the Model: (Constant), SELFDEF
[c] Predictors in the Model: (Constant), SELFDEF, INSANITY

(*Continued*)

TABLE 14.1 (*Continued*)
Multiple Regression Analysis Output

Collinearity Diagnostics[a]

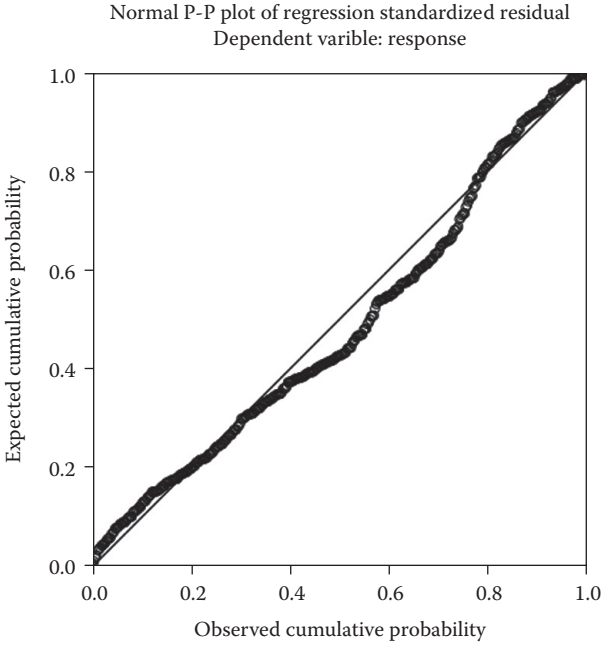
Model Dimension	Eigenvalue	Condition Index	Variance Proportions (Constant)	SELFDEF	INSANITY	PROVOKE
1 1	1.948	1.000	.03	.03		
2	.052	6.118	.97	.97		
2 1	2.815	1.000	.01	.01	.02	
2	.146	4.384	.01	.27	.63	
3	.036	8.534	.99	.72	.35	
3 1	3.746	1.000	.00	.01	.01	.01
2	.148	5.032	.00	.21	.67	.01
3	.073	7.185	.00	.39	.15	.75
4	.034	10.545	.99	.39	.18	.24

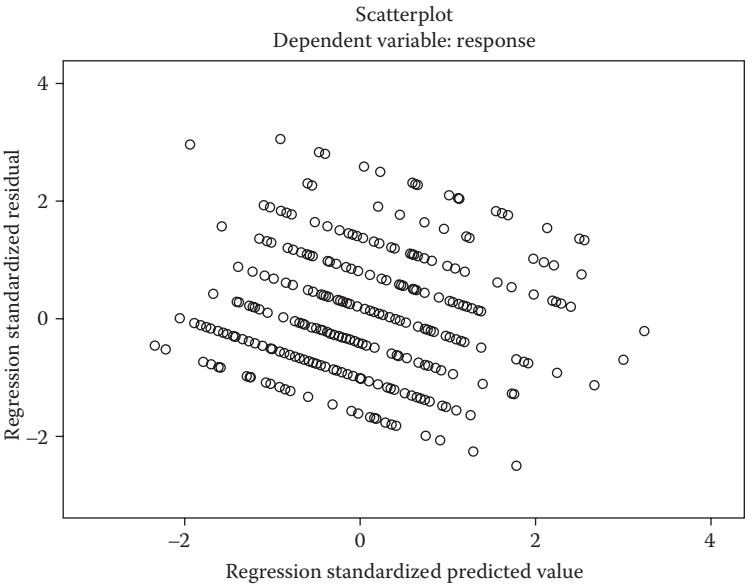
[a] Dependent Variable: respon

Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	1.76	6.35	3.68	.822	391
Residual	−4.150	5.062	.000	1.651	391
Std. Predicted Value	−2.336	3.241	.000	1.000	391
Std. Residual	−2.504	3.054	.000	.996	391

[a] Dependent Variable: respon





14.6.7 Results and Interpretation

14.6.7.1 Testing Assumptions

The scatterplot of **standardized residuals** against the **standardized predicted values** shows no clear relationship pattern between the residuals and the predicted values. This is consistent with the assumptions of linearity and homoscedasticity.

From the **normal P–P plot of standardized residuals,** it can be seen that the plot of the residuals for the dependent variable fits the expected pattern well enough to indicate a relatively normal distribution.

The **Durbin-Watson statistic** is used to test the assumption of independence of error terms (residuals). As the Durbin-Watson $d = 2.052$ is between the two critical values of $1.5 < d < 2.5$, it can be assumed that there is independence of residuals.

14.6.7.2 Prediction Equation (Predicting the Level of Responsibility from the Three Defense Strategies PROVOKE, SELFDEF, and INSANITY)

The prediction equation is $Y' = A + B_1X_1 + B_2X_2 + \ldots B_nX_n$ where:

Y' = The predicted dependent variable

A = Constant

B = Unstandardized regression coefficient

X = Value of the predictor variable

The relevant information for calculating the predicted responsibility attribution is presented in the **Coefficients** table (see Table 14.1). An examination of this table shows that all three predictor variables were entered into the prediction equation (Model 3), indicating that the defense strategies of self-defense, insanity, and provocation are significant predictors of the level of responsibility attributed to the battered woman for her fatal action. In order to predict the level of responsibility attributed from these three defense strategies, use the values presented in the **Unstandardized Coefficients** column for Model 3. Using the **Constant** and **B** (unstandardized coefficient) values, the prediction equation would be:

$$\text{Predicted responsibility attribution} = 5.91 + (-0.41 \times \text{SELFDEF}) + (0.16 \times \text{INSANITY}) + (-0.13 \times \text{PROVOKE})$$

Thus, for a subject who strongly believes that the battered woman's action was motivated by self-defense and provocation (a score of 8 on both scales) and not by an impaired mental state (a score of 1 on the insanity scale), the predicted degree of responsibility attribution would be:

$$\begin{aligned}\text{Predicted responsibility attribution} &= 5.91 + (-0.41 \times 8) + (0.16 \times 1) + (-0.13 \times 8) \\ &= 5.91 - 3.28 + 0.16 - 1.04 \\ &= 1.75\end{aligned}$$

Given that responsibility attribution is measured on an 8-point scale with 1 = not at all responsible, to 8 = entirely responsible, a predicted value of 1.75 would suggest that this subject would attribute little responsibility to the battered woman for her fatal action.

14.6.7.3 Evaluating the Strength of the Prediction Equation

A measure of the strength of the computed prediction equation is ***R*-square**, sometimes called the **coefficient of determination**. In the regression model, *R*-square is the square of the correlation coefficient between *Y*, the observed value of the dependent variable, and *Y′*, the predicted value of *Y* from the fitted regression line. Hence, if for each subject, the researcher computes the predicted responsibility attribution, and then calculates the square of the correlation coefficient between predicted responsibility attribution and observed responsibility attribution values, *R*-square is obtained. If all the observations fall on the regression line, *R*-square is 1 (perfect linear relationship). An *R*-square of 0 indicates no linear relationship between the predictor and dependent variables. To test the hypothesis of no linear relationship between the predictor and dependent variables, that is, *R*-square = 0, the analysis of variance (ANOVA) is utilized. In this case, the results from this test are shown in the ANOVA table (see Table 14.1). The *F* value serves to test how well the regression model (Model 3) fits the data.

If the probability associated with the *F* statistics is small, the hypothesis that *R*-square = 0 is rejected. For this example, the computed *F* statistic is 31.95, with an observed significance level of less than 0.001. Thus, the hypothesis that there is no linear relationship between the predictor and dependent variables is rejected.

14.6.7.4 Identifying Multicollinearity

When the predictor variables are correlated among themselves, the unique contribution of each predictor variable is difficult to assess. This is because of the overlapped or shared variance between the predictor variables, that is, they are multicollinear. For this example, both the "tolerance" values (greater than 0.10) and the "VIF" values (less than 10) are all quite acceptable (see the **Coefficients** table). Thus, multicollinearity does not appear to be a problem in this case.

Another way of assessing if there is too much multicollinearity in the model is to look at the **Collinearity Diagnostics** table. The *condition index* summarizes the findings, and a common rule of thumb is that a condition index over 15 indicates a possible multicollinearity problem and a condition index over 30 suggests a serious multicollinearity problem. For this example, multicollinearity is not a problem.

14.6.7.5 Identifying Independent Relationships

Once it has been established that multicollinearity is not a problem, multiple regression can be used to assess the relative contribution (independent relationship) of each predictor variable by *controlling the effects of other predictor variables* in the prediction equation. The procedure can also be used to measure the size and direction of the obtained independent relationships.

In identifying independent relationships, it is inappropriate to interpret the unstandardized regression coefficients (**B** values in the **Coefficients** table) as indicators of the relative importance of the predictor variables. This is because the **B** values are based on the actual units of measurement, which may differ from variable to variable, that is, one variable may be measured on a five-point scale, while another variable may be measured on an eight-point scale. When variables differ substantially in units of measurement, the sheer magnitude of their coefficients does not reveal anything about their relative importance. Likewise, the unstandardized regression coefficients incorporate the standard deviations of the predictor variables that may differ substantially, and these differences will impact differently on the magnitude of the regression coefficients. One way to make regression coefficients somewhat more comparable is to calculate **Beta** weights, which are the coefficients of the predictor variables when all variables are expressed in standardized form (*Z*-score).

In this example, the **Beta** weights (standardized regression coefficients) for all three defense strategies are presented in the **Coefficients** table. The

size of the Beta weights indicates the strength of their independent relationships. From the table, it can be determined that self-defense has the strongest relationship with responsibility attribution, whereas the other two defense strategies—provocation and insanity—are weaker. The direction of the coefficients also sheds light on the nature of the relationships. Thus, for the defense strategies of self-defense and provocation, the negative coefficients indicate that the more the subjects interpreted the battered woman's motive for her fatal action as owing to self-defense and provocation, the less they held her responsible for her action (self-defense: $\beta = -0.39$, $t = 8.26$, $p < .001$; provocation: $\beta = -0.12$, $t = -2.46$, $p < .05$). Conversely, the positive coefficient associated with the insanity variable shows that the more they interpreted the battered woman's action as owing to an impaired mental state, the more they held the woman responsible for her fatal action, $\beta = 0.13$, $t = 2.84$, $p < .01$.

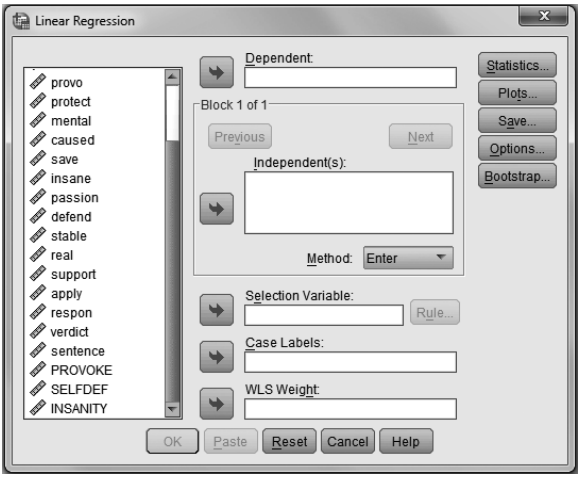
14.7 Example 2: Hierarchical Regression

Another means of measuring the relative importance of predictor variables is to consider the increase in *R*-square when a variable is entered into an equation that already contains the other predictor variables. The increase in *R*-square resulting from the entry of a subsequent predictor variable indicates the amount of unique information in the dependent variable that is accounted for by that variable, *above and beyond* what has already been accounted for by the other predictor variables in the equation.

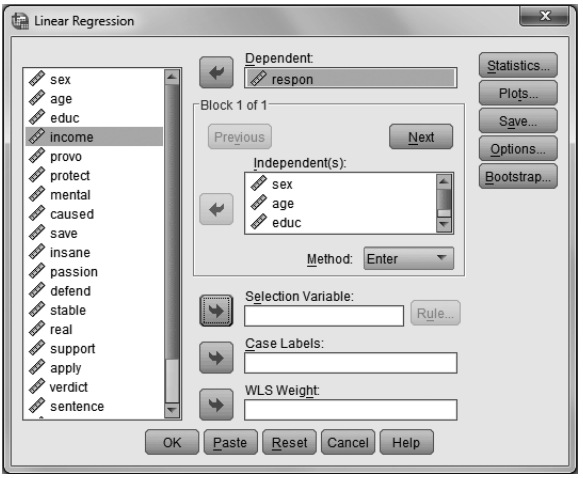
Suppose that the researcher is interested in comparing the relative importance of two sets of variables in predicting responsibility attributed to the battered woman in the previous example. Specifically, the researcher wants to assess the relative importance of a set of variables comprising the three defense strategies (self-defense, provocation, and insanity) and a set of variables comprising the subjects' demographic characteristics (gender, age, educational level, income) in predicting responsibility attribution. This task can be accomplished by the use of the **hierarchical regression** procedure. For this example, the researcher determines the order of entry for the two sets of predictor variables. Assume that the researcher believes that the subjects' demographics would be less strongly related to the dependent variable than the set of defense strategies. On the basis of this assumption, the researcher accords priority of entry into the prediction equation to the set of demographic variables, followed by the set of defense strategies. This order of entry will assess (1) the importance of the demographic variables in predicting the dependent variable of responsibility attribution, and (2) the amount of unique information in the dependent variable that is accounted for by the three defense strategies. This example employs the data set **DOMES.SAV**.

14.7.1 Windows Method

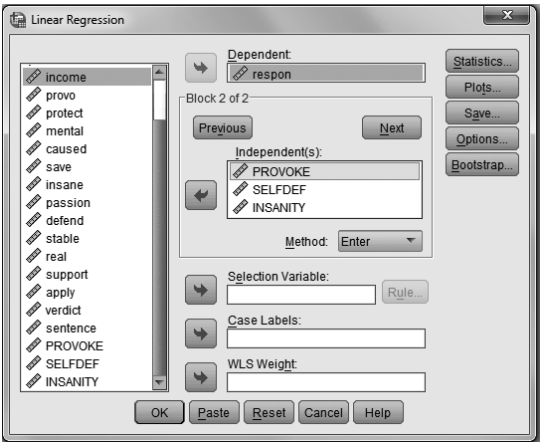
1. From the menu bar, click **Analyze**, then **Regression**, and then **Linear…**. The following **Linear Regression** window will open.



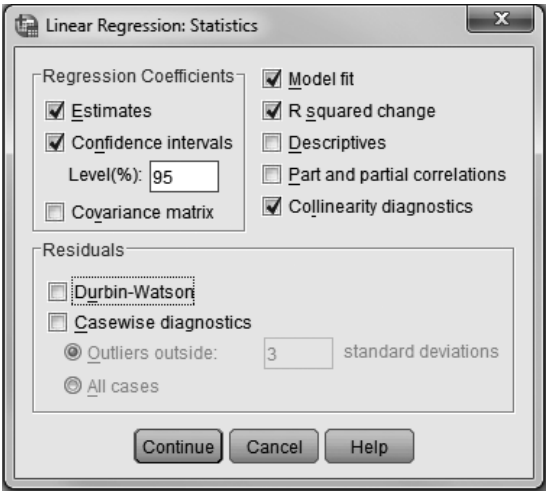
2. Click (highlight) the **RESPON** variable and then click ▶ to transfer this variable to the **Dependent:** field. Since the set of demographic variables (**SEX, AGE, EDUCATIONAL LEVEL, INCOME**) will be entered first into the prediction equation (**Block 1**), click (highlight) these variables and then click ▶ to transfer these variables to the **Independent(s):** field. In the **Method:** cell, select **Enter** from the drop-down list as the method of entry for this set of demographic variables into the prediction equation.



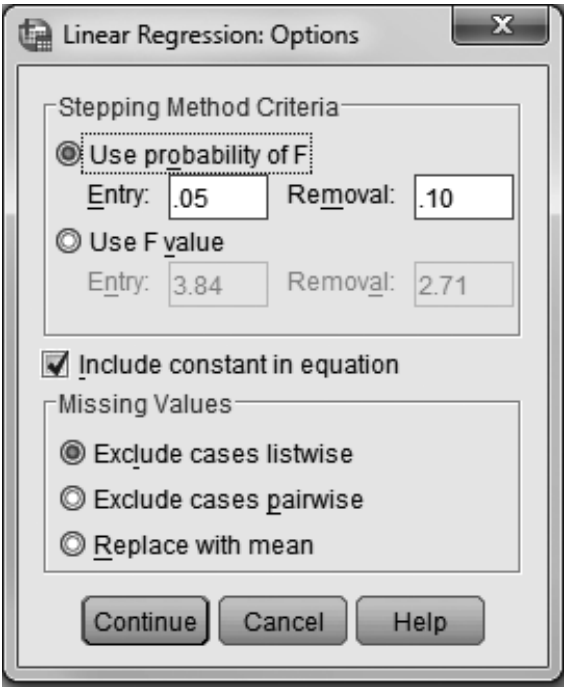
Next, click Next to open **Block 2** in the **Independent(s):** field for entry of the second set of independent variables. Click (highlight) the variables **PROVOKE**, **SELFDEF**, and **INSANITY** and then click ▶ to transfer these variables to the **Independent(s):** field. In the **Method:** field, select **Enter** from the drop-down list as the method of entry for this set of independent (predictor) variables into the prediction equation.



3. Click Statistics... to open the **Linear Regression: Statistics** window. Check the fields to obtain the statistics required. For this example, check the fields for **Estimates, Confidence intervals, Model fit, R squared change**, and **Collinearity diagnostics**. Click Continue when finished.



4. When the **Linear Regression** window opens, click Options... to open the **Linear Regression: Options** window below. Ensure that both the **Use probability of F** and the **Include constant in equation** fields are checked. Under **Missing Values**, check the **Exclude cases listwise** field. Click Continue to return to the **Linear Regression** window.



5. When the **Linear Regression** window opens, click OK to complete the analysis. See Table 14.2 for the results.

14.7.2 SPSS Syntax Method

```
REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS CI(95) R ANOVA COLLIN TOL CHANGE
/CRITERIA = PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT RESPON
/METHOD = ENTER SEX AGE EDUC INCOME/ENTER PROVOKE
INSANITY SELFDEF.
```

14.7.3 SPSS Output

TABLE 14.2

Hierarchical Multiple Regression Output

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	INCOME, EDUC, SEX, AGE[a]	.	Enter
2	PROVOKE, INSANITY, SELFDEF[a]	.	Enter

[a] All requested variables entered.
[b] Dependent Variable: RESPON

Model Summary

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	.241[a]	.058	.048	1.80	.058	5.727	4	372	.000
2	.469[b]	.220	.205	1.65	.162	25.481	3	369	.000

[a] Predictors: (Constant), INCOME, EDUC, SEX, AGE
[b] Predictors: (Constant), INCOME, EDUC, SEX, AGE, PROVOKE, INSANITY, SELFDEF

AVOVA[c]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	74.280	4	18.570	5.727	.000[a]
	Residual	1206.261	372	3.243		
	Total	1280.541	376			
2	Regression	281.286	7	40.184	14.839	.000[b]
	Residual	999.255	369	2.708		
	Total	1280.541	376			

[a] Predictors: (Constant), INCOME, EDUC, SEX, AGE
[b] Predictors: (Constant), INCOME, EDUC, SEX, AGE, PROVOKE, INSANITY, SELFDEF
[c] Dependent Variable: RESPON

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			95% Confidence Interval for B		Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	4.458	.658		6.770	.000	3.163	5.753		
	sex	−.763	.200	−.192	−3.811	.000	−1.157	−.369	.994	1.006
	age	−.019	.013	−.086	−1.511	.132	−.044	.006	.784	1.276
	educ	.126	.086	.076	1.470	.142	−.042	.294	.958	1.044
	income	.221	.139	.089	1.590	.113	−.052	.494	.802	1.246

TABLE 14.2 (*Continued*)

Hierarchical Multiple Regression Output

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			95% Confidence Interval for B		Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound	Tolerance	VIF
2	(Constant)	6.633	.719		9.220	.000	5.219	8.048		
	sex	−.521	.185	−.131	−2.810	.005	−.885	−.156	.969	1.032
	age	−.019	.012	−.086	−1.652	.099	−.042	.004	.783	1.277
	educ	.069	.078	.041	.876	.381	−.086	.223	.950	1.053
	income	107	.128	.043	.835	.404	−.145	.359	.791	1.265
	PROVOKE	−.121	.053	−.109	−2.295	.022	−.224	−.017	931	1.074
	SELFDEF	−.382	.050	−.361	−7.591	.000	−.481	−.283	.936	1.068
	INSANITY	.127	.056	.108	2.290	.023	.018	.236	.943	1.061

[a] Dependent Variable: respon

Excluded Variables[b]

Model		Beta In	F	Sig.	Partial Correlation	Collinearity Statistics Tolerance
1	PROVOKE	−.145[a]	8.434	.004	−.149	.992
	SELFDEF	−.385[a]	66.421	.000	−.390	.967
	INSANITY	.110[a]	4.705	.031	.112	.983

[a] Predictors in the Model: (Constant), INCOME, EDUC, SEX, AGE

[b] Dependent Variable: RESPON

14.7.4 Results and Interpretation

In the **Model Summary** table (see Table 14.2), Model 1 represents entry of the first set of demographic variables, and Model 2 represents entry of the second set of self-defense strategy variables. The results show that Model 1 (demographics) accounted for 5.8% of the variance (***R* Square**) in the subjects' responsibility attribution. Entry of the three defense strategy variables (Model 2) resulted in an ***R* Square Change** of 0.162. This means that entry of the three defense strategy variables increased the explained variance in the subjects' responsibility attribution by 16.2% to a total of 22%. This increase is significant by the **F Change** test, $F(3,369) = 25.48$, $p < .001$ (a test for the increase in explanatory power). These results suggest that the defense strategy variables represent a significantly more powerful set of predictors than the set of demographic variables.

In the **ANOVA** table, the results show that entry of the set of demographic variables alone (Model 1) yielded a significant prediction equation, $F(4,372) = 5.73$, $p < .001$. The addition of the three defense strategy variables (Model 2) resulted in an overall significant prediction equation, $F(7,369) = 14.84$, $p < .001$.

Collinearity Diagnostics[a]

Model Dimension	Eigenvalue	Condition Index	Variance Proportions							
			(Constant)	sex	age	educ	income	PROVOKE	SELFDEF	INSANITY
1 1	4.634	1.000	.00	.00	.00	.00	.01			
2	.208	4.722	.01	.05	.01	.02	.65			
3	.084	7.429	.00	.00	.75	.09	.26			
4	.060	8.786	.01	.70	.05	.23	.08			
5	.014	18.109	.98	.25	.18	.66	.00			
2 1	7.284	1.000	.00	.00	.00	.00	.00	.00	.00	.00
2	.252	5.380	.00	.01	.03	.00	.55	.02	.03	.01
3	.157	6.805	.00	.03	.02	.00	.00	.00	.09	.72
4	.084	9.296	.00	.03	.63	.01	.39	.05	.16	.01
5	.080	9.542	.00	.20	.12	.12	.00	.37	.08	.00
6	.074	9.908	.00	.01	.01	.03	.00	.46	.57	.18
7	.057	11.304	.01	.56	.04	.32	.06	.02	.00	.05
8	.011	25.521	.99	.16	.15	.52	.00	.08	.08	.04

[a] Dependent Variable: respon

Looking at Model 2 in the **Coefficients** table, it can be seen that multicollinearity is not a problem—all tolerance values are above 0.10, all VIF values are below 10, and the condition indices for the seven predictor variables are below 15.

In examining the **Beta** weights (standardized regression coefficients), it can also be seen that all three defense strategy variables are significant predictors of responsibility attribution ($p < .05$), whereas subjects' sex is the only demographic variable that was found to be significant. Thus, the more the subjects perceived the battered woman's fatal action owed to provocation, $\beta = -0.11$, $t = -2.30$, $p < .05$, and to self-defense, $\beta = -0.36$, $t = -7.59$, $p < .001$, the less responsibility they attributed to the woman for her action. Conversely, the more the subjects perceived the battered woman's fatal action owed to insanity, the greater the responsibility they attributed to the woman for her action $\beta = 0.11$, $t = 2.29$, $p < .05$. The finding that subjects' sex was a significant predictor, $\beta = -0.13$, $t = -2.81$, $p < .01$, indicated that males attributed greater responsibility to the woman for her fatal action than females did (code: 1 = male, 2 = female).

14.8 Example 3: Path Analysis

In path analysis, multiple regression is often used in conjunction with a causal theory, with the aim of describing the entire structure of linkages between independent and dependent variables posited from that theory. For example, on the basis of theoretical considerations of the domestic violence example, a researcher has constructed the path model presented in Figure 14.1 to represent the hypothesized structural relationships between the three defense strategies of provocation, self-defense, and insanity, and the attribution of responsibility.

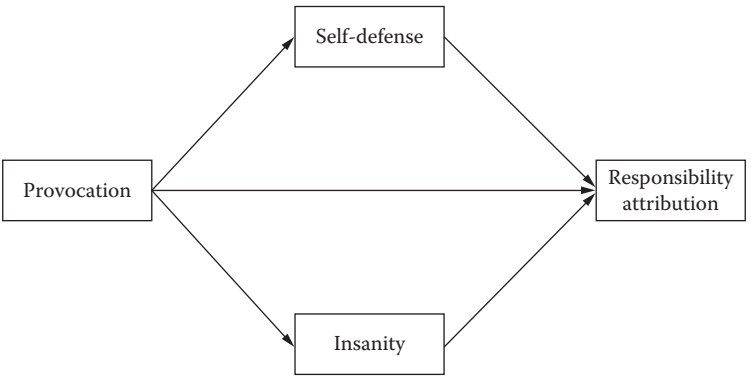
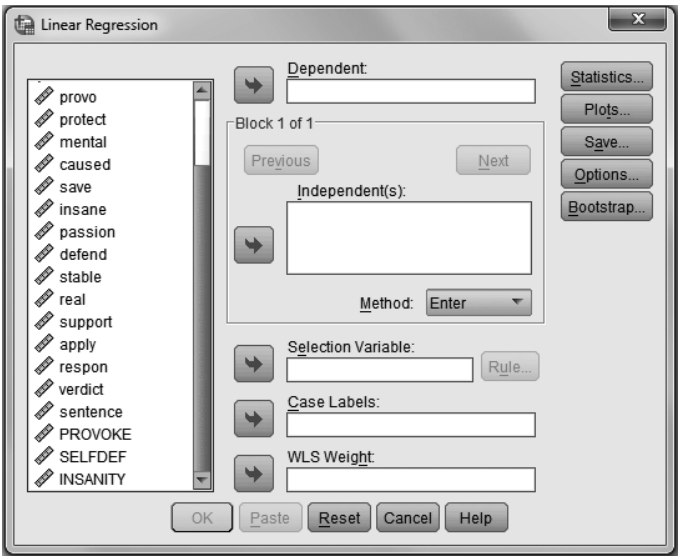


FIGURE 14.1
Path model representing the hypothesized direct and indirect relationships between the defense strategy of provocation on responsibility attribution through the mediators of self-defense and insanity.

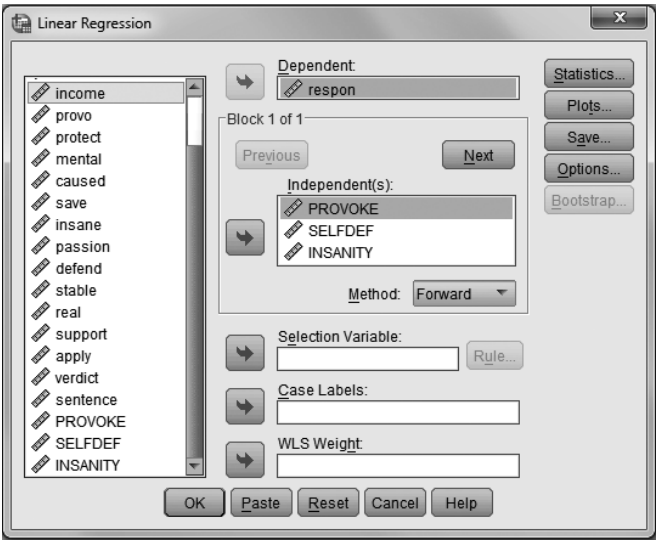
The model specifies an "ordering" among the variables that reflects a hypothesized structure of cause-effect linkages. Multiple regression techniques can be used to determine the magnitude of *direct* and *indirect* influences that each variable has on other variables that follow it in the presumed causal order (as indicated by the directional arrows). Each arrow in the model represents a presumed causal linkage or path of causal influence. Through regression techniques, the strength of each separate path can be estimated. This analysis actually involves three regression equations since (1) *responsibility attribution* is a dependent variable for the three defense variables, and (2) the *self-defense* and *insanity* variables are dependent variables for the defense variable of provocation. This model uses the data set: **DOMES.SAV**.

14.8.1 Windows Method

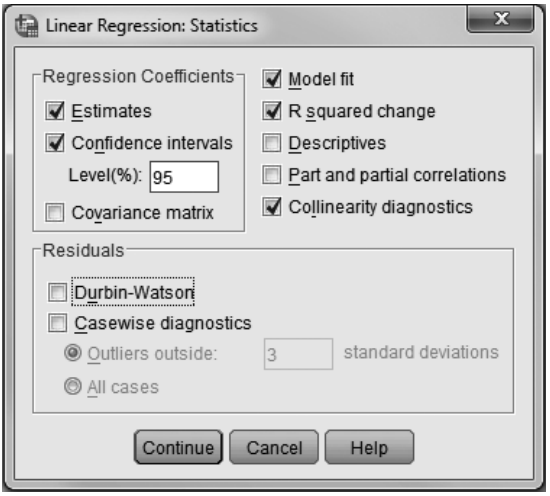
1. The first regression equation involves predicting the level of responsibility (**RESPON**) attributed to the battered woman's action from all three defense strategies (**PROVOKE, SELFDEF, INSANITY**). Click **Analyze**, then **Regression**, and then **Linear…** from the menu bar. The following **Linear Regression** window will open.



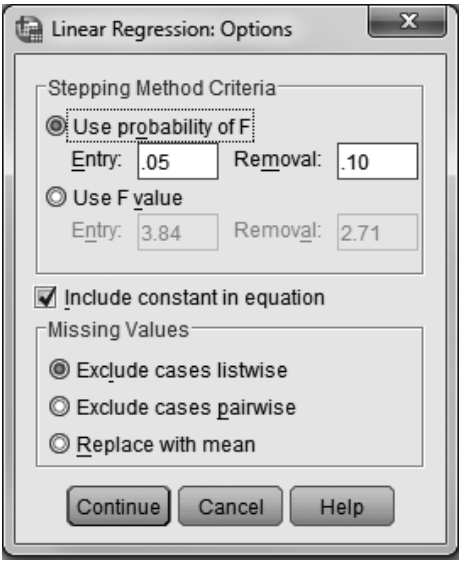
2. Click (highlight) the **RESPON** variable and then click ▶ to transfer this variable to the **Dependent:** field. Next, click (highlight) the **PROVOKE**, **SELFDEF**, and **INSANITY** variables and then click ▶ to transfer these variables to the **Independent(s):** field. In the **Method:** field, select **Forward** from the drop-down list as the method of entry for the three independent (predictor) variables into the prediction equation.



3. Click Statistics... to open the **Linear Regression: Statistics** window. Check the fields to obtain the statistics required. For this example, check the fields for **Estimates, Confidence intervals, Model fit, R squared change**, and **Collinearity diagnostics**. Click Continue when finished.

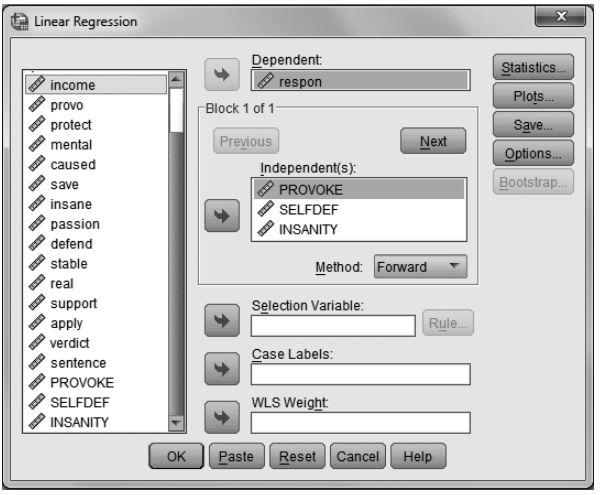


4. When the **Linear Regression** window opens, click Options... to open the **Linear Regression: Options** window below. Ensure that the **Use probability of F**, **Include constant in equation**, and **Exclude cases listwise** fields are checked.

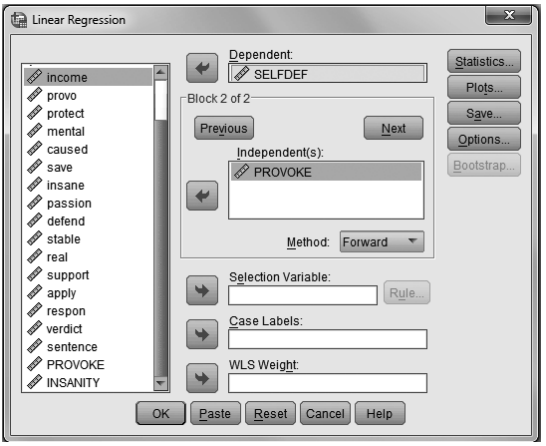


Click Continue to return to the **Linear Regression** window.

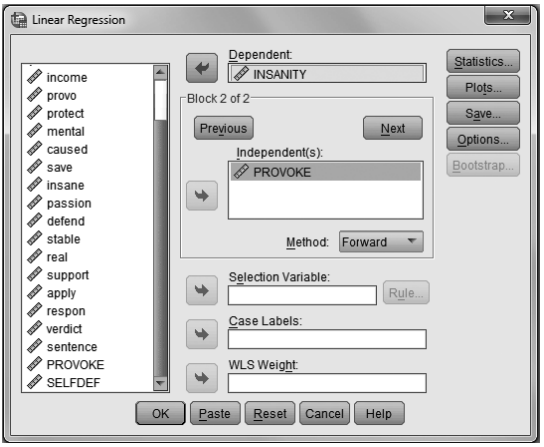
5. When the **Linear Regression** window opens, click OK to complete the analysis.



6. The second prediction equation involves predicting the defense strategy **SELFDEF** from the defense strategy **PROVOKE**. To do this, repeat step 1 to step 5. In step 1, transfer **SELFDEF** to the **Dependent:** field, and the **PROVOKE** variable to the **Independent(s):** field. Click OK to complete the analysis.



7. The third prediction equation involves predicting the defense strategy **INSANITY** from the defense strategy **PROVOKE**. To do this, repeat step 1 to step 5. In step 1, transfer **INSANITY** to the **Dependent:** field, and the **PROVOKE** variable to the **Independent(s):** field. Click OK to complete the analysis. See Table 14.3 for the results.



14.8.2 SPSS Syntax Method

```
REGRESSION VARIABLES = (COLLECT)
  /STATISTICS = DEFAULTS CHA TOL CI COLLIN
  /DEPENDENT = RESPON
  /FORWARD PROVOKE SELFDEF INSANITY.
REGRESSION VARIABLES = (COLLECT)
  /STATISTICS = DEFAULTS CHA TOL CI COLLIN
  /DEPENDENT = SELFDEF INSANITY
  /FORWARD PROVOKE.
```

14.8.3 SPSS Output

TABLE 14.3

Path Analysis Output

Regression

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	SELFDEF	.	Forward (Criterion: Probability-of-F-to-enter <= .050)
2	INSANITY	.	Forward (Criterion: Probability-of-F-to-enter <= .050)
3	PROVOKE	.	Forward (Criterion: Probability-of-F-to-enter <= .050)

[a] Dependent Variable: respon

Model Summary

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	.417[a]	.174	.172	1.678	.174	81.920	1	389	.000
2	.431[b]	.186	.182	1.668	.012	5.740	1	369	.017
3	.446[c]	.198	.192	1.658	.012	6.032	1	387	.014

[a] Predictors: (Constant), SELFDEF

[b] Predictors: (Constant), SELFDEF, INSANITY

[c] Predictors: (Constant), SELFDEF, INSANITY, PROVOKE

AVOVA[d]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	230.785	1	230.785	81.920	.000[a]
	Residual	1095.891	389	2.817		
	Total	1326.675	390			
2	Regression	246.761	2	123.380	44.329	.000[b]
	Residual	1079.915	388	2.783		
	Total	1326.675	390			
3	Regression	263.334	3	87.778	31. 946	.000[c]
	Residual	1063.342	387	2.748		
	Total	1326.675	390			

[a] Predictors: (Constant), SELFDEF

[b] Predictors: (Constant), SELFDEF, INSANITY

[c] Predictors: (Constant), SELFDEF, INSANITY, PROVOKE

[d] Dependent Variable: respon

TABLE 14.3 (*Continued*)

Path Analysis Output

Coefficients[a]

Model	Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	95% Confidence Interval for B Lower Bound	95% Confidence Interval for B Upper Bound	Collinearity Statics Tolerance	Collinearity Statics VIF
1 (Constant)	5.970	.267		22.394	.000	5.446	6.495		
SELFDEF	−.443	.049	−.417	−9.051	.000	−.539	−.346	1.000	1.000
2 (Constant)	5.441	.345		15.769	.000	4.763	6.119		
SELFDEF	−.431	.049	−.406	−8.813	.000	−.527	−.335	.990	1.011
INSANITY	.129	.054	.110	2.396	.017	.023	.234	.990	1.011
3 (Constant)	5.905	.391		15.086	.000	5.135	6.674		
SELFDEF	−.408	.049	−.385	−8.262	.000	−.505	−.311	.956	1.046
INSANITY	.155	.054	.132	2.839	.005	.048	.262	.953	1.050
PROVOKE	−.129	.053	−.115	−2.456	.014	−.232	−.026	.937	1.067

[a] Dependent Variable: respon

Excluded Variables[c]

Model	Beta In	t	Sig.	Partial Correlation	Collinearity Statistics Tolerance	Collinearity Statistics VIF	Collinearity Statistics Minimum Tolerance
1 PROVOKE	−.090[a]	−1.928	.055	−.097	.973	1.028	.973
INSANITY	.110[a]	2.396	.017	.121	.990	1.011	.990
2 PROVOKE	−.115[b]	−2.456	.014	−.124	.937	1.067	.937

[a] Predictors in the Model: (Constant), SELFDEF
[b] Predictors in the Model: (Constant), SELFDEF, INSANITY
[c] Dependent Variable: respon

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions (Constant)	SELFDEF	INSANITY	PROVOKE
1	1	1.948	1.000	.03	.03		
	2	.052	6.118	.97	.97		
3	1	3.746	1.000	.00	.01	.01	.01
	2	.148	5.032	.00	.21	.67	.01
	3	.073	7.185	.00	.39	.15	.75
	4	.034	10.545	.99	.39	.18	.24

[a] Dependent Variable: respon

Regression

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	PROVOKE	.	Forward (Criterion: Probability-of-F-to-enter <= . 050)

[a] Dependent Variable: SELFDEF

(*Continued*)

TABLE 14.3 (*Continued*)

Path Analysis Output

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.158[a]	.025	.023	1.72044	.025	10.117	1	394	.002

[a] Predictors: (Constant), PROVOKE

AVOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	29.946	1	29.946	10.117	.002[a]
	Residual	1166.211	394	2.960		
	Total	1196.157	395			

[a] Predictors: (Constant), PROVOKE
[b] Dependent Variable: SELFDEF

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			95.0% Confidence Interval for B		Collinearity Statics	
		B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	4.310	.286		15.084	.000	3.748	4.872	1.000	1.000
	PROVOKE	.166	.052	.158	3.181	.002	.064	.269		

[a] Dependent Variable: SELFDEF

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions	
				(Constant)	SELFDEF
1	1	1.953	1.000	.02	.02
	2	.047	6.455	.98	.98

[a] Dependent Variable: SELFDEF

Regression

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	PROVOKE	.	Forward (Criterion: Probability-of-F-to-enter <= . 050)

[a] Dependent Variable: INSANITY

TABLE 14.3 (*Continued*)

Path Analysis Output

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.169[a]	.028	.026	1.55785	.028	11.546	1	394	.001

[a] Predictors: (Constant), PROVOKE

AVOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	28.020	1	28.020	11.546	.001[a]
	Residual	956.198	394	2.427		
	Total	984.218	395			

[a] Predictors: (Constant), PROVOKE
[b] Dependent Variable: INSANITY

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			95.0 % Confidence Interval for B		Collinearity Statics	
		B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound	Tolerance	VIF
1	(Constant)	2.795	.259		10.803	.000	2.286	3.304		
	PROVOKE	.161	.047	.169	3.398	.001	.068	.254	1.000	1.000

[a] Dependent Variable: INSANITY

Collinearity Diagnostics[a]

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions	
				(Constant)	PROVOKE
1	**1**	1.953	1.000	.02	.02
	2	.047	6.455	.98	.98

[a] Dependent Variable: INSANITY

14.8.4 Results and Interpretation

The path model depicted in Figure 14.1 hypothesizes that subjects' interpretation of the battered woman's motives as provocation will have both *direct* and *indirect* influences on their responsibility attribution; the indirect influence being mediated by their endorsement of the insanity and self-defense strategies. The direction of the arrows depicts the hypothesized direct and indirect paths. To estimate the magnitude of these paths, a series of regression analyses were carried out.

1. The path coefficients between responsibility attribution and the three defense strategies were obtained by regressing the former on the latter. The results from the **Coefficients** table (see Table 14.3) generated from the first regression analysis show that all three defense strategies entered the prediction equation (Model 3) (i.e., all three defense strategies are significant predictors). The Beta values presented in the **Standardized Coefficients** column represent the standardized regression coefficients between responsibility attribution and the three defense strategies (SELFDEF: $\beta = -0.39$; INSANITY: $\beta = 0.13$; PROVOKE: $\beta = -0.12$).
2. The path coefficient between the self-defense strategy and the provocation defense strategy was obtained by regressing the former on the latter. The results from the **Coefficients** table generated from the second regression analysis show that provocation defense is a significant predictor of self-defense ($\beta = 0.16$).
3. The path coefficient between the insanity defense strategy and provocation is presented in the **Coefficients** table generated from the third regression analysis, and is significant ($\beta = 0.17$).

Figure 14.2 presents the path model together with the estimated standardized regression coefficients (β values) associated with the hypothesized paths.

It can be concluded that the three defense strategies of provocation, self-defense and insanity have direct influences on the subjects' responsibility attribution. The direction of the regression coefficients indicates that (1) the more the subjects endorsed the provocation and self-defense strategies, the less responsibility they attributed to the battered woman for her fatal action, and (2) the more they endorsed the insanity defense, the more they held the battered woman responsible. The results also show that at least

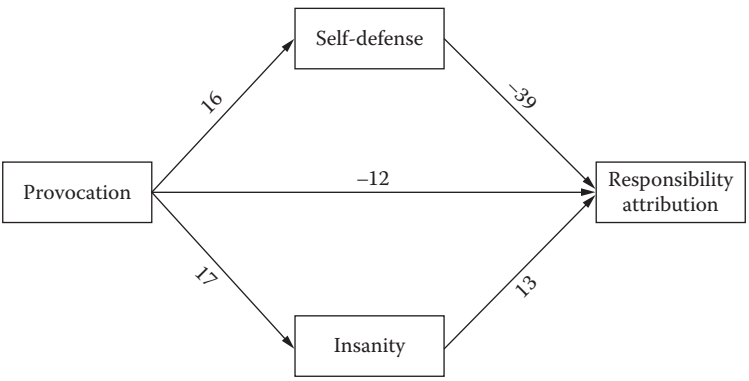


FIGURE 14.2
Path model together with estimated standardized regression coefficients.

part of the influences of provocation on responsibility attribution is indirect, being mediated by the subjects' endorsement of the self-defense and insanity strategies. Thus, the more the subjects endorsed the provocation defense, the more they believed that the woman acted in self-defense, which in turn is associated with a lower level of responsibility attribution. Endorsing the provocation defense also contributed to an increased belief that the battered woman's action was due to an impaired mental state, which in turn is associated with a higher degree of responsibility attributed to her.

14.9 Example 4: Path Analysis—Test of Significance of the Mediation Hypothesis

An important aspect of path modeling is to determine the presence or absence of a mediation effect (the mediation hypothesis). In the previous path analysis example (see Figure 14.1), the predictor exogenous variable of provocation was hypothesized to affect the dependent variable of responsibility attribution indirectly through the mediator variables of self-defense and insanity. Multiple regression analysis was used to determine the magnitude of the *direct* and *indirect* influences that each variable was hypothesized to have on the other variables that follow it in the presumed causal order (as indicated by the directional arrows). However, utilizing a series of regression analyses to determine the statistical significance of the indirect paths posited in the model offers no formal test of significance of the hypothesized indirect effects. The benefit of performing a formal significance test of the indirect effects (test of the mediation hypothesis) can be illustrated below.

Figure 14.3 depicts the mediation model and shows how the *provocation* variable's ***total effect*** (path *c*) can be apportioned into its ***indirect effect*** through the mediators of *self-defense* and *insanity*, and its ***direct effect*** on the dependent variable of *responsibility attribution* (path *c′*).

Figure 14.3a represents the ***total effect*** of *provocation* on *responsibility attribution* (path *c*). Figure 14.3b represents both the direct effect of provocation on responsibility attribution, that is, after controlling for the mediators of self-defense and insanity (path *c′*), and the indirect effects of provocation on responsibility attribution via the mediators of self-defense and insanity. The ***specific indirect effect*** of provocation on responsibility attribution via the mediator of self-defense is defined as the product of the two unstandardized paths linking provocation to responsibility attribution via self-defense. That is, the specific indirect effect of provocation on responsibility attribution through self-defense is quantified as $a1b1$. The ***total indirect effect*** of provocation on responsibility attribution is the sum of the specific indirect effects, that is, $(a1b1) + (a2b2)$. The ***total effect*** of provocation on responsibility attribution is the sum of the direct effect and all of the specific indirect effects,

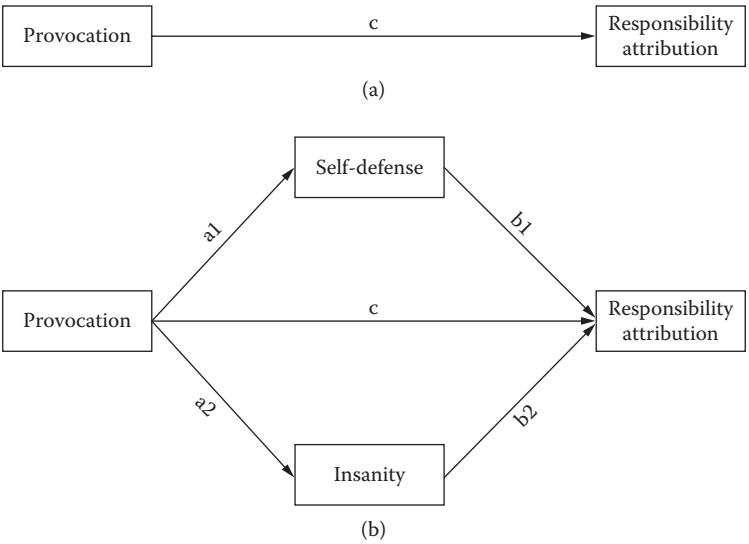


FIGURE 14.3
(a) Mediation model showing provocation's total effect on responsibility attribution. (b) Provocation is hypothesized to exert indirect effects on responsibility attribution through the mediators of self-defence and insanity.

that is, $c = c' + [(a1b1) + (a2b2)]$. The *total indirect effect* can also be calculated as $c - c'$. Performing a formal test of significance of the indirect effects involves merely testing the hypothesis of no difference between the total effect (c) and the direct effect (c'). According to Preacher and Hayes (2004, 2008), a significance test associated with the indirect effect *ab* addresses mediation more directly than a series of separate significance tests not directly involving *ab*, as was carried out in the previous example in which the mediation analysis focuses only on the statistical significance of the *a* and *b* paths. With the focus squarely on the indirect effect *ab* rather than on the separate *a* and *b* paths, emphasis is placed almost entirely on the direction and size of the indirect effects.

14.9.1 Bootstrapping

The test of significance of the *ab* indirect effect is based on the assumption that the distribution of *ab* (or $c - c'$) follows a normal distribution under the null hypothesis. However, as pointed out by Bollen and Stine (1990), this assumption is seldom met, in that not only is the distribution not necessarily normal, often it is not even symmetrical, particularly in small samples. As a consequence of these problems, Preacher and Hayes (2004) proposed bootstrapping the sampling distribution of *ab* to derive a confidence interval with the empirically derived bootstrapped sampling distribution. Bootstrapping is a nonparametric method that allows the researcher to compute estimated

standard errors, confidence intervals and hypothesis testing that makes no assumptions about the shape of the distributions of the variables or the sampling distribution of the statistic (Efron and Tibshirani, 1993). Generally bootstrapping follows the following three basic steps:

1. Resample (with replacement) a given data set a specified number of times.
2. Calculate a specific statistic from each sample.
3. Find the standard deviation of the distribution of that statistic.

To illustrate for the present example, bootstrapping is accomplished by creating many resamples by repeatedly *sampling with replacement* from the data, with each resample being the same size as the original sample size. As a result, each case can be selected as part of a bootstrap resample more than once or not at all. The indirect effect *ab* is then computed in each resample. Assume for the present example that 1,000 bootstrap resamples have been requested. The overall *ab* indirect effect is simply the mean *ab* computed over the 1,000 resamples, and the estimated standard error is the standard deviation of the 1,000 *ab* estimates. Preacher and Hayes (2008) developed macros for SPSS that allow the researcher to request tests of total and specific indirect effects by bootstrapping confidence intervals (CI) at any desired confidence level, for example, 95% or 99% CI. The macros also allow for all possible pairwise contrasts of indirect effects in multiple mediator models. The macros can be downloaded from www.quantpsy.org or from http://www.crcpress.com/e_products/downloads/download.asp?cat_no = C6021.

14.9.2 Steps in Testing the Mediation Hypothesis

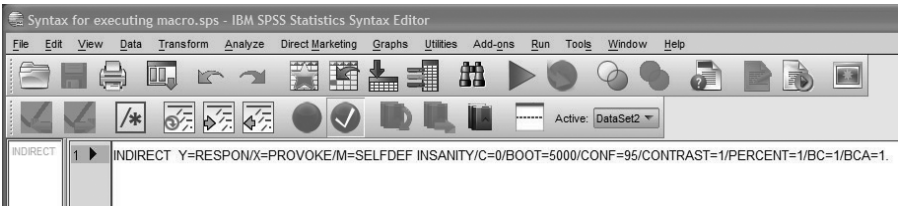
In the present example, the emphasis is on conducting formal tests of significance for the indirect effects *a*1*b*1 and *a*2*b*2 (see Figure 14.3). The analysis involves the following steps.

Step 1. Download the macros from www.quantpsy.org or from www.crcpress.com. Please follow the download instructions carefully and save the macros as an *SPSS syntax file*, that is, with the extension. **sps**. The file will be saved on your computer's desktop under the name **indirect.sps**

Step 2. Double-click on the indirect.sps file to open the syntax file within SPSS. Once the file opens, run the file to execute the macros by highlighting the syntax and clicking ▶.

Step 3. Open the data file **DOMES.SAV**. From the menu bar, click **File** then **New,** and then **Syntax.** Write the following syntax in the syntax window to conduct the analysis (tests of significance).

```
INDIRECT Y = RESPON/X = PROVOKE/M = SELFDEF INSANITY/C = 0/
BOOT = 5000/CONF = 95/CONTRAST = 1/PERCENT = 1/BC = 1/BCA = 1.
```



To execute the syntax, highlight the syntax and click ▶. See Table 14.4 for the results.

TABLE 14.4

Test of Mediation Hypothesis Output

```
indirect y = respon/x = provoke/m = selfdef insanity/c = 0/
boot = 5000/conf = 95/contrast = 1/percent = 1/bc = 1/bca = 1.
Dependent, Independent, and Proposed Mediator Variables:
DV = respon
IV = PROVOKE
MEDS = SELFDEF
INSANITY
Sample size
391
```

IV to Mediators (a paths)

	Coeff	se	t	P
SELFDEF	.1725	.0526	3.2759	.0011
INSANITY	.1650	.0478	3.4519	.0006

Direct Effects of Mediators on DV (b paths)

	Coeff	se	t	p
SELFDEF	-.4081	.0494	-8.2617	.0000
INSANITY	.1545	.0544	2.8395	.0048

Total Effect of IV on DV (c path)

	Coeff	se	t	p
PROVOKE	-.1738	.0559	-3.1080	.0020

Direct Effect of IV on DV (c′ path)

	Coeff	se	t	p
PROVOKE	-.1290	.0525	-2.4559	.0145

Model Summary for DV Model

R-sq	Adj R-sq	F	df1	df2	p
.1985	.1923	31.9465	3.0000	387.0000	.0000

TABLE 14.4 (*Continued*)

Test of Mediation Hypothesis Output

BOOTSTRAP RESULTS FOR INDIRECT EFFECTS

Indirect Effects of IV on DV through Proposed Mediators (ab paths)

	Data	Boot	Bias	SE
TOTAL	−.0449	−.0444	.0005	.0308
SELFDEF	−.0704	−.0698	.0006	.0261
INSANITY	.0255	.0254	−.0001	.0129
C1	−.0959	−.0952	.0006	.0272

Bias Corrected and Accelerated Confidence Intervals

	Lower	Upper
TOTAL	−.1087	.0149
SELFDEF	−.1290	−.0246
INSANITY	.0059	.0588
C1	−.1551	−.0477

Bias Corrected Confidence Intervals

	Lower	Upper
TOTAL	−.1087	.0147
SELFDEF	−.1293	−.0249
INSANITY	.0060	.0592
C1	−.1563	−.0478

Percentile Confidence Intervals

	Lower	Upper
TOTAL	−.1068	.0166
SELFDEF	−.1250	−.0218
INSANITY	.0041	.0542
C1	−.1516	−.0449

```
************************************************************
Level of Confidence for Confidence Intervals:
95
Number of Bootstrap Resamples:
5000
************************************************************
```

INDIRECT EFFECT CONTRAST DEFINITIONS: Ind_Eff1 MINUS Ind_Eff2

Contrast	IndEff_1	IndEff_2
C1	SELFDEF	INSANITY

```
********************************* NOTES
 ******************************
— — — END MATRIX— — —
```

14.9.3 SPSS Output

14.9.4 Results and Interpretation

To reiterate, the focus of the analysis is to determine the statistical significance of the indirect effects (*a1b1*, *a2b2*) posited in the path model (Figure 14.3b). The interpretation of the analysis results follows.

14.9.4.1 Total Indirect Effect

Referring to the **Bootstrap Results for Indirect Effects** table, it can be seen that the bootstrapped estimate of the ***total indirect effect*** of provocation on responsibility attribution is similar to the point estimate computed from the conventional regression analysis of the raw data (Boot = −0.0444, Data = −0.0449). To determine the significance of the bootstrapped estimate, compute the critical ratio for the total indirect effect of provocation on responsibility attribution.

$$Z = \frac{Bootstrapped\ estimate}{Standard\ error\ (SE)} = \frac{-0.0444}{0.0308} = -1.442$$

As this value is less than −1.96 we fail to reject the null hypothesis that the total indirect effect is zero ($p > .05$).

The above result from the critical ratio test is supported by the bootstrapped 95% confidence intervals (CIs) (BCa: −0.1087 to 0.0149; BC: −0.1087 to 0.0147; percentile: −0.1068 to 0.0166). As zero is contained within the 95% confidence intervals, it can be concluded that the total indirect effect is not significantly different from zero at $p > .05$ (two-tailed). The interpretation of these results is that, taken as a set, *self-defense* and *insanity* do not mediate the effect of *provocation* on *responsibility attribution*.

14.9.4.2 Specific Indirect Effects

In testing the mediation hypothesis, interest is directed not only at the total indirect effect of provocation on responsibility attribution but also with the specific indirect effects. The specific indirect effects (obtained via bootstrapping) are *a1b1* = −0.0698 (through *self-defense*) and *a2b2* = 0.0254 (through *insanity*) (see the **Bootstrap Results for Indirect Effects** table). The critical ratio tests of significance for these indirect effects are computed as follows:

$$\text{Self-defense: } Z = \frac{Bootstrapped\ estimate}{Standard\ error\ (SE)} = \frac{-0.0698}{0.0261} = -2.674$$

$$\text{Insanity: } Z = \frac{Bootstrapped\ estimate}{Standard\ error\ (SE)} = \frac{0.0254}{0.0129} = 1.969$$

As these values are greater than 1.96, we can reject the null hypothesis that the specific indirect effect is zero ($p < .05$). That is, it can be concluded that both self-defense and insanity are significant mediators of provocation on responsibility attribution.

The above results are supported by the bootstrapped 95% confidence intervals (CIs) for both self-defense (BCa: −0.1290 to −0.0246; BC: −0.1293 to −0.0249; percentile: −0.1250 to −0.0218) and insanity (BCa: 0.0059 to 0.0588; BC: 0.0060 to 0.0592; percentile: 0.0041 to 0.0542). As zero is not contained within the 95% confidence intervals, it can be concluded that the specific indirect effects (through self-defense and insanity) are significantly different from zero at $p < .05$ (two-tailed). Thus, although results from the test of total indirect effect show that taken as a set, *self-defense* and *insanity* do not mediate the effect of *provocation* on *responsibility attribution*, individually both *self-defense* and *insanity* are significant mediators of the provocation > responsibility attribution relationship.

14.9.4.3 Pairwise Contrast of Indirect Effects

The above results showed that the specific indirect effects of provocation on responsibility attribution through both self-defense and insanity were significantly different from zero. It may be of interest to see whether these two indirect effects differ significantly. Results from the pairwise contrast (**C1**) are presented under the **Bootstrap Results for Indirect Effects** table. The 95% confidence intervals (BCa, BC, percentile) for the contrast were computed as follows—BCa: −0.1551 to −0.0477; BC: −0.1563 to −0.0478; percentile: −0.1516 to −0.0449. As zero is not contained in the intervals, it can be concluded that the two indirect effects differ significantly in terms of magnitude, with the specific indirect effect through *self-defense* (−0.0698) being larger than the specific indirect effect through *insanity*.

15

Multiple Discriminant Analysis

15.1 Aim

Discriminant function analysis is a statistical technique that allows the researcher to determine which continuous variables discriminate between two or more naturally occurring groups. When the analysis involves two groups, the technique is referred to as *discriminant analysis*. When there are more than two groups, the technique is referred to as *multiple discriminant analysis* (MDA). For example, a medical researcher may want to investigate which variables discriminate between smokers, ex-smokers, and non-smokers. For this purpose, the researcher could collect data on a number of variables known to influence the uptake and maintenance of smoking behavior (e.g., age, educational level, occupation status). Multiple discriminant analysis could then be used to:

1. Optimize the *predictive discriminant functions* calculated from the predictor variables that maximally discriminate the levels of the dependent variable (smokers, ex-smokers, and non-smokers).
2. Derive the *classification function* that will classify the subjects into different groups with a better than chance accuracy.

15.2 Checklist of Requirements

- The dependent variable must be categorical (nominal) and can consist of two groups (e.g., male versus female; low versus high) or more than two groups (e.g., Chinese, Americans, Australians). The independent variables are metric (continuous).
- The size of the sample has a direct impact on the stability of the results. Specifically, the results become unstable as the sample size

decreases relative to the number of independent variables. While the minimum sample size recommended is five observations per independent variable, the "rule of thumb" suggests a ratio of 20 observations for each predictor variable (Hair, Anderson, Tatham, and Black, 1995). In addition, the sample size for each group (levels of the dependent variable) must also be considered. As a practical guideline, each group should possess at least 20 observations. Unequal group sample sizes are acceptable. However, group sizes that are widely different may impact the estimation of the discriminant function and the categorization of observations (e.g., larger groups have a disproportionately higher chance of classification) (Hair et al., 1995). If group sizes do vary markedly, then randomly sample from the larger group(s) to obtain group sizes comparable to the smaller group(s).

15.3 Assumptions

- **Linearity**—Similar to other multivariate techniques that employ a variate (i.e., linear combination that represents the weighted sum of two or more predictor variables that discriminate best between a priori defined groups), an implicit assumption is that all relationships among all pairs of predictors within each group are linear. However, violation of this assumption is less serious than others in that it tends to lead to reduced power rather than increase Type I error (Tabachnick and Fidell, 2001).
- **Multivariate normality**—The assumption is that scores on each predictor variable are normally distributed (*univariate normality*) and that the sampling distribution of the combination of two or more predictors is also normally distributed (*multivariate normality*). Multivariate normality is difficult to test and currently there are no specific tests capable of testing the normality of all linear combinations of sampling distributions of predictors. However, since multivariate normality implies univariate normality (although the reverse is not necessarily true), a situation in which all variables exhibit univariate normality will help gain, although not guarantee, multivariate normality (Hair et al., 1995).
- **Homogeneity of variance-covariance matrices**—When sample sizes are unequal and small, unequal covariance matrices can adversely affect the results of significance testing. Even with decently sized samples, heterogeneity of variance-covariance matrices can affect

the classification process whereby cases are "overclassified" into groups with greater variability (Tabachnick and Fidell, 2001). A test of this assumption can be made via Box's M. As this test is overly sensitive (increases the probability of Type I error), an alpha level of .001 is recommended.

- **Multicollinearity**—As with multiple regression analysis, multicollinearity denotes the situation where the independent/predictor variables are highly correlated. When independent variables are multicollinear, there is "overlap" or sharing of predictive power so that one variable can be highly explained or predicted by the other variable(s). Thus, that predictor variable adds little to the explanatory power of the entire set.
- **Outliers**—MDA is highly sensitive to the presence of outliers in that their inclusion can have a substantial impact on the classification accuracy of the discriminant analysis results. Therefore, it is prudent to examine all results for the presence of both univariate and multivariate outliers and to eliminate significant outliers before conducting MDA.

15.4 Example 1: Two-Group Discriminant Analysis

A study was conducted to identify the factors underlying dangerous driving among young Australian males (Ho and Gee, 2008). A number of factors were identified, namely *driving fast/risk taking, disrespect for traffic laws, sensation-seeking, and danger assessment*. The statements written to reflect these four factors were rated on six-point scales, with high scores indicating strong endorsement of these factors. A total of 380 subjects provided responses to these statements. Discriminant analysis will be used to determine how these factors discriminate between young drivers who had been involved in a motor vehicle accident (GROUP = 1) and those who had not (GROUP = 2). These factors (together with their SPSS variable name) are listed below.

- **FAST_RIS**—A desire to drive fast and/or to take risks while driving.
- **DISRESP**—A negative attitude (disrespect) toward traffic laws.
- **DANGER**—The extent to which a variety of activities are considered to be dangerous.
- **SEN_SEEK**—The need for varied, novel, and complex sensations and experiences.

15.4.1 Data Entry Format

The data set is named **DRIVE_1.**

Variables	Column(s)	Code
Age	1	In years
Educ	2	1 = primary, to 6 = tertiary
Income	3	1 = <$10,000 per year, 5 = ≥$60,001 per year
Employ	4	1 = full-time paid employment, 6 = students (study and work)
m1 to m40	5–44	1 = strongly disagree, 6 = strongly agree
ss1 to ss10	45–54	1 = strongly disagree, 6 = strongly agree
d1 to d19	55–73	1 = not at all dangerous, 6 = very dangerous
fast_ris	74	1 = strongly disagree, 6 = strongly agree
disresp	75	1 = strongly disagree, 6 = strongly agree
group	76	1 = accident, 2 = no accident
sen_seek	77	1 = strongly disagree, 6 = strongly agree
danger	78	1 = not at all dangerous, 6 = very dangerous
group_1	79	1 = accident-charged, 2 = accident-no charge, 3 = no accident

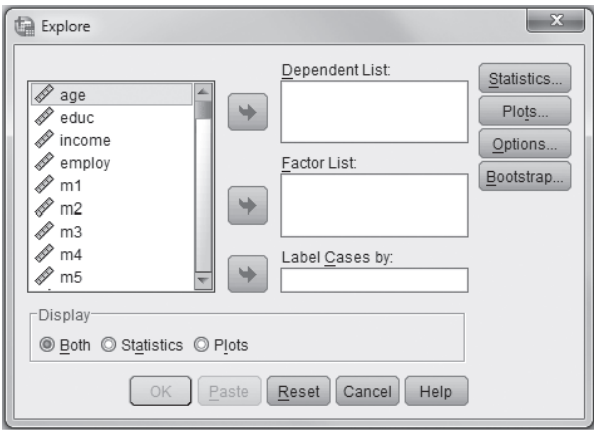
15.4.2 Testing Assumptions

15.4.2.1 Multivariate Normality

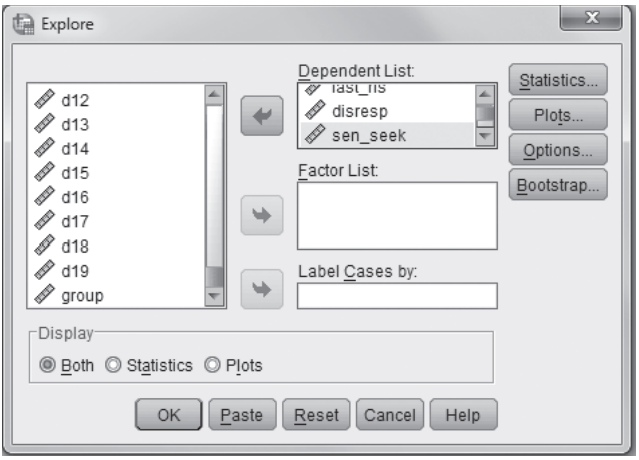
As there are no specific tests that test for multivariate normality, univariate normality for the four predictor variables will be tested instead.

15.4.2.1.1 Windows Method

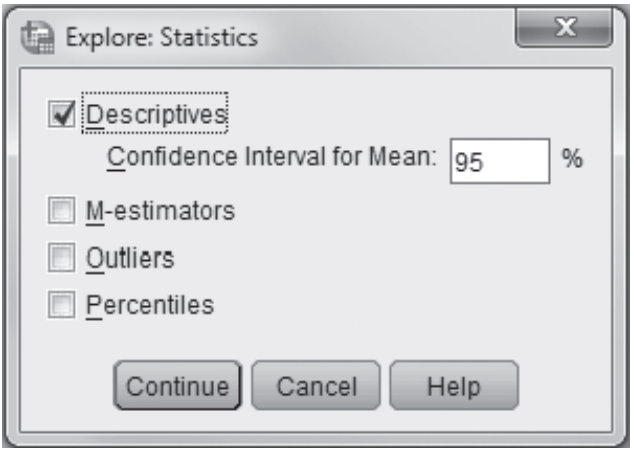
1. From the menu bar, click **Analyze**, then **Descriptive Statistics**, and then **Explore…**. The following **Explore** window will open.



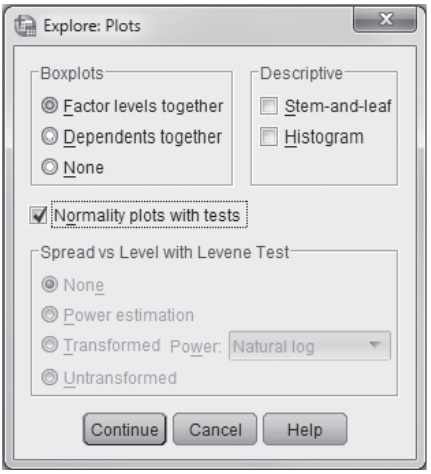
2. Transfer the **FAS_RIS**, **DISRESP**, **SEN_SEEK**, and **DANGER** variables to the **Dependent List:** field by clicking these variables (highlight) and then clicking .



3. Click Statistics... to open the **Explore: Statistics** window. Check the **Descriptives** field and click Continue to return to the **Explore** window.



4. In the **Explore** window click Plots... to open the **Explore: Plots** window. Check the **Factor levels together** field and the **Normality plots with tests** field. Click Continue to return to the **Explore** window.
5. When the **Explore** window opens, click OK to complete the analysis. See Table 15.1 for the results.



15.4.2.1.2 SPSS Syntax Method

```
EXAMINE VARIABLES = FAST_RIS DISRESP SEN_SEEK DANGER
/PLOT NPPLOT
/COMPARE GROUPS
/STATISTICS DESCRIPTIVES
/CINTERVAL 95
/MISSING LISTWISE
/NOTOTAL.
```

15.4.2.1.3 SPSS Output

TABLE 15.1

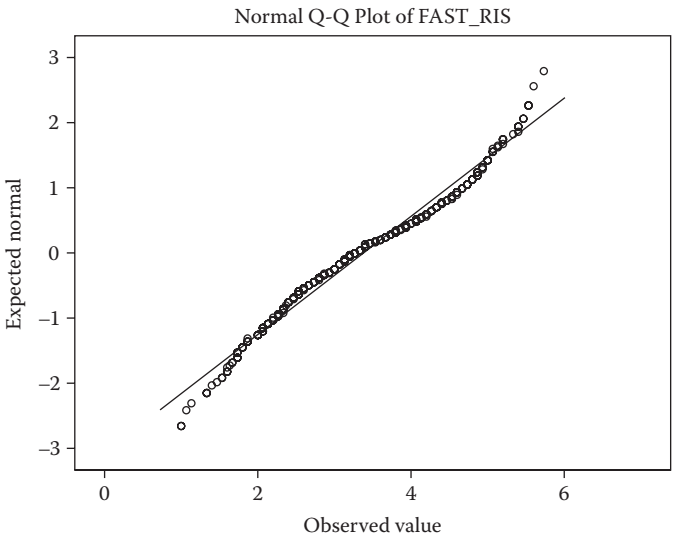
Explore Analysis (Selected) Output

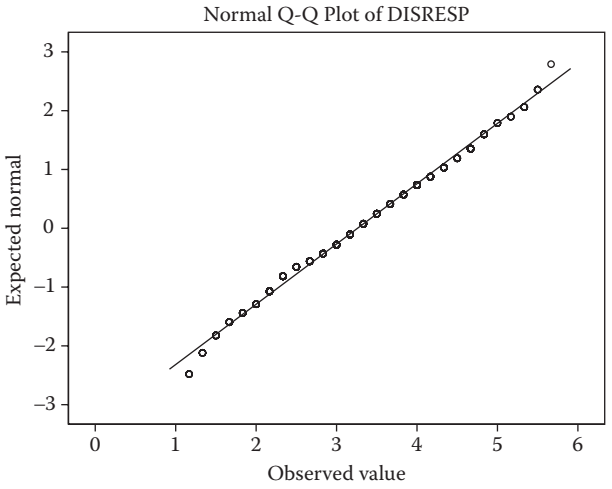
Descriptives				
			Statistic	**Std. Error**
fast_ris	Mean		3.3809	.05655
	95% confidence interval for mean	Lower bound	3.2697	
		Upper bound	3.4921	
	Median		3.2667	
	Variance		1.215	
	Std. deviation		1.10239	
	Skewness		.092	.125
	Kurtosis		−.954	.250
disresp	Mean		3.2605	.05004
	95% confidence interval for mean	Lower bound	3.1621	
		Upper bound	3.3589	
	Median		3.3333	

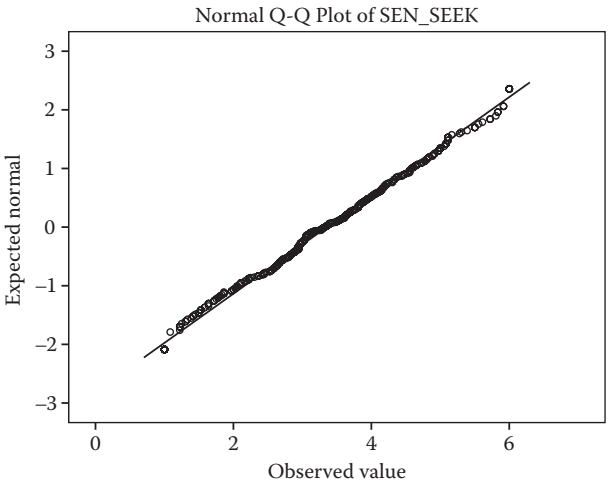
TABLE 15.1 (*Continued*)

Explore Analysis (Selected) Output

Descriptives				
			Statistic	**Std. Error**
	Variance		.951	
	Std. deviation		.97542	
	Skewness		.073	.125
	Kurtosis		−.544	.250
Sen_ seek	Mean		3.3552	.06118
	95% confidence interval for mean	Lower bound	3.2349	
		Upper bound	3.4755	
	Median		3.3194	
	Variance		1.422	
	Std. deviation		1.19265	
	Skewness		.028	.125
	Kurtosis		−.497	.250
Danger	Mean		3.0025	.04613
	95% confidence interval for mean	Lower bound	2.9118	
		Upper bound	3.0932	
	Median		3.0595	
	Variance		.809	
	Std. deviation		.89929	
	Skewness		−.056	.125
	Kurtosis		−.368	.250







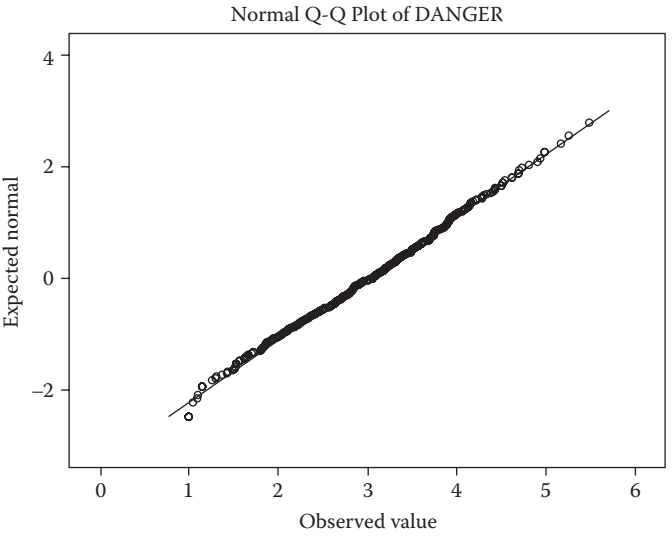
15.4.2.1.4 Interpretation

A simple diagnostic test for normality is based on the skewness and kurtosis values. The statistical *z* value for the skewness value is calculated as:

$$Z_{\text{skewness}} = \frac{\text{skewness}}{\sqrt{\text{s.e. skewness}}}$$

The statistical *z* value for the kurtosis value is calculated as:

$$Z_{\text{kurtosis}} = \frac{\text{kurtosis}}{\sqrt{\text{s.e. kurtosis}}}$$



If the calculated z value exceeds the specified critical probability value, then the distribution is nonnormal. For example, a calculated z value exceeding ±2.58 will result in a rejection of the assumption of normality at the .01 critical probability (alpha) level. A calculated z value exceeding ±1.96 will result in a rejection of the assumption of normality at the .05 alpha level. On the basis of the obtained skewness statistics, the z values for the four predictor variables **FAST_RIS** (.74), **DISRESP** (.58), **SEN_SEEK** (.22), and **DANGER** (−.13) are less than ±1.96. Thus, it can be concluded that the distribution of these four variables does not depart significantly from normality.

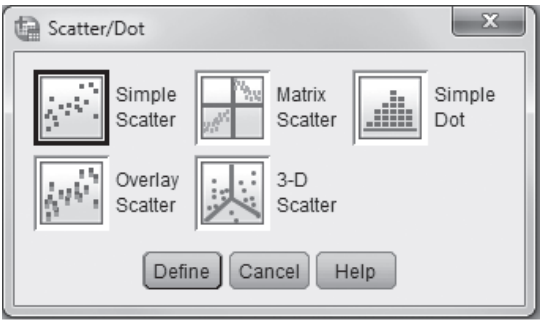
Another diagnostic test for normality is a visual check of the **normal probability plot**, which compares the cumulative distribution of the observed values with the expected values derived from the normal distribution. The normal distribution forms a straight diagonal line, and if a variable's distribution is normal, the data distribution will fall more or less on the diagonal. Inspection of the normal probability plots shows very little departure from normality for all four predictor variables.

15.4.2.2 Linearity

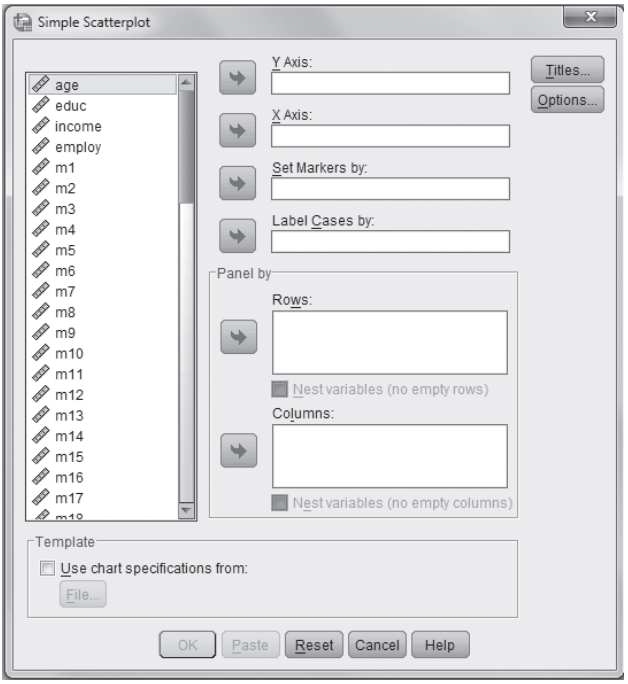
Linearity can be assessed by examining the scatterplots of the variables.

15.4.2.2.1 Windows Method

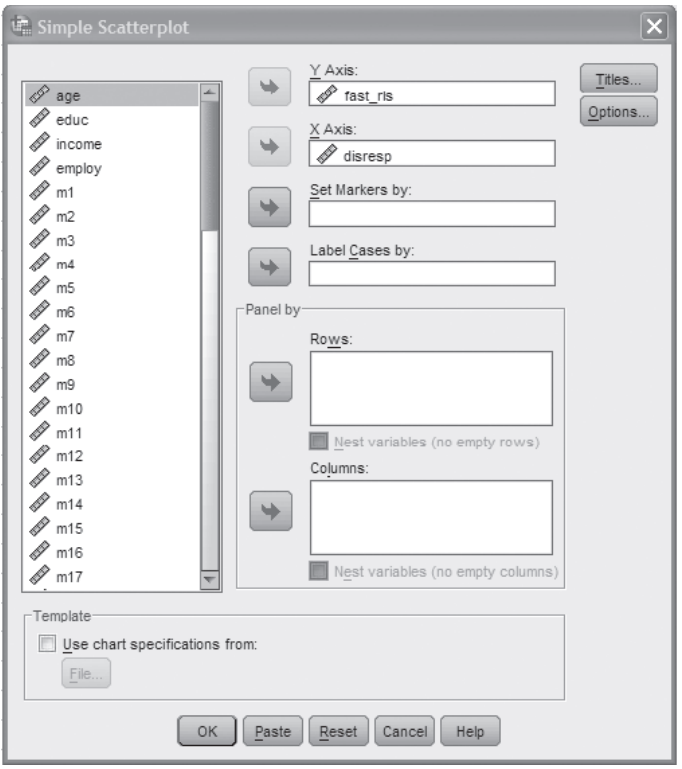
1. From the menu bar, click **Graphs**, then **Legacy Dialogs**, and then **Scatter/Dot…**. The following **Scatter/Dot** window will open.



2. Click **Simple Scatter** and then Define. The following **Simple Scatterplot** window will open.



3. To obtain a scatterplot for the variables **FAST_RIS** and **DISRESP**, transfer the **FAST_RIS** variable to the **Y Axis:** field and the **DISRESP** variable to the **X Axis**: field by clicking these variables (highlight) and then clicking .
4. Click OK to complete the analysis. Repeat steps 1–4 to obtain scatterplots for all other pairs of independent variables. See Figure 15.1 for the scatterplots.



15.4.2.2.2 SPSS Syntax Method

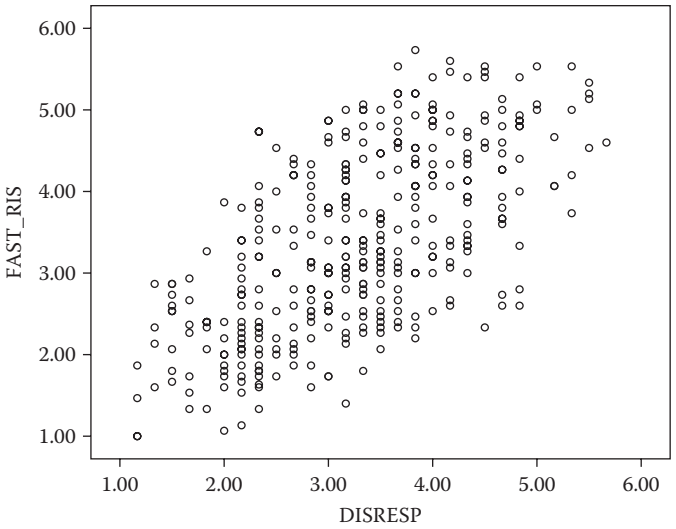
```
GRAPH
/SCATTERPLOT(BIVAR) = DISRESP WITH FAST_RIS
/MISSING = LISTWISE.
GRAPH
/SCATTERPLOT(BIVAR) = DISRESP WITH SEN_SEEK
/MISSING = LISTWISE.
GRAPH
/SCATTERPLOT(BIVAR) = DISRESP WITH DANGER
/MISSING = LISTWISE.
GRAPH
/SCATTERPLOT(BIVAR) = FAST_RIS WITH SEN_SEEK
/MISSING = LISTWISE.
GRAPH
/SCATTERPLOT(BIVAR) = FAST_RIS WITH DANGER
/MISSING = LISTWISE.
GRAPH
/SCATTERPLOT(BIVAR) = SEN_SEEK WITH DANGER
/MISSING = LISTWISE.
```

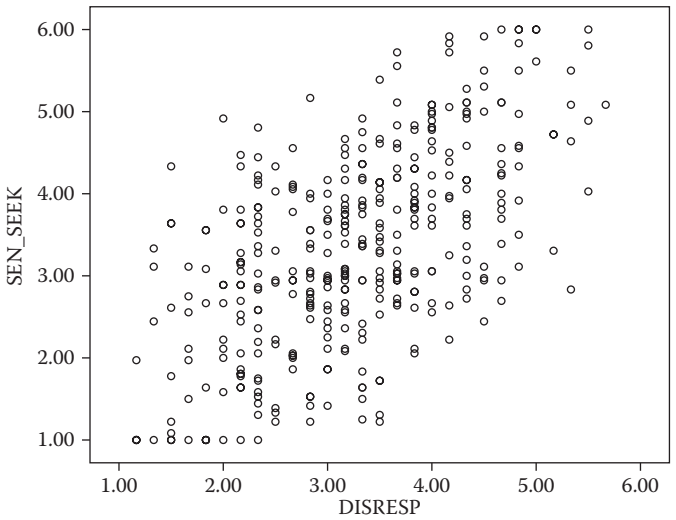
15.4.2.2.3 SPSS Output

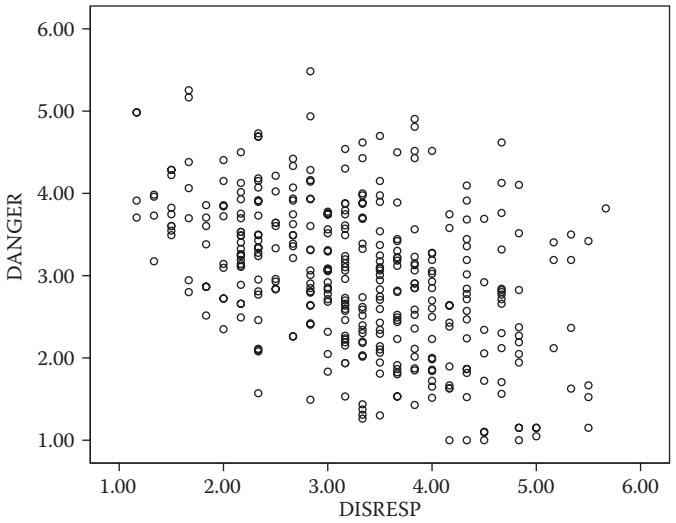
Examination of the scatterplots (Figure 15.1) showed no serious nonlinearity between the pairs of independent variables.

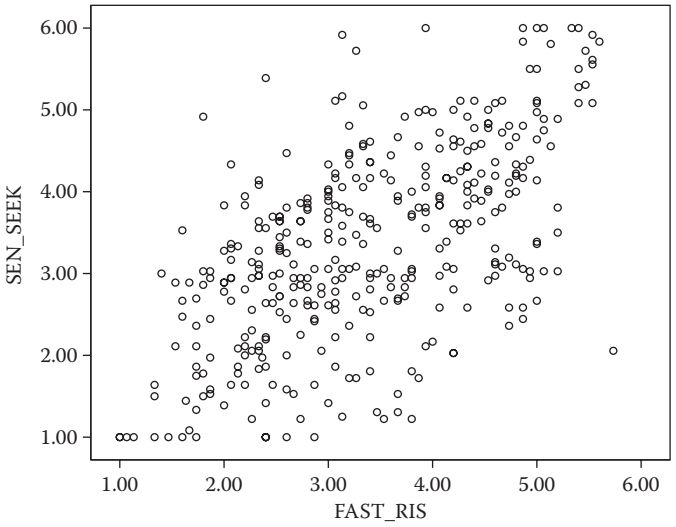
15.4.2.3 Univariate Outliers

Univariate outliers are cases with very large z (standardized) scores on the variables. For small samples (80 or fewer observations), identify those cases with z scores of 2.5 or greater as outliers ($p < .01$). For large samples, cases









with z scores of 3.29 or greater ($p < .001$) are potential outliers. The equation for converting data values to standard scores is:

$$z = \frac{\text{data point-mean}}{\text{standard deviation}}$$

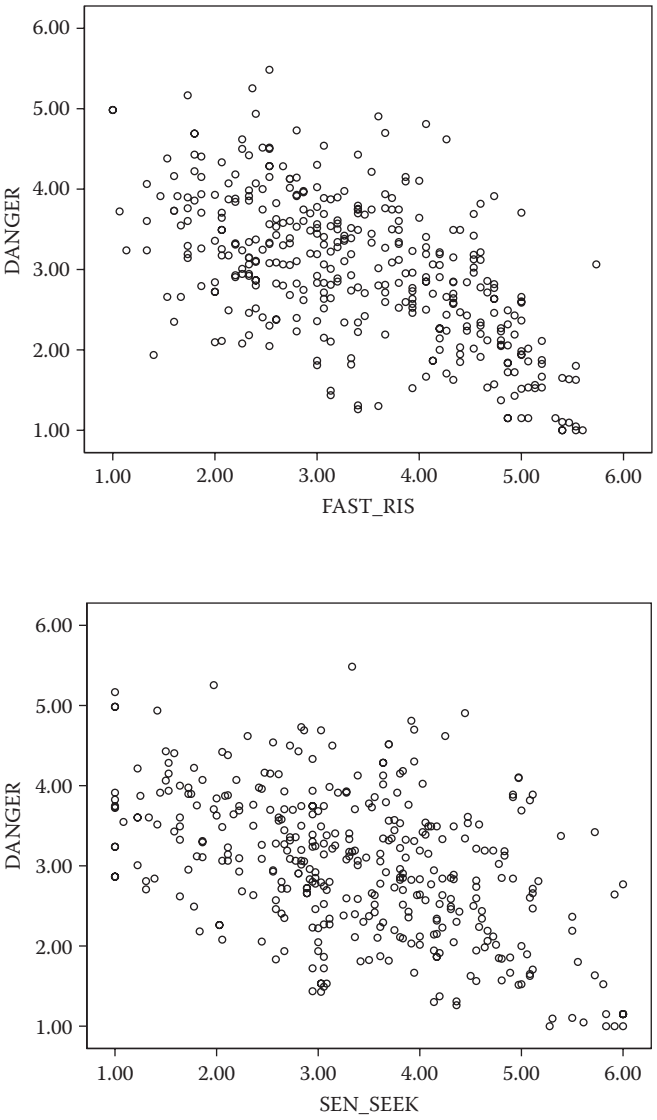
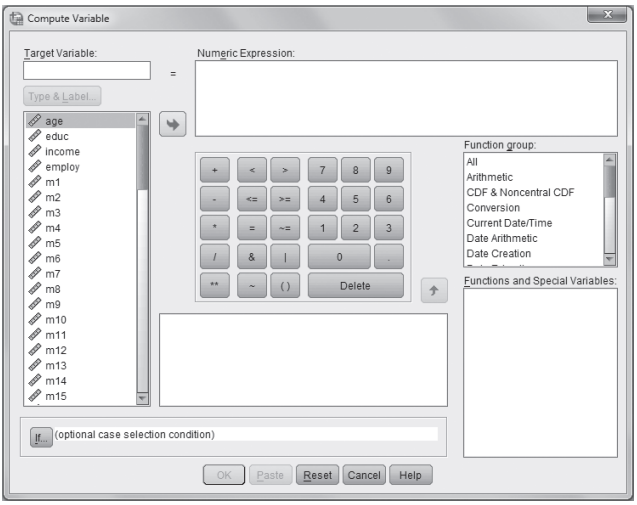


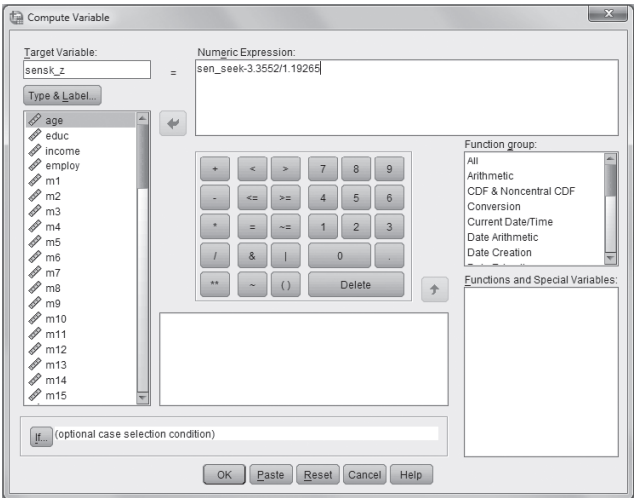
FIGURE 15.1
Scatterplots for the variables of FAS_RIS, DISRESP, SEN_SEEK, and DANGER.

15.4.2.3.1 Windows Method

1. From the menu bar, click **Transform**, and then **Compute Variable…**. The following **Compute Variable** window will open.
2. To compute the *z* scores for the **SEN_SEEK** variable, it will be necessary to create a new variable (e.g., **SENSK_Z**) to represent these *z* scores. To do this, obtain the mean (3.3552) and the

standard deviation (1.19265) for the **SEN_SEEK** variable, which can be found in Table 15.1 (Explore Analysis Output). In the





Target Variable: field, type the name **SENSK_Z**. In the **Numeric Expression:** field, type the equation for computing the *z* scores, that is, SEN_SEEK—3.3552/1.19265.

3. Click OK to complete the analysis. The newly created variable **SENSK_Z** will be added to the end of the data file. Repeat steps 1 and 2 to compute *z* scores for the other three independent variables (**DANGER_Z**, **FASTRK_Z**, **DISRESP_Z**).

15.4.2.3.2 SPSS Syntax Method

```
COMPUTE SENSK _ Z = SEN _ SEEK-3.3552/1.19265.
COMPUTE DANGER _ Z = DANGER-3.0025/.89929.
COMPUTE FASTRK _ Z = FAST _ RIS-3.3809/1.10239.
COMPUTE DISRESP _ Z = DISRESP-3.2605/.97542.
EXECUTE.
```

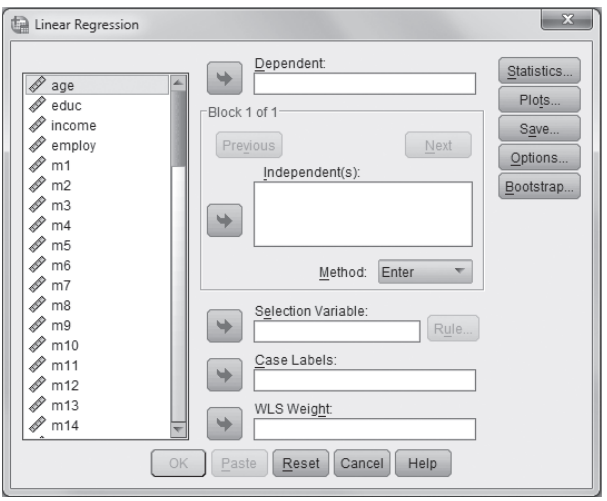
sensk_z	danger_z	fastrk_z	disresp_z
2.16	.76	.93	1.49
1.66	.27	.13	−.18
1.27	.15	1.27	−.68
1.30	−.57	1.67	−1.01
.21	−1.81	2.00	.32

15.4.2.3.3 Interpretation

A section of the data file containing the computed z scores for the variables **SENSK_Z**, **DANGER_Z**, **FASTRK_Z**, and **DISRESP_Z** is presented here. Examination of these z values shows no univariate outliers, that is, there are no cases with z scores of 3.29 or greater ($p < .001$).

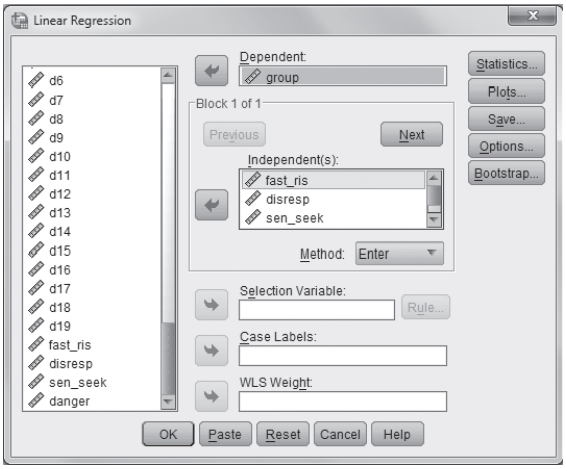
15.4.2.4 Multivariate Outliers

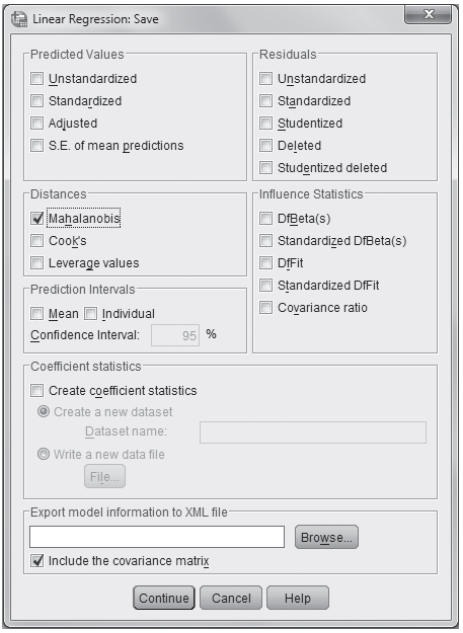
Multivariate outliers can be assessed using the **Mahalanobis distance** statistic obtained from regression analysis. The variable GROUP (coded 1 = accident, 2 = no accident) will be utilized as the dependent variable (Note: Multivariate outliers among independent variables are unaffected by the dependent variable.)



15.4.2.4.1 Windows Method: Regression Analysis

1. From the menu bar, click **Analyze**, then **Regression**, and then **Linear…**. The following **Linear Regression** window will open.
2. Click (highlight) the **GROUP** variable and then click ▶ to transfer this variable to the **Dependent:** field. Click (highlight) the variables **FAS_RIS**, **DISRESP**, **SEN_SEEK**, and **DANGER** and then click ▶ to transfer these variables to the **Independent(s):** field. In the **Method:** cell, select **Enter** as the method of entry for this set of independent variables.





3. Click Save... to open the **Linear Regression: Save** window. Check the **Mahalanobis** cell and the **Include the covariance matrix** cell.

   Click Continue to return to the **Linear Regression** window.
4. When the **Linear Regression** window opens, click OK to complete the analysis. The Mahalanobis distances will be added to the data file.

15.4.2.4.2 SPSS Syntax Method

```
REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA
/CRITERIA = PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT GROUP
/METHOD = ENTER FAST _ RIS DISRESP SEN _ SEEK DANGER
/SAVE MAHAL.
```

15.4.2.4.3 Interpretation

A section of the data file containing the saved Mahalanobis distances is presented here.

MAH_1
8.80414
2.83608
4.90521
7.01874
6.16360
5.81314
5.57188
7.02920
7.76000
7.59630

The significance of the Mahalanobis distance is evaluated by the chi-square test. For four independent variables (d$f = 3$), the critical value of chi-square at an alpha level of $p < .001$ is 16.268. An examination of the saved Mahalanobis distances shows no outlying cases.

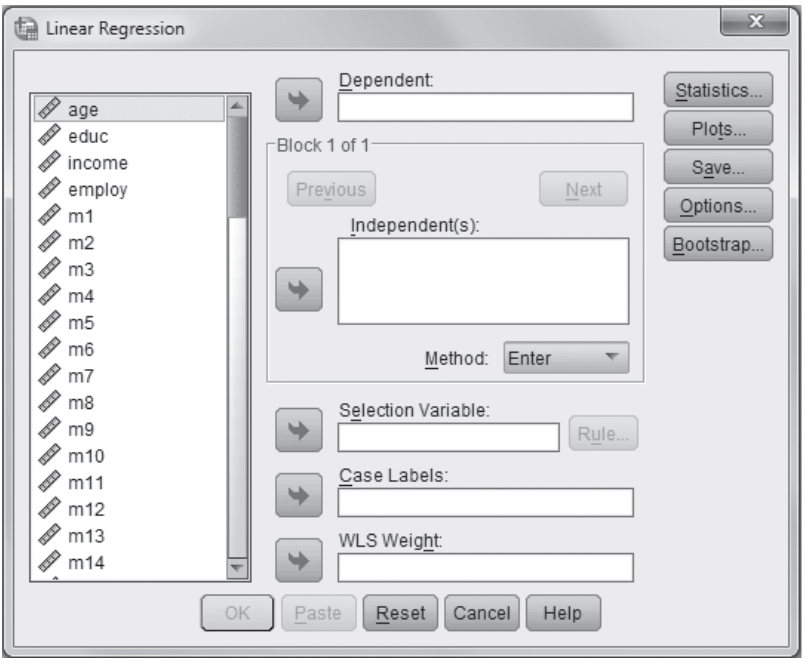
15.4.2.5 Multicollinearity

Section 14.5 in Chapter 14 describes how multicolliearity can be assessed via multiple regression analysis. From the set of four predictor variables (**FAS_RIS, DISRESP, SEN_SEEK, DANGER**), use one (say, **FAS_RIS**) as the dependent variable, and **DISRESP**, **SEN_SEEK**, and **DANGER** as predictors. Compute R^2

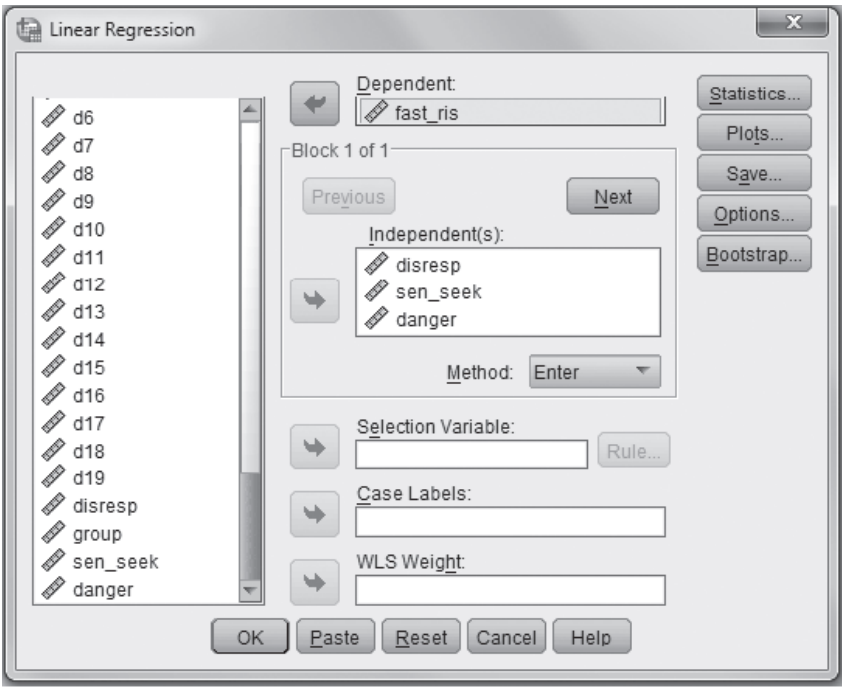
(the proportion of variance that **DISRESP**, **SEN_SEEK**, and **DANGER** explain in **FAS_RIS**), and then take 1 – R^2. Thus, the "tolerance" value is an indication of the percent of variance in the predictor **FAS_RIS** that cannot be accounted for by the other predictors. Hence, very small values (< .10) indicate "overlap" or sharing of predictive power (i.e., the predictor is redundant).

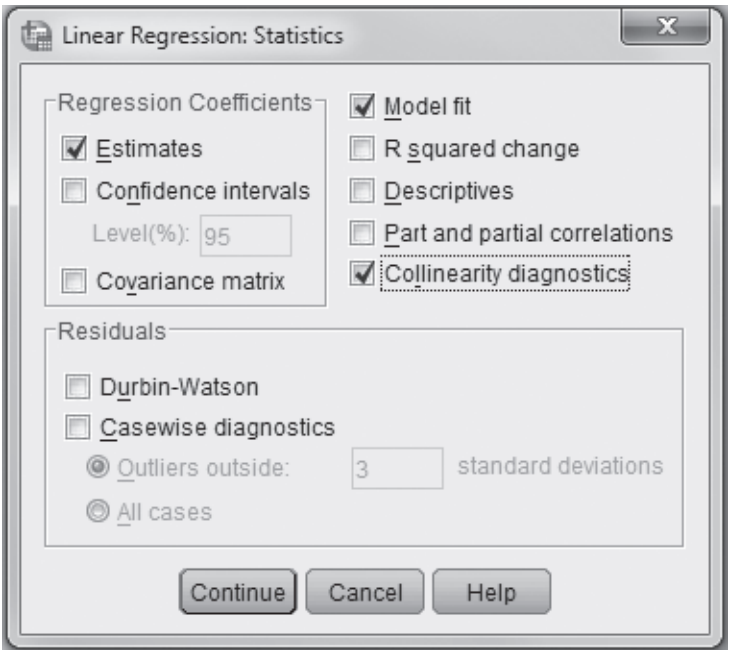
15.4.2.5.1 Windows Method

1. From the menu bar, click **Analyze**, then **Regression**, and then **Linear…**. The following **Linear Regression** window will open.



2. Click (highlight) the **FAS_RIS** variable and then click ▶ to transfer this variable to the **Dependent:** field. Click (highlight) the variables **DISRESP**, **SEN_SEEK**, and **DANGER** and then click ▶ to transfer these variables to the **Independent(s):** field. In the **Method:** cell, select **Enter** as the method of entry for this set of independent variables.
3. Click Statistics... to open the **Linear Regression: Statistics** window. Check the **Estimates** cell, **Model fit** cell, and the **Collinearity diagnostic** cell. Click Continue to return to the **Linear Regression** window.
4. When the **Linear Regression** window opens, click OK to complete the analysis.
5. Repeat steps 1–4 for the other three predictor variables.


Linear Regression
d6
d7
d8
d9
d10
d11
d12
d13
d14
d15
d16
d17
d18
d19
disresp
group
sen_seek
danger
Dependent:
fast_ris
Block 1 of 1
Previous
Next
Independent(s):
disresp
sen_seek
danger
Method:
Enter
Selection Variable:
Rule...
Case Labels:
WLS Weight:
Statistics...
Plots...
Save...
Options...
Bootstrap...
OK
Paste
Reset
Cancel
Help


Linear Regression: Statistics
Regression Coefficients
Estimates
Confidence intervals
Level(%): 95
Covariance matrix
Model fit
R squared change
Descriptives
Part and partial correlations
Collinearity diagnostics
Residuals
Durbin-Watson
Casewise diagnostics
Outliers outside: 3 standard deviations
All cases
Continue
Cancel
Help

15.4.2.5.2 SPSS Syntax Method

```
REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA TOL
/CRITERIA = PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT FAST_RIS
/METHOD = ENTER DISRESP SEN_SEEK DANGER.
REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA TOL
/CRITERIA = PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT DISRESP
/METHOD = ENTER FAST_RIS SEN_SEEK DANGER.
REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA TOL
/CRITERIA = PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT SEN_SEEK
/METHOD = ENTER DISRESP DANGER FAST_RIS.
REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA TOL
/CRITERIA = PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT DANGER
/METHOD = ENTER DISRESP FAST_RIS SEN_SEEK.
```

15.4.2.5.3 Interpretation

The "tolerance" diagnostic results for multicollinearity are presented as follows.

1. Dependent variable = **FAST_RIS**; predictors = **DISRESP**, **SEN_SEEK**, **DANGER**

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	2.488	.268		9.267	.000		
	disresp	.387	.050	.343	7.750	.000	.616	1.623
	sen_seek	.231	.042	.249	5.543	.000	.595	1.680
	danger	−.381	.050	−.311	−7.637	.000	.728	1.373

[a] Dependent Variable: fast_ris

2. Dependent variable = **DISRESP**; predictors = **FAST_RIS**, **SEN_SEEK**, **DANGER**

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	1.386	.276		5.024	.000		
	fast_ris	.356	.046	.402	7.750	.000	.526	1.903
	sen_seek	.261	.039	.319	6.645	.000	.615	1.626
	danger	−.067	.051	−.062	−1.315	.189	.633	1.579

[a] Dependent Variable: disresp

3. Dependent variable = **SEN_SEEK**; predictors = **FAST_RIS**, **DISRESP**, **DANGER**

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	1.543	.346		4.462	.000		
	disresp	.403	.061	.330	6.645	.000	.594	1.684
	danger	−.203	.063	−.153	−3.226	.001	.648	1.543
	fast_ris	.328	.059	.303	5.543	.000	.490	2.040

[a] Dependent Variable: sen_seek

4. Dependent variable = **DANGER**; predictors = **SEN_SEEK, FAST_RIS**, **DISRESP**

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	4.860	.138		35.138	.000		
	disresp	−.068	.052	−.074	−1.315	.189	.534	1.873
	fast_ris	−.353	.046	−.432	−7.637	.000	.524	1.910
	sen_seek	−.132	.041	−.176	−3.226	.001	.566	1.768

[a] Dependent Variable: danger

Examination of the four **Coefficients** tables shows that all the "tolerance" values are greater than 0.10. Thus, multicollinearity does not seem to be a problem for the four predictor variables.

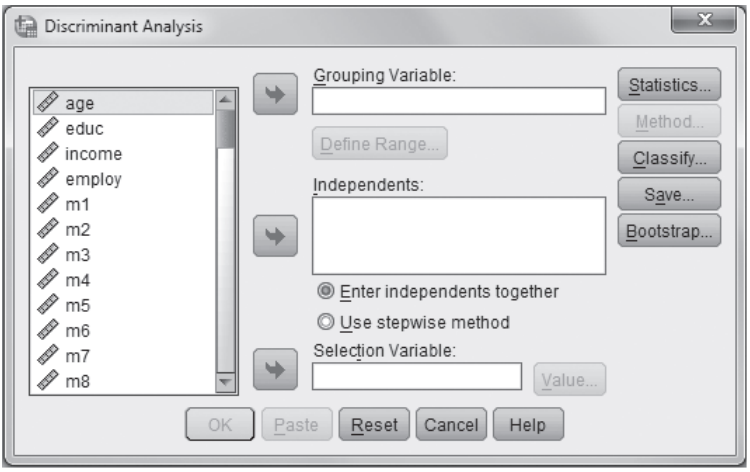
15.4.2.6 Homogeneity of Variance-Covariance Matrices

This assumption will be tested with Box's *M* test as part of the discriminant analysis. As the test is highly sensitive, an alpha level of .001 should be used.

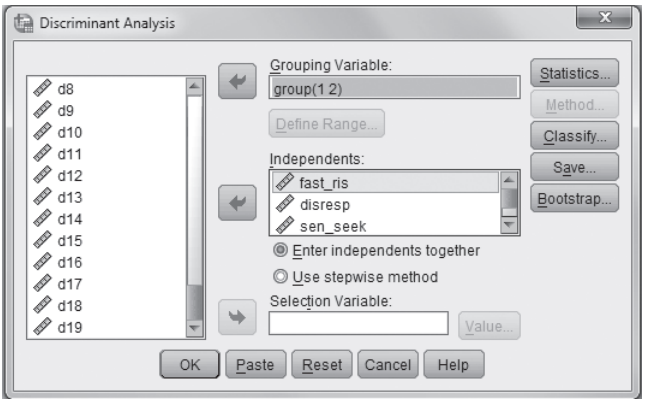
15.4.3 Two-Group Discriminant Analysis

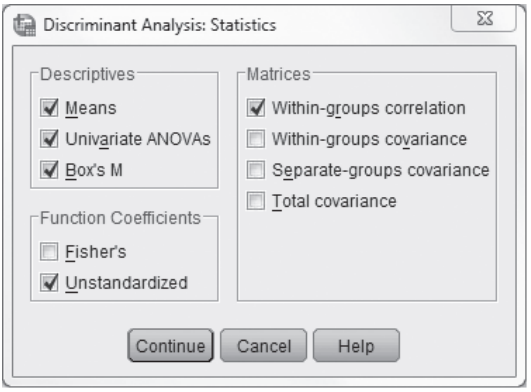
15.4.3.1 Windows Method

1. From the menu bar, click **Analyze**, then **Classify**, and then **Discriminant**. The following **Discriminant Analysis** window will open.

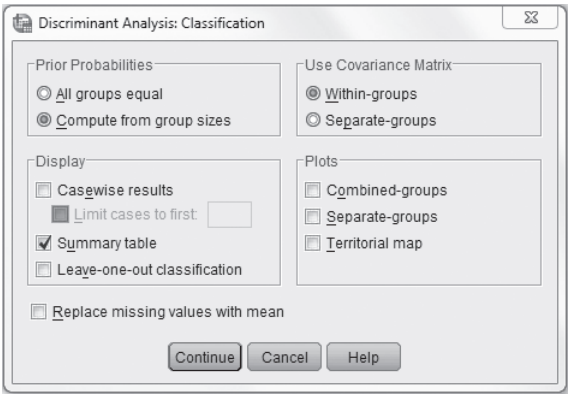


2. Click (highlight) the **GROUP** variable and then click ▶ to transfer this variable to the **Grouping Variable:** field. Click Define Range... to define the range of the **GROUP** variable (Minimum = 1, Maximum = 2). Click (highlight) the variables **FAS_RIS**, **DISRESP**, **SEN_SEEK**, and **DANGER** and then click ▶ to transfer these variables to the **Independent(s):** field. Select **Enter independents together** as the method of entry for this set of independent variables.
3. Click Statistics... to open the **Discriminant Analysis: Statistics** window. Check the cells for **Means, Univariate ANOVAs, Box's M, Within-groups correlation,** and **Unstandardized.** Click Continue to return to the **Discriminant Analysis** window.





4. When the **Discriminant Analysis** window opens, click Classify... . This will open the **Discriminant Analysis: Classification** window below. In the **Prior Probabilities** field check the **Compute from group sizes** cell. In the **Display** field, check the **Summary table** cell. In the **Use Covariance Matrix** field, check the **Within-groups** cell.



Click Continue to return to the **Discriminant Analysis** window.

5. When the **Discriminant Analysis** window opens, click OK to complete the analysis. See Table 15.2 for the results.

15.4.3.2 SPSS Syntax Method

```
DISCRIMINANT
/GROUPS = GROUP(1 2)
/VARIABLES = FAST_RIS DISRESP SEN_SEEK DANGER
/ANALYSIS ALL
/PRIORS SIZE
/STATISTICS = MEAN STDDEV UNIVF BOXM RAW CORR TABLE
/CLASSIFY = NONMISSING POOLED.
```

15.4.3.3 SPSS Output

TABLE 15.2

Discriminant Analysis Output

Discriminant

Analysis Case Processing Summary

Unweighted Cases		***N***	**Percent**
Valid		380	100.0
Excluded	Missing or out-of-range group codes	0	.0
	At least one missing discriminating variable	0	.0
	Both missing or out-of-range group codes and at least one missing discriminating variable	0	.0
	Total	0	.0
Total		380	100.0

Group Statistics

				Valid N (listwise)	
Group		**Mean**	**Std. Deviation**	**Unweighted**	**Weighted**
Accident	fast_ris	3.6912	1.07884	170	170.000
	disresp	3.3696	.87396	170	170.000
	sen_seek	3.5559	1.18966	170	170.000
	Danger	2.7484	.87958	170	170.000
No accident	fast_ris	3.1297	1.05873	210	210.000
	disresp	3.1722	1.04413	210	210.000
	sen_seek	3.1927	1.17291	210	210.000
	Danger	3.2082	.86365	210	210.000
Total	fast_ris	3.3809	1.10239	380	380.000
	disresp	3.2605	.97542	380	380.000
	sen_seek	3.3552	1.19265	380	380.000
	danger	3.0025	.89929	380	380.000

(*Continued*)

TABLE 15.2 (*Continued*)

Discriminant Analysis Output

Discriminant

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
fast_ris	.936	25.979	1	378	.000
disresp	.990	3.876	1	378	.050
sen_seek	.977	8.892	1	378	.003
danger	.935	26.198	1	378	.000

Pooled Within-Groups Matrices

		fast_ris	disresp	sen_seek	danger
Correlation	fast_ris	1.000	.627	.587	−.555
	disresp	.627	1.000	.584	−.440
	sen_seek	.587	.584	1.000	−.460
	danger	−.555	−.440	−.460	1.000

Analysis 1

Box's Test of Equality of Covariance Matrices

Log Determinants

Group	Rank	Log Determinant
Accident	4	−1.483
No accident	4	−1.337
Pooled within-groups	4	−1.334

The ranks and natural logarithms of determinants printed are those of the group covariance matrices.

Test Results

Group	Rank	Log Determinant
Box's M		25.685
F	Approx.	2.539
	df1	10
	df2	619229.996
	Sig.	.005

Tests null hypothesis of equal population covariance matrices.

Summary of Canonical Discriminant Functions

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative%	Canonical Correlation
1	.099[a]	100.0	100.0	.300

[a] First 1 canonical discriminant functions were used in the analysis.

TABLE 15.2 (*Continued*)

Discriminant Analysis Output

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	d*f*	Sig.
1	.910	35.572	4	.000

Standardized Canonical Discriminant Function Coefficients

	Function 1
fast_ris	−.761
disresp	.426
sen_seek	−.016
danger	.594

Structure Matrix

	Function 1
danger	.836
fast_ris	−.832
sen_seek	−.487
disresp	321

Polled within-groups correlations between discriminating variables and standardization canonical discriminating functions.
Variables ordered by absolute size of correlation within function.

Canonical Discriminant Function Coefficients

	Function 1
fast_ris	−.713
disresp	.439
sen_seek	−.013
danger	.682
(Constant)	−1.024

Unstandardized coefficients

Functions at Group Centroids

Group	Function 1
Accident	−.349
No accident	.283

Unstandardized canonical discriminant functions evaluated at group means

(*Continued*)

TABLE 15.2 (*Continued*)

Discriminant Analysis Output

Classification Statistics

Classification Processing Summary

Processed		380
Excluded	Missing or out-of-range group codes	0
	At least one missing discriminating variable	0
Used in Output		380

Prior Probabilities for Groups

		Cases Used in Analysis	
Group	**Prior**	**Unweighted**	**Weighted**
Accident	.447	170	170.000
No accident	.553	210	210.000
Total	1.000	380	380.000

Classification Results[a]

			Predicted Group Membership		
		Group	**Accident**	**No Accident**	**Total**
Original	Count	Accident	84	86	170
		No accident	51	159	210
	%	Accident	49.4	50.6	100.0
		No accident	24.3	75.7	100.0

[a] 63.9% of original grouped cases correctly classified.

15.4.3.4 Results and Interpretation

15.4.3.4.1 Test of Homogeneity of Variance-Covariance Matrices

This assumption is tested with Box's *M* test and the results are presented under the heading **Box's Test of Equality of Covariance Matrices**. The results indicate that the Box's *M* value of 25.685 ($F = 2.539$) is associated with an alpha level of .005. As mentioned earlier, Box's *M* is highly sensitive to factors other than just covariance differences (e.g., normality of the variables and large sample size). As such, an alpha level of .001 is recommended. On the basis of this alpha level, the computed level of .005 is not significant ($p > .001$). Thus, the assumption of equality of covariance matrices has not been violated.

15.4.3.4.2 Group Means of the Independent Variables

The **Group Statistics** table presents group (accident versus no accident) means and standard deviations for each of the four independent variables. The **Tests of Equality of Group Means** tables presents the Wilks' lambda and univariate *F* test used to assess differences between the mean scores of the four independent variables for the two groups of drivers. Significant group differences are found for all four independent variables ($p \leq .05$). Thus, all four independent variables are potentially important for discriminating between drivers who had been involved in accidents and those who had not.

15.4.3.4.3 Canonical Discriminant Function

A major purpose of discriminant functions is to obtain the combinations of predictor variables that maximally separate various groups from each other. As there are two groups of drivers in the present analysis, MDA produces one discriminant function (number of groups − 1). The **Wilks' Lambda** table presented under the heading **Summary of Canonical Discriminant Functions** shows that the discriminant function is highly significant by the chi-square test, $\chi^2\,(df = 4) = 35.57, p < .001$. The **Eigenvalues** table also displays a canonical correlation of .30. Squaring this correlation ($.30^2$) yields a value of .09. Thus, 9% of the variance in the dependent variable (GROUP) can be accounted for by this model that includes the four predictor variables.

15.4.3.4.4 Standardized Canonical Discriminant Function Coefficients

These discriminant coefficients/weights are presented in the **Standardized Canonical Discriminant Function Coefficients** table. Each coefficient represents the relative contribution of its associated predictor variable to the discriminant function such that predictor variables with relatively larger coefficients contribute more to the discriminating power of the function than do variables with smaller coefficients. Thus, the function coefficients in this table show that the predictor variables FAST_RIS, DISRESP, and DANGER contribute the most to the discriminant function in discriminating between the two groups of drivers. It should be noted that these coefficients are interpreted in the same way as Beta weights are interpreted in multiple regression analysis and are used primarily to calculate predicted group membership. Nevertheless, given their common problems with Beta weights in regression analysis (e.g., multicollinearity), they are less preferred for interpreting the results of discriminant analysis than the discriminant loadings presented in the structure matrix (Hair et al., 1995).

15.4.3.4.5 Structure Matrix

The **Structure Matrix** table presents the correlation (discriminant loading) of each predictor variable with the discriminant function. These discriminant loadings are similar to factor loadings in factor analysis and are ordered

in descending magnitude. The interpretation of these loadings in discriminating between the two groups of drivers (dependent variable) follows.

The researcher is interested in knowing which predictor variables are substantive discriminators of the GROUP dependent variable. In general, any variables with loadings of ≥±.30 are considered substantive. Inspection of the structure matrix shows that all four predictor variables exceed this standard. In interpreting the individual variables that have both practical and statistical significance the researcher needs to (1) identify the variables that are statistically significant and (2) understand what the different group means for each predictor variable indicate. Table 15.3 presents the discriminant loadings for the four predictor variables, their group means, and the univariate *F* test of significance. Please note that this information can be found in the discriminant analysis output in Table 15.2.

From Table 15.3, the researcher can use the discriminant loadings and the *F* values to determine which of the four predictor variables are significant discriminators of the two groups of drivers, as well as their rankings. Thus, the results show that all four predictor variables are statistically significant discriminators ($p \leq .05$). Moreover, of the four variables in the function, the perception of DANGER discriminates the most and DISRESPECT for traffic laws discriminates the least. Thus, in terms of rankings, the perception of danger, the desire to drive fast/take risks, and the need for sensation-seeking are the strongest discriminators, with disrespect for traffic laws being the weakest.

Inspection of the group means clearly shows that the two groups of drivers (accident versus no accident) differed significantly in the way they assessed danger, their desire to drive fast and to take risks, their need for sensation seeking, and their level of disrespect for traffic laws. More specifically, those drivers who had been involved in car accidents, compared to those who had not been involved in accidents, tended to have a lower perception of danger,

TABLE 15.3

Discriminant Loadings, Group Means, and Univariate *F* Test

Predictor Variables	Discriminant Loadings	Group Means		*F* Ratio	*p*
		Accident	No Accident		
DANGER	.836	2.748	3.208	26.198	<.001
FAST_RIS	−.832	3.691	3.129	25.979	<.001
SEN_SEEK	−.487	3.555	3.192	8.892	<.01
DISRESP	−.321	3.369	3.172	3.876	≤.05

a higher desire to drive fast/take risks, a higher need for sensation-seeking, and a higher disrespect for traffic laws.

15.4.3.4.6 Group Centroids

Group centroids are standardized joint means based on the linear combination of the predictor variables for each group, and are used to interpret group differences. The group centroids are presented in the **Functions at Group Centroids** table. It can be seen from the table that the group centroid for the no-accident group is .283, whereas the group centroid for the accident group is −.349. Therefore, it can be determined that the accident group drivers differ from the no-accident group drivers in their lower perception of danger, higher desire to drive fast/take risks, higher need for sensation-seeking, and higher disrespect for traffic laws.

15.4.3.4.7 Classification Matrix

An important step in discriminant analysis is to assess the predictive accuracy of the discriminant function. This is required because the statistical tests employed to assess the significance of the discriminant function do not inform the researcher as to how well the function predicts. For example, if the sample sizes are large, then the group means (centroids) could be similar and still achieve statistical significance. Thus, the level of significance (e.g., $p < .01$) can be a very poor indication of the function's ability to discriminate between the two groups (Hair et al., 1995). To determine the predictive power of a discriminant function, the researcher must examine the classification matrix to determine what proportion of cases is correctly classified and what proportion of cases is misclassified.

Examination of the **Classification Results** table shows that 63.9% of the cases were correctly classified. That is, the obtained discriminant function correctly predicted group membership for 63.9% of the cases. Inspection of the classification results shows that there was 75.7% correct classification (24.3% misclassification) for the no-accident group but just 49.4% correct classification (50.6% misclassification) for the accident group. Thus, the two groups of drivers appear to be quite different with the non-accident group more likely to be correctly classified by the discriminant function than the accident group.

15.4.3.4.8 Classification Accuracy

The final step in determining the predictive ability of a discriminant function is to test its classification accuracy. That is, is the discriminatory power of the classification statistically better than chance (50% assignment)? The researcher can use Press's *Q* statistic to compare with the chi-square critical

value of 6.63 with 1 degree of freedom ($p < .01$). If Q exceeds this critical value, the classification can be regarded as significantly better than chance.

$$\text{Press's } Q = \frac{\left[N - (nK)\right]^2}{N(K-1)}$$

where
N = total sample size
n = number of observations correctly classified
K = number of groups
Thus,

$$\text{Press's } Q = \frac{\left[380 - (243 \times 2)\right]^2}{380(2-1)} = 29.57$$

Given that Press's $Q = 29.57 > 6.63$, it can be concluded that the classification results exceed the classification accuracy expected by chance at a statistically significant level ($p < .01$).

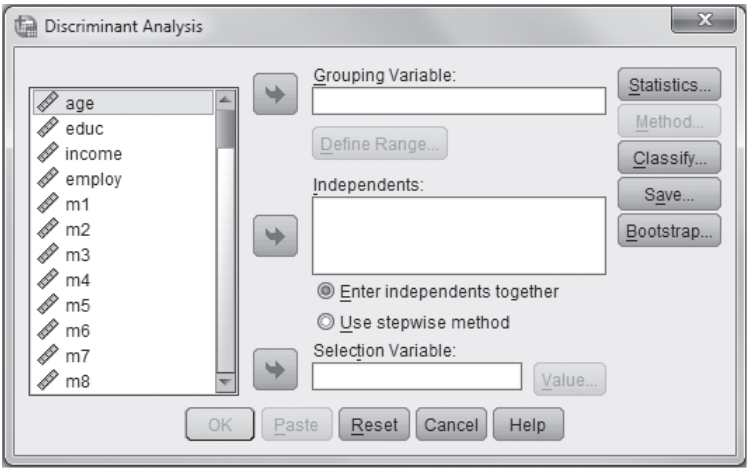
15.5 Example 2: Three-Group Discriminant Analysis

This example extends on the previous two-group analysis example by using the same four factors of *driving fast/risk taking*, *disrespect for traffic laws*, *sensation-seeking*, *and danger assessment* to discriminate between three groups of drivers (**GROUP_1**): those who had been involved in a motor vehicle accident and charged by the police (coded 1), those who had been involved in a motor vehicle accident but were not charged by the police (coded 2), and those who had not been involved in a motor vehicle accident (coded 3). The same data set (**DRIVE_1**) will be used for this example. See Section 15.4.1 for the SPSS coding of the four predictor variables (**FAST_RIS**, **DISRESP**, **SEN_SEEK**, **DANGER**) and the dependent variable **GROUP_1**.

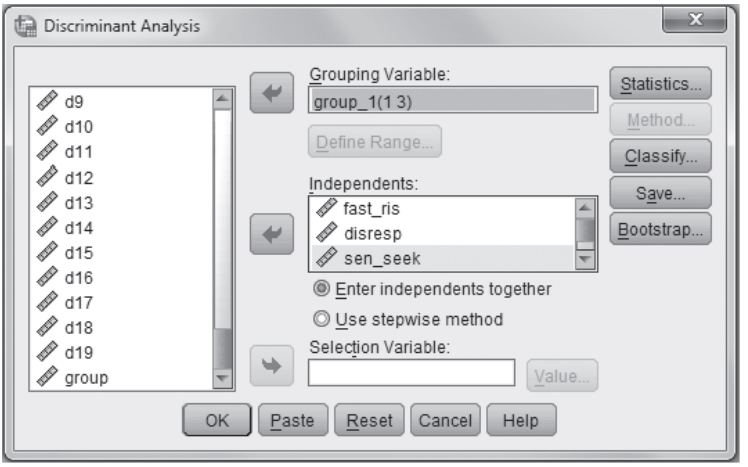
15.5.1 Three-Group Discriminant Analysis

15.5.1.1 Windows Method

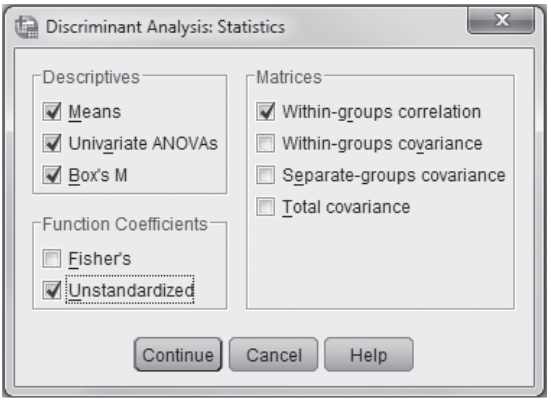
1. From the menu bar, click **Analyze**, then **Classify**, and then **Discriminant**. The following **Discriminant Analysis** window will open.
2. Click (highlight) the **GROUP_1** variable and then click ▸ to transfer this variable to the **Grouping Variable:** field. Click Define Range... to open



the **Discriminant Analysis: Define Range** window. To define the range of the grouping variable **GROUP_1,** type **1** in the **Minimum** cell and **3** in the **Maximum** cell, and then click Continue. Click (highlight) the variables **FAS_RIS**, **DISRESP**, **SEN_SEEK**, and **DANGER** and then click ▶ to transfer these variables to the **Independent(s):** field. Select **Enter independents together** as the method of entry for this set of independent variables.

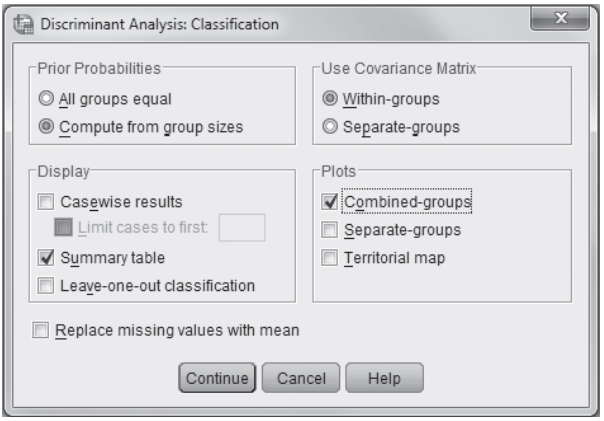


3. Click Statistics... to open the **Discriminant Analysis: Statistics** window. Check the **Means**, **Univariate ANOVAs**, **Box's M**, **Within-groups correlation**, and **Unstandardized** cells.



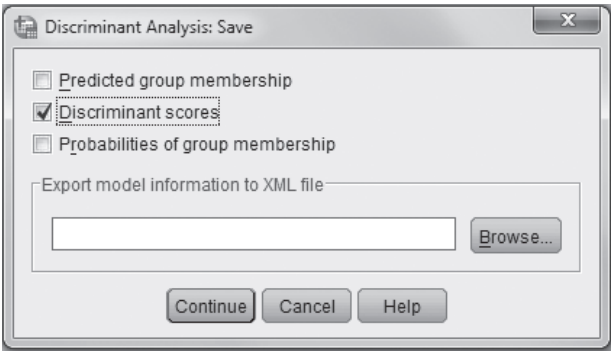
Click Continue to return to the **Discriminant Analysis** window.

4. When the **Discriminant Analysis** window opens, click Classify... . This will open the **Discriminant Analysis: Classification** window below. In the **Prior Probabilities** field check the **Compute from group sizes** cell. In the **Display** field, check the **Summary table** cell. In the **Use Covariance Matrix** field, check the **Within-groups** cell. In the **Plots** field, check the **Combined-groups** cell.



Click Continue to return to the **Discriminant Analysis** window.

5. When the **Discriminant Analysis** window opens, click Save... to open the **Discriminant Analysis: Save** window. In this window, check the **Discriminant scores** cell. This will create and save two new variables **Dis1_1** and **Dis2_1** at the end of the data set. These are the drivers' mean scores on the two discriminant functions. Pairwise comparisons on the means of the discriminant functions can be used to check whether the discriminant functions provide differences among all of the three groups of drivers.



Click Continue to return to the **Discriminant Analysis** window.

6. When the **Discriminant Analysis** window opens, click OK to complete the analysis. See Table 15.4 for the results.

15.5.1.2 SPSS Syntax Method

```
DISCRIMINANT
/GROUPS = GROUP_1(1 3)
/VARIABLES = FAST_RIS DISRESP SEN_SEEK DANGER
/ANALYSIS ALL
/SAVE = SCORES
/PRIORS SIZE
/STATISTICS = MEAN STDDEV UNIVF BOXM COEFF RAW CORR TABLE
/PLOT = COMBINED
/CLASSIFY = NONMISSING POOLED.
```

15.5.1.3 SPSS Output

TABLE 15.4

Discriminant Analysis Output

Discriminant			
Analysis Case Processing Summary			
Unweighted Cases		***N***	**Percent**
Valid		380	100.0
Excluded	Missing or out-of-range group codes	0	.0
	At least one missing discriminating variable	0	.0
	Both missing or out-of-range group codes and at least one missing discriminating variable	0	.0
	Total	0	.0
Total		380	100.0

(*Continued*)

TABLE 15.4 (*Continued*)

Discriminant Analysis Output

Discriminant

Group Statistics

Group 1		Mean	Std. Deviation	Valid N (listwise) Unweighted	Weighted
Accident-charged	fast_ris	3.8938	1.11293	102	102.000
	disresp	3.5458	.89926	102	102.000
	sen_seek	3.7113	1.22647	102	102.000
	danger	2.5545	.87585	102	102.000
Accident-no charge	fast_ris	3.3873	.95520	68	68.000
	disresp	3.1054	.76777	68	68.000
	sen_seek	3.3227	1.10018	68	68.000
	danger	3.0393	.80743	68	68.000
No accident	fast_ris	3.1297	1.05873	210	210.000
	disresp	3.1722	1.04413	210	210.000
	sen_seek	3.1927	1.17291	210	210.000
	danger	3.2082	.86365	210	210.000
Total	fast_ris	3.3809	1.10239	380	380.000
	disresp	3.2605	.97542	380	380.000
	sen_seek	3.3552	1.19265	380	380.000
	danger	3.0025	.89929	380	380.000

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
fast_ris	.913	17.970	2	377	.000
disresp	.968	6.250	2	377	.002
sen_seek	.966	6.718	2	377	.001
danger	.904	20.041	2	377	.000

Pooled Within-Groups Matrices

		fast_ris	disresp	sen_seek	danger
Correlation	fast_ris	1.000	.619	.581	−.542
	disresp	.619	1.000	.578	−.425
	sen_seek	.581	.578	1.000	−.451
	danger	−.542	−.425	−.451	1.000

TABLE 15.4 (*Continued*)

Discriminant Analysis Output

Discriminant

Analysis 1

Box's Test of Equality of Covariance Matrices

Log Determinants

Group 1	Rank	Log Determinant
Accident-charged	4	−1.668
Accident-no charge	4	−1.587
No accident	4	−1.337
Pooled within-groups	4	−1.365

The ranks and natural logarithms of determinants printed are those of the group covariance matrices.

Test Results

Box's M		39.485
F	Approx.	1.936
	df1	20
	df2	167036.989
	Sig.	.007

Tests null hypothesis of equal population covariance matrices.

Summary of Canonical Discriminant Functions

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	.134[a]	92.9	92.9	.344
2	.010[a]	7.1	100.0	.100

[a] First 2 canonical discriminant functions were used in the analysis.

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1 through 2	.873	51.079	8	.000
2	.990	3.807	3	.283

Standardized Canonical Discriminant Function Coefficients

	Function	
	1	2
fast_ris	−.635	−.763
disresp	.198	1.293
sen_seek	.030	−.250
danger	.641	−.214

(*Continued*)

TABLE 15.4 (*Continued*)

Discriminant Analysis Output

Discriminant

Structure Matrix

	Function	
	1	**2**
danger	.888*	−.237
fast_ris	−.843*	.008
sen_seek	−.514*	.150
disresp	−.450	.768*

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions. Variables ordered by absolute size of correlation within function.

* Largest absolute correlation between each variable and any discriminant function

Canonical Discriminant Function Coefficients

	Function	
	1	**2**
fast_ris	−.601	−.723
disresp	.206	1.344
sen_seek	.025	−.212
danger	.748	−.249
(Constant)	−.967	−.477

Unstandardized coefficients

Functions at Group Centroids

	Function	
group 1	**1**	**2**
Accident-charged	−.576	.049
Accident-no charge	−.009	−.215
No accident	.283	.046

Unstandardized canonical discriminant functions evaluated at group means

Classification Statistics

Classification Processing Summary

Processed		380
Excluded	Missing or out-of-range group codes	0
	At least one missing discriminating variable	0
Used in Output		380

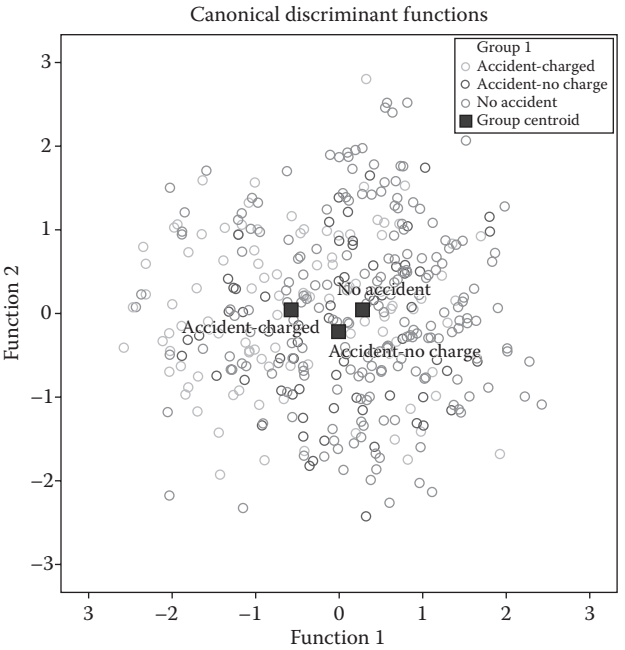
TABLE 15.4 (*Continued*)

Discriminant Analysis Output

Discriminant

Prior Probabilities for Groups

		Cases Used in Analysis	
Group 1	**Prior**	**Unweighted**	**Weighted**
Accident-charged	.268	102	102.000
Accident-no charge	.179	68	68.000
No accident	.553	210	210.000
Total	1.000	380	380.000



Classification Results[a]

Predicted Group Membership

		Group 1	Accident-Charged	Accident-No Charge	No Accident	Total
Original	Count	Accident-charged	42	0	60	102
		Accident-no charge	11	0	57	68
		No accident	27	0	183	210

(*Continued*)

TABLE 15.4 (*Continued*)

Discriminant Analysis Output

Classification Results[a]					
		Predicted Group Membership			
	Group 1	Accident-Charged	Accident-No Charge	No Accident	Total
%	Accident-charged	41.2	.0	58.8	100.0
	Accident-no charge	16.2	.0	83.8	100.0
	No accident	12.9	.0	87.1	100.0

[a] 59.2% of original grouped cases correctly classified.

15.5.1.4 Results and Interpretation

15.5.1.4.1 Test of Homogeneity of Variance-Covariance Matrices

This assumption is tested with Box's M test and the results are presented under the heading **Box's Test of Equality of Covariance Matrices**. The results indicate that the Box's M value of 39.485 is associated with an alpha level of .007. As mentioned earlier, Box's M is highly sensitive to factors other than just covariance differences (e.g., normality of the variables and large sample size). As such, an alpha level of .001 is recommended. On the basis of this alpha level, the computed level of .007 is not significant ($p > .001$). Thus, the assumption of equality of covariance matrices has not been violated.

15.5.1.4.2 Group Means of the Independent Variables

The **Group Statistics** table presents group (accident-charged, accident-no charge, no accident) means and standard deviations for each of the four independent variables. The **Tests of Equality of Group Means** table presents Wilks' lambda and the univariate *F* test used to assess differences between the mean scores of the four independent variables for the three groups of drivers. Significant group differences are found for all four independent variables ($p \leq .002$). Thus, all four independent variables are potentially important for discriminating between the three groups of "accident-charged" drivers, "accident-no charge" drivers, and "no accident" drivers.

15.5.1.4.3 Summary of Canonical Discriminant Functions

A major purpose of discriminant functions is to obtain the combinations of predictor variables that maximally separate various groups from each other. As there are three groups of drivers in the present analysis, MDA produces two discriminant functions (number of groups − 1). The first discriminant function explains the largest amount of variance

(difference) between the three driver groups. The second discriminant function (orthogonal and independent of the first) explains the largest amount of the remaining (residual) variance after the variance for the first function is removed. The **Eigenvalues** table presented under the heading **Summary of Canonical Discriminant Functions** shows that the first function accounts for 92.9% of the variance explained by the two functions; the second function accounts for the remaining 7.1% of the variance. The **Wilks' Lambda** table shows the test of significance of the two functions. **Function 1 through 2** indicates the significance of both functions (with no functions removed). The associated chi-square value is highly significant, χ^2 (df = 8) = 51.08, $p < .001$, and indicates that the two functions together discriminate between the three groups of drivers well. After the first function is extracted, the chi-square is recalculated. The results show that this function is not significant, χ^2 (df = 3) = 3.81, $p > .05$, indicating that no significant differences are present in the remaining variance.

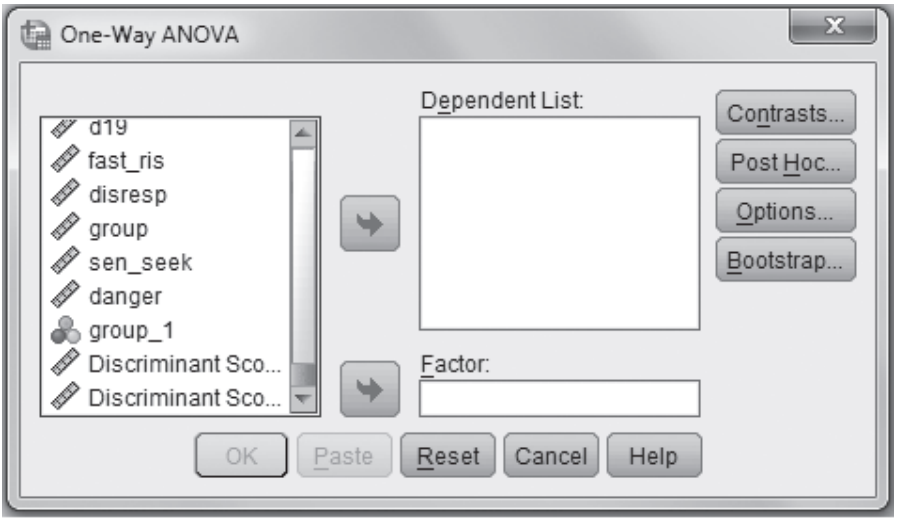
The **Eigenvalues** table also displays a canonical correlation of .344 for the first function. Squaring this correlation ($.344^2$) yields a value of .1183. Thus, 11.83% of the variance in the dependent variable (GROUP_1) can be accounted for by the first function. The second function has a canonical correlation of .10. Squaring this correlation ($.10^2$) yields a value of .01. Thus, the second function explains 1% of the remaining variance of 88.17% (100–11.83). Therefore, the total variance accounted for by both functions is .1183 + (.01 × .8817) =.1271, or 12.71% of the total variance in the GROUP_1 dependent variable.

15.5.1.5 Evaluating Group Differences

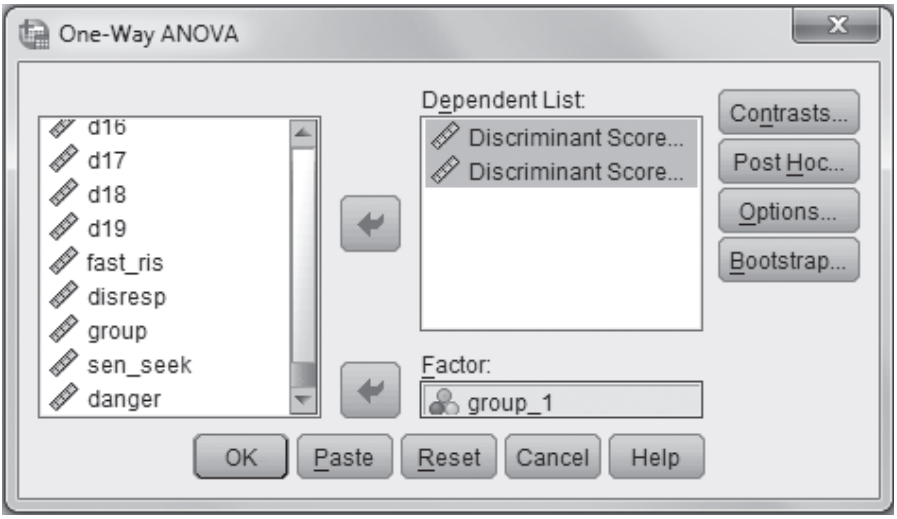
While the two discriminant functions together are statistically significant, it is important to understand whether the discriminant functions separately provide differences among all of the three groups of drivers. It is possible to have statistically significant functions, but have at least one pair of groups not be statistically significant (i.e., not discriminated between) (Hair et al., 1995). This problem frequently arises as a direct function of the number of groups included in the analysis (i.e., as the number of groups increases or decreases). The following procedure will test for group differences between each pair of groups (accident-charged drivers versus accident-no charge drivers; accident-charged drivers versus no-accident drivers; accident-no charge drivers versus no-accident drivers).

15.5.1.6 Windows Method

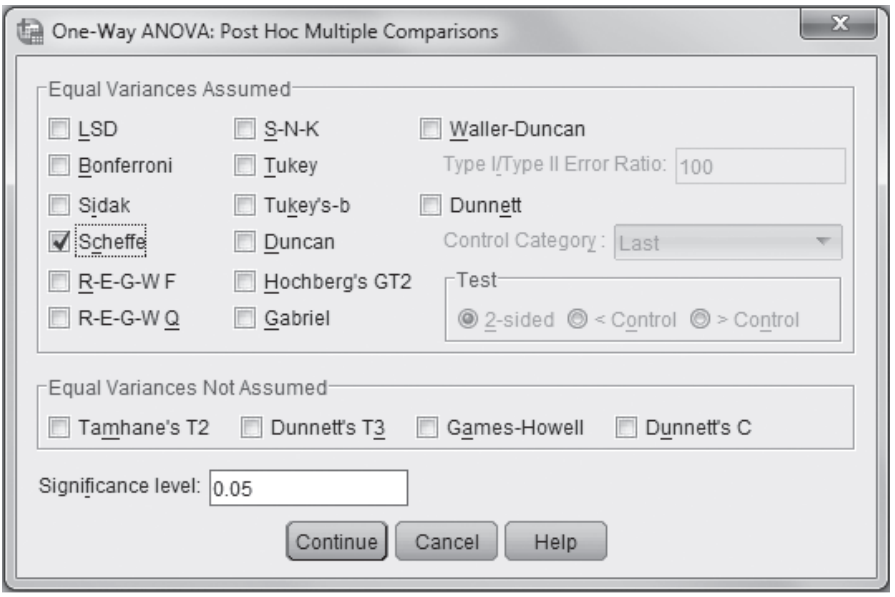
1. From the menu bar, click **Analyze**, then **Compare Means**, and then **One-Way ANOVA**. The following **One-Way ANOVA** window will open.



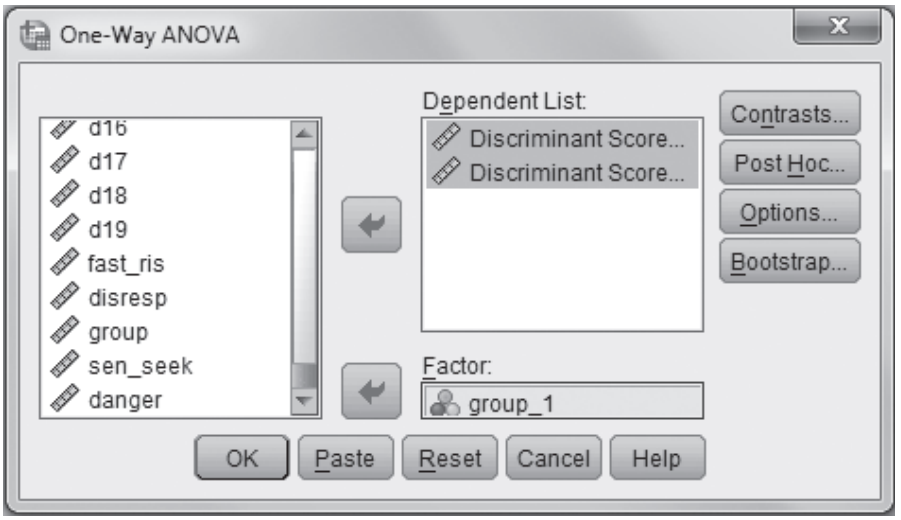
2. Transfer the variables **Dis1_1** and **Dis2_1** to the **Dependent List:** field by clicking (highlight) the variables and then clicking . Transfer the **GROUP_1** variable to the **Factor:** field by clicking (highlight) the variable and then clicking .



3. To obtain pairwise comparisons between the three groups of drivers, the researcher needs to perform a post hoc comparison test. Click Post Hoc... to achieve this. When the following **One-Way ANOVA: Post Hoc Multiple Comparisons** window opens, check the **Scheffe** cell to run the Scheffé post hoc test. Next, click Continue.



4. When the following **One-Way ANOVA** window opens, run the analysis by clicking OK. See Table 15.5 for the results.



15.5.1.7 SPSS Syntax Method

```
ONEWAY DIS1 _ 1 DIS2 _ 1 BY GROUP _ 1
/MISSING ANALYSIS
/POSTHOC = SCHEFFE ALPHA(0.05).
```

TABLE 15.5

Multiple Comparisons Output

Multiple Comparisons							
Scheffe							
						95% Confidence Interval	
Dependent Variable	(I) Group 1	(J) Group 1	Mean Difference (I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
Discriminant Scores from Function 1 for Analysis 1	Accident-charged	Accident-no charge	−.56656870*	.15655607	.002	−.9513059	−.1818315
		No accident	−.85825284*	.12068897	.000	−1.1548465	−.5616592
	Accident-no charge	Accident-charged	−.56656870*	.15655607	.002	.1818315	.9513059
		No accident	−.29168414	.13952701	.114	−.6345724	.0512041
	No accident	Accident-charged	−.85825284*	.12068897	.000	.5616592	1.1548465
		Accident-no charge	.29168414	.13952701	.114	−.0512041	.6345724
Discriminant Scores from Function 2 for Analysis 1	Accident-charged	Accident-no charge	−26415223	.15655607	.242	−.1205850	.6488895
		No accident	.00275912	.12068897	1.000	−.2938345	.2993528
	Accident-no charge	Accident-charged	−.26415223	.15655607	.242	−.6488895	.1205850
		No accident	−.26139311	.13952701	.174	−.6042813	.0814951
	No accident	Accident-charged	−.00275912	.12068897	1.000	−.2993528	.2938345
		Accident-no charge	.26139311	.13952701	.174	−.0814951	.6042813

* The mean difference is significant at the 0.05 level.

The multiple comparisons show that the "accident-charged" group differs significantly from the other two groups of drivers ("accident no-charge" and "no accident") on the first discriminant function. On the second discriminant function, the three groups of drivers do not differ significantly from each other. These findings denote that while the discriminant functions created separation in an overall sense, it is merely the first function that provided information regarding which of the specific groups of drivers the function discriminated/not discriminated between.

15.5.1.7.1 Standardized Canonical Discriminant Function Coefficients

These discriminant coefficients/weights are presented in the **Standardized Canonical Discriminant Function Coefficients** table. As shown in the previous two-group model, each coefficient represents the relative contribution of its associated predictor variable to the discriminant function such that predictor variables with relatively larger coefficients contribute more to the discriminating power of the function than do variables with smaller coefficients. Thus, the first function coefficients in this table show that the predictor variables of FAST_RIS and DANGER contribute the most to the discriminant function in discriminating between the three groups of drivers. After extraction of the first function, the second function coefficients show that the predictor variable FAST_RIS contributes the most to the discriminant function in discriminating between the three groups of drivers.

15.5.1.7.2 Structure Matrix

The **Structure Matrix** table presents the correlation (discriminant loading) of each predictor variable with the discriminant function. The interpretation of these loadings in discriminating between the three groups of drivers (dependent variable) is as follows.

As with the previous two-group example, the researcher is interested in knowing which predictor variables are substantive discriminators of the GROUP_1 dependent variable. In general, any variables with loadings of $\geq \pm .30$ are considered substantive. Inspection of **Function 1** shows that all four predictor variables exceed this criterion. **Function 2** shows that the DISRESP variable exceeds this criterion. In interpreting the individual variables that have both practical and statistical significance the researcher needs to (1) identify the variables that are statistically significant and (2) understand what the different group means for each predictor variable indicate. Table 15.6 presents the discriminant loadings for the four predictor variables, their group means, and the univariate *F* test of significance. Please note that this information can be found in the discriminant analysis output in Table 15.4.

From Table 15.6, the researcher can use the discriminant loadings and the *F* values to determine which of the four predictor variables are significant discriminators of the three groups of drivers, as well as their rankings. Thus, the results show that all four predictor variables are statistically significant discriminators ($p < .001$). Moreover, of the four variables in the first function,

TABLE 15.6

Discriminant Loadings, Group Means, and Univariate *F* Test

Predictor Variables	Discriminant Loadings		Group Means				
	Func 1	Func 2	Accident-Charged	Accident-No Charge	No Accident	*F* Ratio	*p*
DANGER	.888	−.237	2.55	3.04	3.21	20.04	<.001
FAST_RIS	−.843	.008	3.89	3.39	3.13	17.97	<.001
DISRESP	−.514	.768	3.54	3.10	3.17	6.25	<.01
SEN_SEEK	−.450	.150	3.71	3.32	3.19	6.72	<.01

the perceptions of DANGER and FAST_RIS (the desire to drive fast and to take risks) discriminate the most and DISRESP (disrespect) for traffic laws and SEN_SEEK (sensation seeking) discriminate the least. Thus, in terms of rankings, the perception of danger, the desire to drive fast/take risks, and disrespect for traffic laws are the strongest discriminators, with the need for sensation seeking being the weakest.

Inspection of the group means clearly shows that the three groups of drivers (accident-charged, accident-no charge, no accident) differed significantly in the way they assessed danger, their desire to drive fast and to take risks, their level of disrespect for traffic laws, and their need for sensation seeking. More specifically, those drivers who had been involved in car accidents and were charged, compared to those who had been involved in accidents but were not charged, and those who had not been involved in accidents, tended to have a lower perception of danger, a higher desire to drive fast/take risks, a higher disrespect for traffic laws, and a higher need for sensation-seeking.

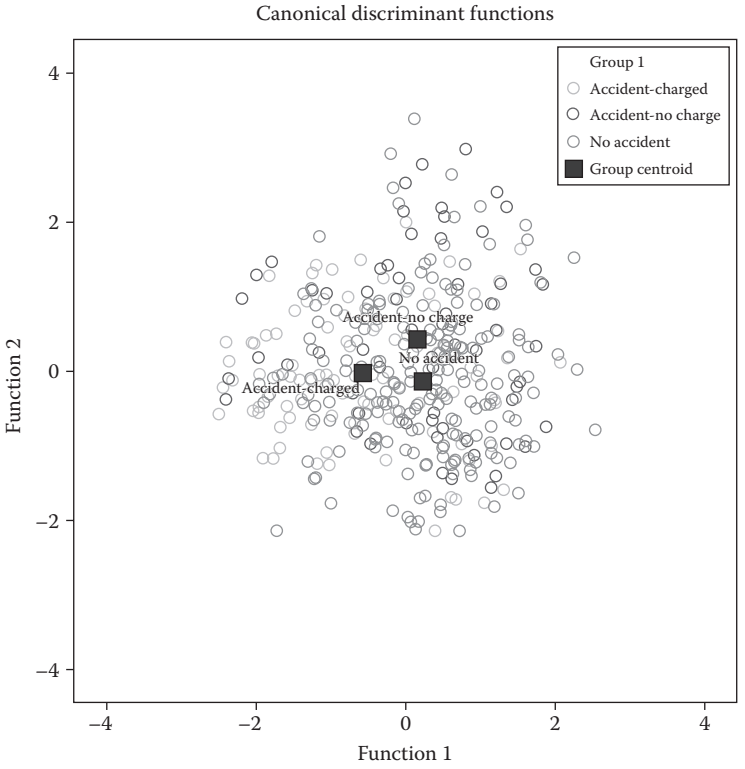
Moreover, by identifying the largest loadings for each discriminant function, the researcher gains insight into how to label each function. Hence, the perceptions of DANGER and FAST_RIS have the highest loadings on the first function that suggest a label of "low perceived danger and risk-taking" as the function that discriminates between the three groups of drivers (accident-charged, accident-no charge, no accident). For the second function, the variable of DISRESP has the highest loading, which suggests that a label of "disrespect for traffic laws" best describes this function that discriminates between the three groups of drivers.

15.5.1.7.3 Group Centroids

Group centroids are standardized joint means based on the linear combination of the predictor variables for each group, and are used to interpret group differences. The group centroids are presented in the **Functions at Group Centroids** table. These group centroids can be used to determine between which of the three groups each function discriminates. The easiest approach is to look at the group centroids and see where the differences lie.

For example, function 1 shows that the centroid for the accident-charged group is –.576, for the accident no-charge group it is –.009, and for the no accident group it is .283. From this, it can be concluded that the primary source of differences for this function is between the "accident-no charge" and "no accident" groups versus the "accident-charged" group. Function 2 can also be used in this way but because it represents substantially less variance than function 1, caution must be taken in determining the impact of variables on the basis of their loadings on this function (Hair et al., 1995).

The analysis also produces a *combined-groups plot* showing the group centroids and the participants' scores. An examination of the map clearly indicates that the "accident-charged" group differs from the "accident no-charge" and "no accident" groups on function 1. The differences between the three groups are less clearly defined on function 2.



15.5.1.7.4 Classification Matrix

An important step in discriminant analysis is to assess the predictive accuracy of the discriminant functions. As with the previous two-group example, this is required because the statistical tests employed to assess the significance of the

discriminant functions do not inform the researcher as to how well the functions predict. For example, if the sample sizes are large, then the group means (centroids) could be similar and still achieve statistical significance. Thus, the level of significance (e.g., $p < .01$) can be a very poor indication of the functions' ability to discriminate between the three groups (Hair et al., 1995). To determine the predictive power of the discriminant functions, the researcher must examine the classification matrix to determine what proportion of cases is correctly classified and what proportion of cases is misclassified.

Examination of the **Classification Results** table shows that 59.2% of the cases were correctly classified. That is, the obtained discriminant functions correctly predicted group membership for 59.2% of the cases. Inspection of the classification results shows that there was (1) 41.2% correct classification (58.8% misclassification) for the accident-charged group, (2) 0% correct classification (100% misclassification) for the accident-no charge group, and (3) 87.1% correct classification (12.9% misclassification) for the no-accident group. Thus, the three groups of drivers appear to be quite different with the no-accident group more likely to be correctly classified by the discriminant functions than the accident-charged and accident-no charge groups.

15.5.1.7.5 Classification Accuracy

The final step in determining the predictive power of a discriminant function is to test its classification accuracy. That is, is the discriminatory power of the classification statistically better than chance? The researcher can use Press's Q statistic to compare with the chi-square critical value of 6.63 with 1 degree of freedom ($p < .01$). If Q exceeds this critical value, the classification can be regarded as significantly better than chance.

$$\text{Press's } Q = \frac{\left[N - (nK)\right]^2}{N(K-1)}$$

where

N = total sample size
n = number of observations correctly classified
K = number of groups
Thus,

$$\text{Press's } Q = \frac{(380 - [225 \times 3])^2}{380(3-1)} = 114.51$$

Given that Press's Q = 114.51 (>6.63), it can be concluded that the classification results exceed the classification accuracy expected by chance at a statistically significant level ($p < .01$).

16

Logistic Regression

16.1 Aim

Similar to discriminant analysis, logistic regression allows one to predict a discrete outcome, such as group membership, from a set of variables that may be continuous, discrete, dichotomous, or a mix of any of these. Generally, the dependent variable is dichotomous, such as male/female, smoker/nonsmoker or success/failure. While discriminant analysis is also used to predict group membership with only two groups, logistic regression is more flexible in that it has no assumptions about the distributions of the predictor variables—the predictor variables do not have to be normally distributed, linearly related, or of equal variance/covariance across the groups.

There are three primary uses of logistic regression:

1. *Prediction of group membership and outcome*. The goal is to correctly predict the category of the outcome of individual cases. Thus, the research question asked is whether an outcome can be predicted from a selected set of independent variables. For instance, in epidemiological studies, can the development of lung cancer be predicted from the incidence and duration of smoking as well as from demographic variables such as gender, age, and social and economic status (SES)?
2. *Logistic regression provides knowledge of the relationships and strengths among the variables.* The goal is to identify which independent variables predict the outcome, that is, increase or decrease the probability of the outcome or have no effect. For example, does inclusion of information about the incidence and duration of smoking improve prediction of lung cancer, and is a particular variable associated with an increase or decrease in the probability that a case has lung cancer? These parameter estimates (the coefficients of the predictors included in a model) can also be used to calculate and interpret the *odds ratio*. For instance, what are the odds that a person has lung cancer at age 65, given that he has smoked 10 packs a day for the past 30 years?

3. *Classification of cases*. The goal is to understand how reliable the logistic regression model is in classifying cases for whom the effect is known. For instance, how many people with or without lung cancer are diagnosed correctly? The researcher establishes a cut point of say .5, and then asks, for instance: How many people with lung cancer are correctly classified if everyone with a predicted probability of .5 or more is diagnosed as having lung cancer (Tabachnick and Fidell, 2001)?

16.2 Checklist of Requirements

- The dependent variable must be categorical (nominal) and should consist of two groups (e.g., male versus female; low versus high). The independent variables can be continuous, discrete, dichotomous, or a mixture of any of these.
- The size of the sample can affect model convergence. When there are too few cases relative to the number of predictor variables, logistic regression may produce extremely large parameter estimates and standard errors, and possibly, failure of convergence when combinations of discrete variables result in too many cells with no cases (Tabachnick and Fidell, 2001).

16.3 Assumptions

Although assumptions regarding the distribution of predictors are not required for logistic regression, multivariate normality and linearity among the predictors may enhance the power (Tabachnick and Fidell, 2001).

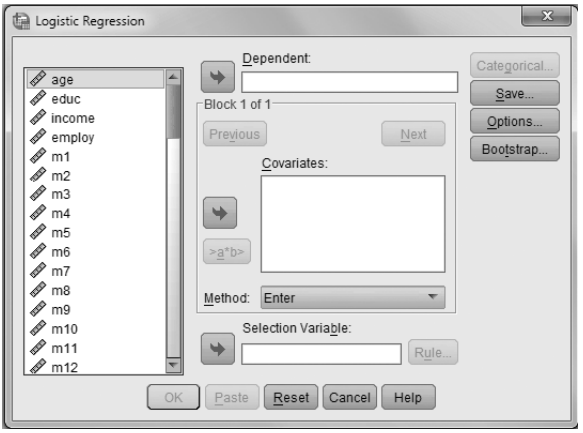
16.4 Example: Two-Group Logistic Regression

The following example is identical to the two-group discriminant analysis discussed in Chapter 15 (see Section 15.4), with logistic regression used this time to determine how the factors of *driving fast/risk taking, disrespect for traffic laws, sensation-seeking,* and *danger assessment* discriminate between young drivers who had been involved in a motor vehicle accident and those who had not. The analysis allows different steps in the logistic regression model. The difference between the steps is the predictors that are included. This is similar to blocking variables into groups and then entering them into the

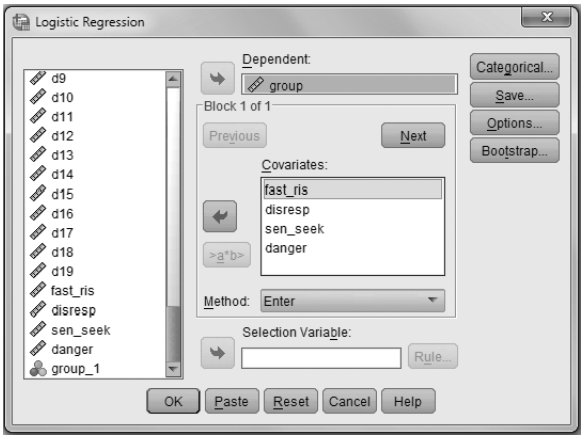
equation one group at a time. In the present case, there are no strong theoretical grounds for blocking the predictor variables into groups for entry. As such, all four predictors will be entered into the logistic regression model as a single block. The data set is **DRIVE_1**.

16.4.1 Windows Method

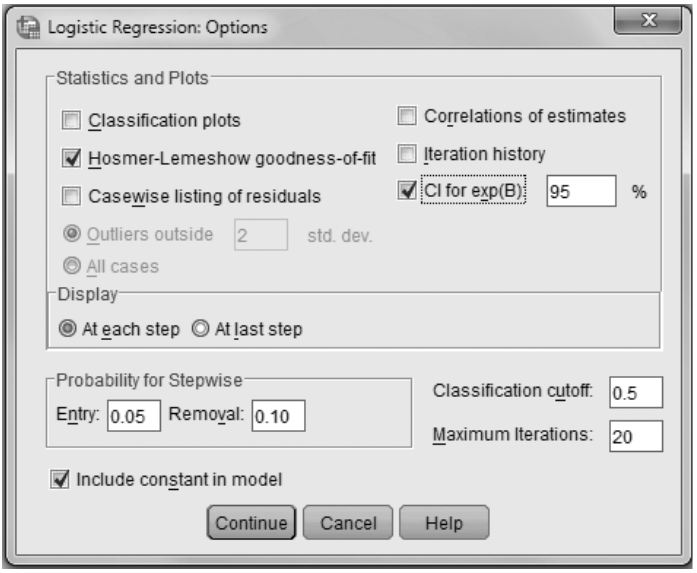
1. From the menu bar, click **Analyze**, then **Regression**, and then **Binary Logistic**. The following **Logistic Regression** window will open.



2. Click (highlight) the **GROUP** variable and then click ▶ to transfer this variable to the **Dependent:** field. Click (highlight) the variables **FAS_RIS**, **DISRESP**, **SEN_SEEK**, and **DANGER** and then click ▶ to transfer these variables to the **Covariates:** field. Select **Enter** as the method of entry for this set of independent variables.

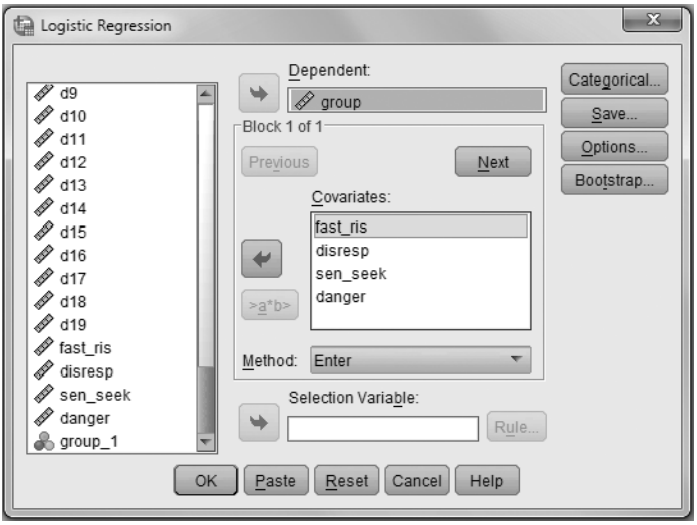


3. Click Options... to open the **Logistic Regression: Options** window. Check the **Hosmer-Lemeshow goodness-of-fit** cell, the **CI for exp(B): 95%** cell, and the **Include constant in model** cell.



Click Continue to return to the **Logistic Regression** window.

4. When the **Logistic Regression** window opens, click OK to complete the analysis. See Table 16.1 for the results.



16.4.2 SPSS Syntax Method

```
LOGISTIC REGRESSION VARIABLES GROUP
/METHOD = ENTER FAST_RIS DISRESP SEN_SEEK DANGER
/PRINT = GOODFIT CI(95)
/CRITERIA = PIN(0.05) POUT(0.10) ITERATE(20) CUT(0.5).
```

16.4.3 SPSS Output

TABLE 16.1

Logistic Regression Output

Logistic Regression

Case Processing Summary

Unweighted cases[a]	***N***	**Percent**
Selected cases		
Included in analysis	380	100.0
Missing cases	0	.0
Total	380	100.0
Unselected cases	0	.0
Total	380	100.0

[a] If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original value	**Internal value**
Accident	0
No accident	1

Block 0: Beginning Block

Classification Table[a,b]

			Predicted		
			group		
Observed			**Accident**	**No accident**	**Percentage correct**
Step 0	Group	Accident	0	170	.0
		No accident	0	210	100.0
	Overall percentage				55.3

[a] Constant is inluded in model

[b] The cut value is .500

Variables in the Equation

		B	**Standard error**	**Wald**	**D*f***	**Significance**	**exp(*B*)**
Step 0	Constant	.211	.103	4.195	1	.041	1.235

(*Continued*)

TABLE 16.1 (*Continued*)

Logistic Regression Output

Variables Not in the Equation

			Score	d*f*	Significance
Step 0	Variables	fast_ris	24.437	1	.000
		disresp	3.857	1	.050
		sen_seek	8.734	1	.003
		danger	24.629	1	.000
	Overall statistics		34.302	4	.000

Block 1: Method = Enter

Omnibus Tests of Model Coefficients

		Chi-square	d*f*	Significance
Step 1	Step	35.401	4	.000
	Block	35.401	4	.000
	Model	35.401	4	.000

Model Summary

Step	–2 Log Likelihood	Cox & Snell *R* Square	Nagelkerke *R* Square
1	487.173[a]	.089	.119

[a] Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	D*f*	Significance
1	8.960	8	.346

Contingency Table for Hosmer and Lemeshow Test

		Group = accident		Group = no accident		
		Observed	Expected	Observed	Expected	Total
Step 1	1	28	27.391	10	10.609	38
	2	24	23.790	14	14.210	38
	3	26	21.357	12	16.643	38
	4	12	19.262	26	18.738	38
	5	18	16.919	20	21.081	38

(*Continued*)

TABLE 16.1 (*Continued*)

Logistic Regression Output

	6	14	15.352	24	22.648	38
	7	16	13.842	22	24.158	38
	8	13	12.466	25	25.534	38
	9	10	11.035	28	26.965	38
	10	9	8.586	29	29.414	38

Classification Table[a]

			Predicted		
			Group		
	Observed		Accident	No accident	Percentage correct
Step 1	Group	Accident	84	86	49.4
		No accident	51	159	75.7
	Overall percentage				63.9

[a] The cut value is .500.

Variables in the Equation

								95% C.I for EXP(B)	
		B	S.E	Wald	d*f*	ig	Exp(B)	Lower	Upper
step 1[a]	fast_ris	-434	-.146	8.855	1	.003	.648	.487	.867
	disresp	.272	.153	3.166	1	.075	1.313	.973	1.772
	sen_seek	-.011	.122	.008	1	.930	.989	.779	1.257
	danger	.428	-154	7.658	1	.006	1.534	1.133	2.076
	constant	-441	.842	.275	1	.600	.643		

[a] Variable(s) entered on step 1: fast_ris, disresp, sen_seek, danger.

16.4.4 Results and Interpretation

16.4.4.1 Model Estimation

By default, logistic regression is run in two steps. The first step, called **Step 0**, estimates a base model that is used to provide a standard for comparison and includes no predictors and just the intercept. Often, this model is not interesting to researchers. Nevertheless, the results yielded by this base model can be used for prediction. For example, take a look at the classification results presented in the **Classification Table** presented in the **Block 0: Beginning Block** output. Given the base rates of the two groups of drivers (210/380 = 55.26% were not involved in accidents, 44.74% were involved in accidents) and no other information, the best strategy is to predict, for every case, that the driver was not involved in a motor vehicle accident. Using this strategy, we would be correct 55.26% of the time.

The results presented in the **Block 1: Method = Enter** output are those yielded in the second step (or model) with the four predictors (driving fast/risk taking; disrespect for traffic laws; sensation-seeking; and danger assessment) entered into it. In this case, it is the full model that was set in the logistic regression command. More steps are possible if the researcher uses stepwise or blocking of variables.

In the **Omnibus Tests of Model Coefficients** table, the inclusion of the four predictor variables yielded a chi-square value of 35.40 with 4 d*f*, $p < .001$. Thus, the overall model is statistically significant, which means that adding the four predictor variables to the model has significantly increased our ability to predict whether our subjects had or had not been involved in a motor vehicle accident.

The **Model Summary** table presents the **–2 log likelihood** statistic, which is a measure of how poorly the overall model predicts variations in the outcome of interest (accident versus no-accident drivers). The –2 log likelihood has a chi-squared distribution with smaller values indicating better model fit. In this example, the –2 log likelihood value is 487.173. By itself, this number is not very informative and is used primarily for nested model comparisons. In comparison with the base model that accepts only the intercept, inclusion of the four predictor variables reduced the –2 log likelihood statistic by 35.40, the chi-square value presented in the **Omnibus Tests of Model Coefficients** table. It should be noted that while SPSS does not give the –2 log likelihood (LL) statistic for the base model that accepts only the intercept, it can be easily calculated: since the full model's –2 LL statistic is 487.173, and adding the four predictor variables reduced the base model's –2 LL statistic by 35.40, therefore the base model's –2 LL statistic is 487.173 + 35.40 = 522.57.

16.4.4.2 Assessing Overall Model Fit

There are various measures available to assess model fit. First, we can use the reduction in the –2 log likelihood statistic to test the significance of the difference between any two models, as long as one model is nested inside the other. In this case, the base model that contains only the intercept has a –2 LL statistic of 522.57. Adding the four predictor variables produced a decrease of 35.40. This difference, which represents an improvement in model fit, is a chi-square value with 4 d*f* (one d*f* for each predictor variable) and is highly significant, $p < .001$.

Second, two measures of the *pseudo R*-square are available—the **Cox and Snell R^2** (.089) and the **Nagelkerke R^2** (.119). They are called *pseudo R*-square because in logistic regression analysis they are estimated from the maximum likelihood iteration process and are not calculated to minimize variance in the same way that *R*-square calculated from ordinary least-squares (OLS) does in multiple regression analysis. As such, they cannot be interpreted in the same way as one would interpret an OLS *R*-square. Their main purpose is to assess the goodness-of-fit of logistic models, with higher values indicating a better model fit.

The final measure of model fit is the **Hosmer and Lemeshow** test, which measures the correspondence of the actual and predicted values of the dependent variable. The better model fit is indicated by a smaller difference in the observed and predicted classification. A chi-square statistic is computed comparing the observed frequencies against the expected frequencies. A non-significant chi-square value indicates good model fit.

For the overall logistic regression model (that includes the four predictor variables), all of the measures of model fit improved over the base model that contains only the intercept. The –2 LL value decreased to 487.173, χ^2 ($df = 4$) = 35.40, $p < .001$. The *pseudo* R^2 values ranged from .089 to .119 and represent improvement over the base model. The **Hosmer and Lemeshow** test showed non-significance, indicating no difference in the distribution of the actual and predicted dependent values, χ^2 ($df = 8$) = 8.96, $p > .05$.

16.4.4.3 Classification Matrix

The **Classification Table** shows that 159/210 = 75.7% of the no-accident sub jects were correctly classified, while only 84/170 = 49.4% of the accident subjects were correctly classified. Overall, the classification results show that the model predicted correctly 243 out of 380 times, for an overall success rate of 63.9%. This is an improvement over the success rate of 55.3% for the base model with intercept only.

16.4.4.4 Test of Relationships and Strengths among the Variables

The goal is to identify which independent variables predict the outcome, that is, whether our subjects had or had not been involved in a motor vehicle accident. The **Variables in the Equation** table presents the **Wald** chi-square statistic, which is used to test the statistical significance of each coefficient (β) in the model. More specifically, the Wald chi-square statistic tests the unique contribution of each predictor—that is, holding constant the other predictors. On the basis of the Wald test, two predictors are found to be significant predictors of whether our subjects had or had not been involved in a motor vehicle accident—"driving fast/risk taking," (β = –.43), Wald χ^2 ($df = 1$) = 8.85, $p < .01$; "danger assessment," (β = .43), Wald χ^2 ($df = 1$) = 7.65, $p < .01$. Given that the dependent variable is coded 1 = accident and 2 = no accident, these coefficients imply that (1) those who had been involved in motor vehicle accidents have higher tendencies to drive fast and to take risks, and (2) those who had not been involved in motor vehicle accidents have a higher perception of risk.

16.4.4.5 Interpreting Odds Ratios

The **odds ratio** is a relative measure of risk and informs us how much more likely it is that somebody who is exposed to the predictor variable under study will develop the outcome as compared to someone who is not exposed. More specifically, it is a ratio of the odds at two values of the independent variable that are one unit apart. It indicates how many times higher the odds of occurrence are for each one-unit increase in the independent variable

(Sweet and Grace-Martin, 2002). In the present example, we can use the **odds ratio** shown in the column **Exp(B)** in the **Variables in the Equation** table to gauge the strength of the above significant relationships.

The odds ratio for FAST_RIS (motive to drive fast and to take risks) is .648. Given that the relationship between FAST_RIS and the dependent variable (coded 1 = accident and 2 = no accident) is negative (β = –.43), this shows that a higher FAST_RIS score decreases the odds of not being involved in an accident. How much is the decrease? Each one-unit increase on the FAST_RIS scale decreases the odds of not being involved in an accident by a factor of .648. Thus, a driver who likes to drive fast and to take risks (e.g., with a score of 5 on the 6-point Likert scale) is only .648 times as likely of not being involved in an accident as a driver who is less likely to drive fast and to take risks (with a score of 4 on the 6-point Likert scale).

The odds ratio for DANGER (danger assessment) is 1.534. Given that the relationship between DANGER and the dependent variable (coded 1 = accident and 2 = no accident) is positive (β = .42), this shows that a higher danger assessment score increases the odds of not being involved in an accident. How much will the increase be? Each one-unit increase on the DANGER scale increases the odds of not being involved in an accident by a factor of 1.534. Thus, a driver who assesses the danger highly (e.g., with a score of 5 on the 6-point Likert scale) is 1.534 times as likely of not being involved in an accident as a driver who assesses the danger lowly (with a score of 4 on the 6-point Likert scale).

16.4.4.6 Predictions on the Basis of Probabilities from Logistic Regression Coefficients

Similar to linear regression, the obtained logistic regression coefficients can be used to make predictions for the dependent variable. For example, it is possible to calculate the exact probability of not being involved in a motor vehicle accident for a driver who possesses the extreme tendency to drive fast and to take risks (FAST_RIS = 6). To do this, we need to consider the logistic regression equation:

$$\text{Log} - \text{odds} = a + b(x)$$

a = constant/intercept

b = regression coefficient

x = value of predictor variable

Because the regression equation represents the log-odds, it will be necessary to take the antilog (exp) to change the equation to probability.

Step 1: Calculate the odds that a driver who possesses the extreme tendency to drive fast and to take risks (FAST_RIS = 6) will not be involved in an accident.

The odds prediction equation is $\text{ODDS} = e^{a + bx}$. Thus, given that the driver scored an extreme high score of 6 on the FAST_RIS variable,

$$\text{ODDS} = e^{-.441 + (-.434 \times 6)} = e^{-3.045} = 0.048$$

That is, the driver who scored extremely high (6) on the FAST_RIS variable is only .048 as likely to not to be involved in an accident as he/she is likely to be involved in an accident.

Alternatively, if the driver scored an extreme low score of 1 on the FAST_RIS variable,

$$\text{ODDS} = e^{-.441 + (-.434 \times 1)} = e^{-0.875} = 0.417$$

That is, this driver who scored low (1) on the FAST_RIS variable (i.e., low tendency to drive fast and to take risks) is .417 as likely to not to be involved in an accident as he/she is likely to be involved in an accident.

Step 2: Converting odds to probabilities

The formula for converting odds to probabilities is Probability = (ODDS)/(1 + ODDS). For the driver who scored extremely high (6) on the FAST_RIS variable,

$$\text{Probability} = \frac{\text{ODDS}}{1 + \text{ODDS}} = \left(\frac{.048}{1.048}\right)$$

Thus, the probability of a "fast driving-risk taking" driver not to be involved in an accident is .046 or approximately 4.6 chances in a hundred.

Alternatively, for the driver who scored extremely low (1) on the FAST_RIS variable,

$$\text{Probability} = \frac{\text{ODDS}}{1 + \text{ODDS}} = \left(\frac{.417}{1.417}\right)$$

Thus, the probability of a "slow driving-low risk taking" driver not to be involved in an accident is .294, or approximately 29.4 chances in a hundred.

16.4.4.7 Summary of Probability Findings

For this case, the probability findings can be succinctly expressed in the following manner:

The tendencies to drive fast and to take risks are significantly associated with being involved in a motor vehicle accident. The probability of a driver who has an extremely high tendency to drive fast and to take risks to not to be involved in a vehicle accident is only .046. In comparison, drivers who tend to drive slower and not to take risks have a much higher probability (.294) of not being involved in a vehicle accident.

17

Canonical Correlation Analysis

17.1 Aim

Canonical correlation analysis (CCA) is a statistical technique that facilitates the study of interrelationships among sets of multiple dependent variables and multiple independent variables. Whereas multiple regression analysis (Chapter 14) is used to predict the value of a single (metric) dependent variable from a linear function of a set of independent variables, canonical correlation analysis predicts multiple dependent variables from multiple independent variables. For example, a researcher may be interested in how intelligence predicts academic performance among high school students. The researcher may have four different measures of intelligence in the predictor variable set and three different measures of academic performance in the criterion variable set. The research question of interest then will be whether there is a relationship between intelligence and academic performance as multi-operationalized in the two *latent* variable sets. These two latent variables are created in CCA by applying a linear equation (i.e., $Y' = \beta 1X1 + \beta 2X2$; β = standardized beta weight, X = observed scores in Z score form) to the measured predictor variables to create a single latent predictor variable and another linear equation to the observed dependent variables to create a single latent criterion variable. It should be noted that these two equations are generated to yield the largest possible correlation between the two latent variables. Thus, the most central statistic in a CCA is the *canonical correlation* that identifies the optimum structure or dimensionality of each of the latent predictor and latent criterion variables (weighted on the basis of the relationships between the measurement variables within the sets) that maximizes the relationship (conceptualized as a simple bivariate correlation—Pearson r) between them. Indeed, everything that occurs in the CCA is designed to maximize this simple correlation (Sherry and Henson, 2005).

In its most general form, canonical correlation analysis is conducted to answer two basic research questions:

1. Is there a noteworthy relationship between the predictor and criterion latent variable sets? That is, does the canonical model sufficiently

capture the relationship between the predictor and criterion variable sets to warrant interpretation? After all, if the relationship is small (and uninterpretable), it will make no sense to try to identify variables that contributed to that relationship.

2. If an important relationship is found to exist between the predictor and criterion latent variable sets, then what is the relative contribution of each variable to the canonical functions (relationships) extracted?

17.2 Checklist of Requirements

- The size of the sample has a direct impact on the "practical" significance of a study's outcome. Small sample sizes will not represent the correlations as well and may obscure any meaningful relationships. Very large samples, on the other hand, have the tendency to produce statistically significant outcomes even for small, unimportant effects that have no practical significance. Hair et al. (1995) and Tabachnick and Fidell (2001) recommend at least 10 cases per variable to avoid "overfitting" the data.

17.3 Assumptions

- **Linearity**—Similar to other multivariate techniques that employ a variate (i.e., *linear combination that represents the weighted sum of two or more variables*), an implicit assumption is that all relationships among all pairs of variables are linear. The assumption of linearity affects two aspects of the canonical correlation results. First, the analysis is performed on correlation or variance-covariance matrices that reflect only linear relationships. If the relationship between two variables is nonlinear, it is not captured by these statistics (Tabachnick and Fidell, 2001). Second, the canonical correlation is the linear relationship between the variates. If the variates relate in a nonlinear fashion, the relationship will not be captured by the canonical correlation (Hair et al., 1995).
- **Multivariate normality**—Canonical correlation analysis imposes no strict requirement that the variables be normally distributed. However, normality is desirable because interpretation derived from the statistical inference test of the significance of each canonical

function is based on the assumption of multivariate normality. This assumption requires that all variables and all linear combinations of variables are normally distributed. However, multivariate normality is difficult to test, although the likelihood of multivariate normality is increased if the variables are all normally distributed.

- **Homoscedasticity**—Canonical analysis is best when relationships among pairs of variables are homoscedastic, that is, when one variable exhibits similar amount of variance across the range of values for the other variable. If the variances are different from each other (exhibit heteroscedasticity), the probability of obtaining a statistically significant result if the null hypothesis is true is greater than the desired alpha level, that is, it increases the probability of committing a Type I error.

17.4 Key Terms in Canonical Correlation Analysis

The following lists some of the key terms encountered in CCA. Review of these terms should assist in developing an understanding of the key concepts underlying CCA (see Hair et al., 1995; Sherry and Henson, 2005).

- *Canonical correlation coefficient* (*Rc*)—The bivariate (Pearson *r*) correlation between the two canonical variates. It is a measure of the strength of the overall relationships between the linear composites (variates) for the predictor and dependent variables. *Rc* is directly analogous to the multiple *R* in regression.
- *Squared canonical correlation* (Rc^2)—Also known as *canonical root* and *eigenvalue*, this is the square of the canonical correlation and provides an estimate of the amount of shared variance between the two latent sets of predictor and dependent variables. It is directly analogous to the R^2 statistic in multiple regression.
- *Canonical variates*—The linear combinations that represent the weighted sum of the variables that comprise either the dependent or independent latent variables.
- *Canonical function*—The relationship between two canonical variates, one variate for the dependent variables and one variate for the independent variables. The strength of the relationship is given by the canonical correlation coefficient (*Rc*). CCA will extract as many functions as there are variables in the smaller variable set. For example, if there are five independent variables and three dependent variables, the maximum number of canonical functions that can be extracted is three. The extracted functions are orthogonal (independent) of each other.

- *Standardized canonical function coefficients*—Also known as *canonical weights*, these are standardized coefficients used in the linear equations to combine the observed predictor and dependent variables into two respective latent variables. The sign and magnitude of the coefficient assigned to each variable in its canonical variate aid in the interpretation of the canonical functions (relationship between the dependent and independent variants). Variables with relatively larger coefficients (weights) contribute more to the variates, and vice versa. These weights can be interpreted in the same way as beta weights in multiple regression.
- *Structure coefficient* (*rs*)—Also known as *canonical loading*, it is a measure of the simple bivariate (Pearson *r*) correlation between an observed variable (e.g., a predictor variable) and the canonical function scores (canonical variates) for the variable set (e.g., the latent variable created from all the predictor variables via the linear equation). Like factor loadings, it aids in the interpretation of the structure of the computed latent variable, by identifying which observed variables contributed to the creation of the latent variable.
- *Squared canonical structure coefficients* (rs^2)—These are the square of the structure coefficients. This statistic indicates the amount of shared variance between an observed variable and the latent variable generated from the observed variable's set.

17.5 An Example of Canonical Correlation Analysis

Ho (2000) investigated the efficacy of two competing Protection Motivation models (Ordered Protection Motivation versus Protection Motivation) in predicting intention for safe sex behavior (condom use) following HIV/AIDS communication. Items were written to represent the protection motivation variables of *perceived risk, perceived severity, self-efficacy, response efficacy, maladaptive coping,* and *fear,* and the "intention for condom use" variables of *willing to use, likely to use, intend to use,* and *certain to use*. These 10 items, together with their full descriptors, are listed below. Canonical correlation analysis will be used to investigate whether there is a relationship between protection motivation and intention for condom use as multi-operationalized in the two *latent* variable sets.

- Protection motivation variables

  Perceived_risk—Not using a condom (or not insisting that one be used during sexual intercourse) will lead to an increased risk of getting AIDS.

Perceived_severity—Contracting AIDS is equivalent to receiving a death sentence.

Self_efficacy—By practicing safe sex (e.g., using a condom or insisting that one be used), I am confident that I can avoid getting AIDS.

Response_efficacy—I can lower the probability of contracting AIDS by always practicing safe sex (e.g., using a condom or insisting that one be used).

Maladaptive_coping—I don't worry about AIDS because I am capable of getting HIV-free partners.

Fear—The fact that so many people have died of AIDS really frightens me.

- Intention for condom use

Willing_use—I am willing to use a condom (or insist that one be used) when I have sex in the future.

Likely_use—It is highly probable that I will use a condom (or insist that one be used when I have sex in the future.

Intend_use—I intend to use a condom (or insist that one be used) when I have sex in the future.

Certain_use—I am certain that I will use a condom (or insist that one be used) when I have sex in the future.

17.5.1 Data Entry Format

The data set is named **HIV.SAV**.

Variables	Code
• Sex	1 = male, 2 = female
• Age	in years
• Involve	1 = yes, 2 = no
• Healedu	1 = yes, 2 = no
• **Perceived_risk**	1 = strongly disagree, 5 = strongly agree
• **Perceived_severity**	1 = strongly disagree, 5 = strongly agree
• **Self_efficacy**	1 = strongly disagree, 5 = strongly agree
• **Maladaptive_coping**	1 = strongly disagree, 5 = strongly agree
• **Willing_use**	1 = strongly disagree, 5 = strongly agree
• **Response_efficacy**	1 = strongly disagree, 5 = strongly agree
• **Likely_use**	1 = strongly disagree, 5 = strongly agree
• **Fear**	1 = strongly disagree, 5 = strongly agree
• **Intend_use**	1 = strongly disagree, 5 = strongly agree
• **Certain_use**	1 = strongly disagree, 5 = strongly agree

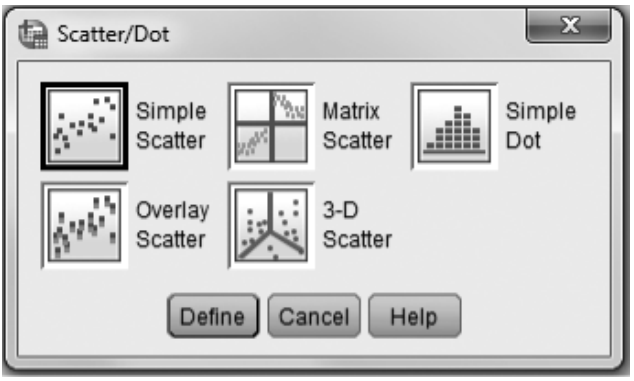
17.5.2 Testing Assumptions

17.5.2.1 Linearity

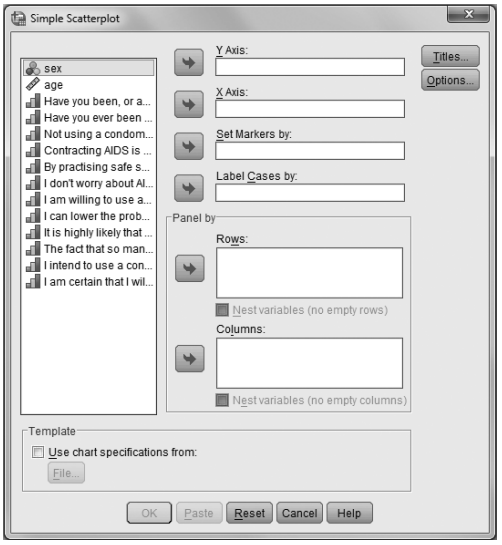
Linearity can be assessed by examining the scatterplots of the variables.

17.5.2.1.1 Windows Method

1. From the menu bar, click **Graphs**, then **Legacy Dialogs**, and then-**Scatter/Dot...**. The following **Scatter/Dot** window will open.

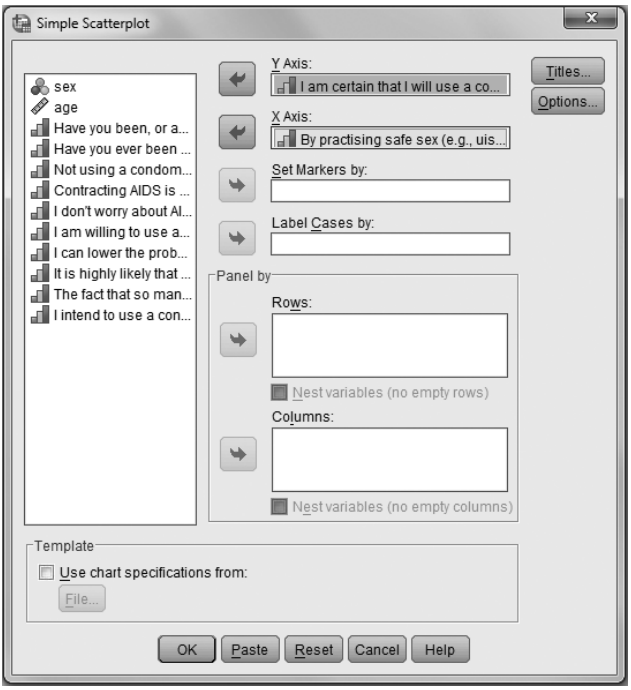


2. Click **Simple Scatter** and then Define**.** The following **Simple Scatterplot** window will open.



3. To obtain a scatterplot for the **SELF_EFFICACY** and **CERTAIN_USE** variables, transfer the **CERTAIN_USE** variable to the **Y Axis:** field

and the **SELF_EFFICACY** variable to the **X Axis**: field by clicking these variables (highlight) and then clicking ▶.



4. Click OK to complete the analysis. Repeat steps 1 to 4 to obtain scatterplots for all other pairs of variables. See Figure 17.1 for the **SELF_EFFICACY** with **CERTAIN_USE** scatterplot.

17.5.2.1.2 SPSS Syntax Method

```
GRAPH
/SCATTERPLOT(BIVAR) = SELF_EFFICACY WITH CERTAIN_USE
/MISSING = LISTWISE.
```

NOTE: To obtain scatterplots for other pairs of variables, substitute the names of the variables in the above syntax file with the names of the other variables.

17.5.2.1.3 SPSS Output

Examination of the scatterplot shows no serious nonlinearity between the pair of variables.

NOTE: In the interests of brevity and page space, the other 23 scatterplots are not presented.

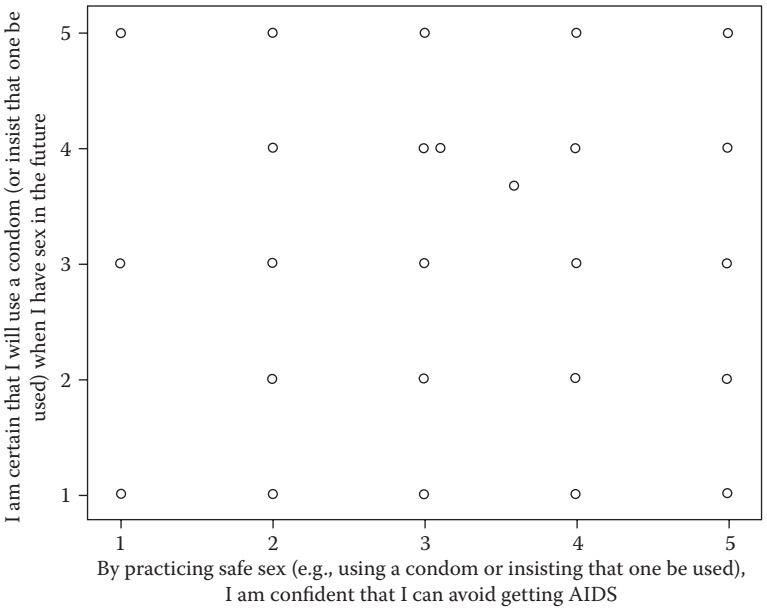


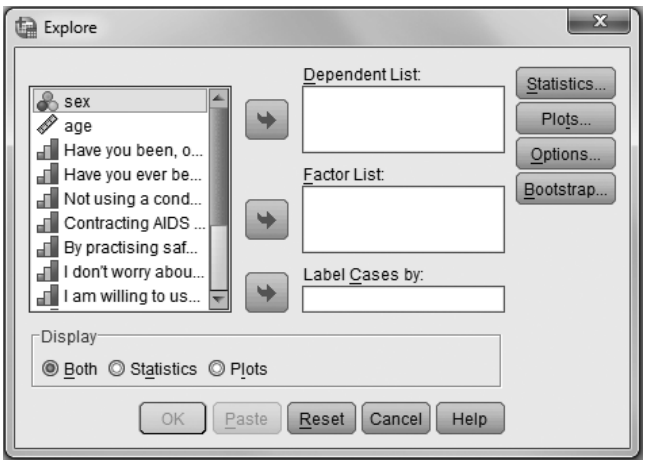
FIGURE 17.1
Scatterplot for the variables of SELF_EFFICACY and CERTAIN_USE.

17.5.2.2 Multivariate Normality

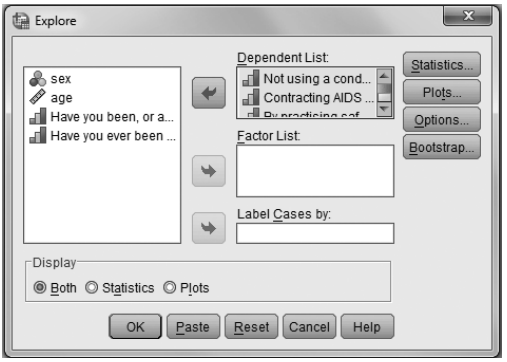
As there are no specific tests that test for multivariate normality, univariate normality for the ten variables will be tested instead.

17.5.2.2.1 Windows Method

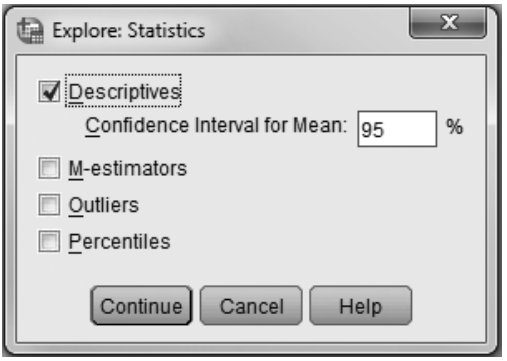
1. From the menu bar, click **Analyze**, then **Descriptive Statistics**, and then **Explore**. The following **Explore** window will open.



2. Transfer the **PERCEIVED_RISK**, **PERCEIVED_SEVERITY**, **SELF_EFFICACY**, **RESPONSE_EFFICACY**, **MALADAPTIVE_COPING**, **FEAR**, **WILLING_USE**, **LIKELY_USE**, **INTEND_USE**, and **CERTAIN_USE** variables to the **Dependent List:** field by clicking these variables (highlight) and then clicking  .



3. Click  to open the **Explore: Statistics** window. Check the **Descriptives** field and click  to return to the **Explore** window.

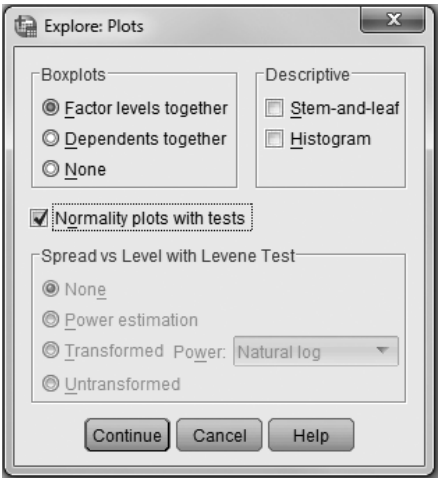


4. In the **Explore** window click  to open the **Explore: Plots** window. Check the **Factor levels together** field and the **Normality plots with tests** field. Click  to return to the **Explore** window.
5. When the **Explore** window opens, click  to complete the analysis. See Figure 17.2 for the results.

17.5.2.2.2 SPSS Syntax Method

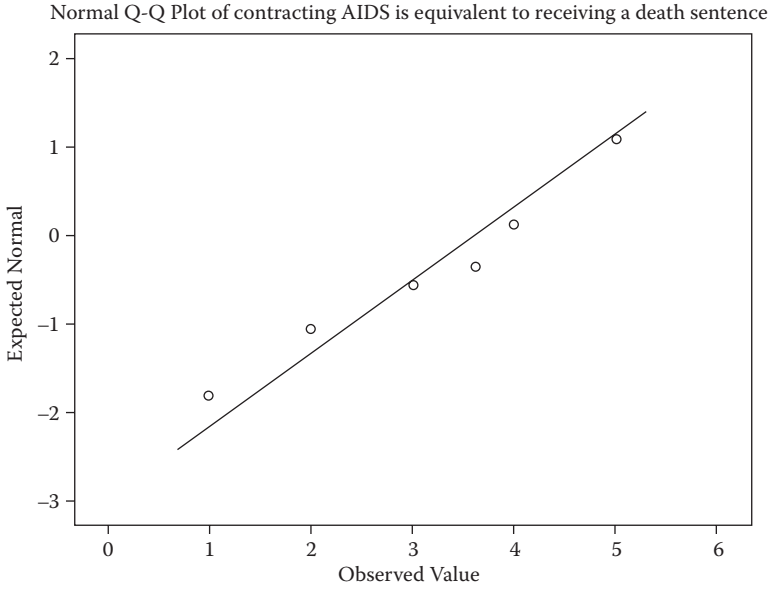
```
EXAMINE VARIABLES = PERCEIVED_RISK PERCEIVED_SEVERITY SELF_
EFFICACY RESPONSE_EFFICACY MALADAPTIVE_COPING FEAR WILLING_USE
LIKELY_USE INTEND_USE CERTAIN_USE
```

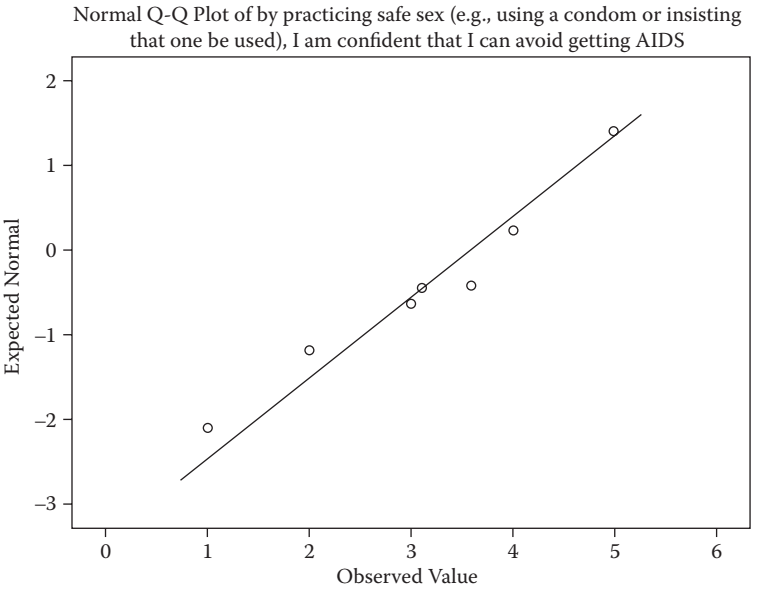
```
/PLOT BOXPLOT NPPLOT
/COMPARE GROUPS
/STATISTICS DESCRIPTIVES
/CINTERVAL 95
/MISSING LISTWISE
/NOTOTAL.
```

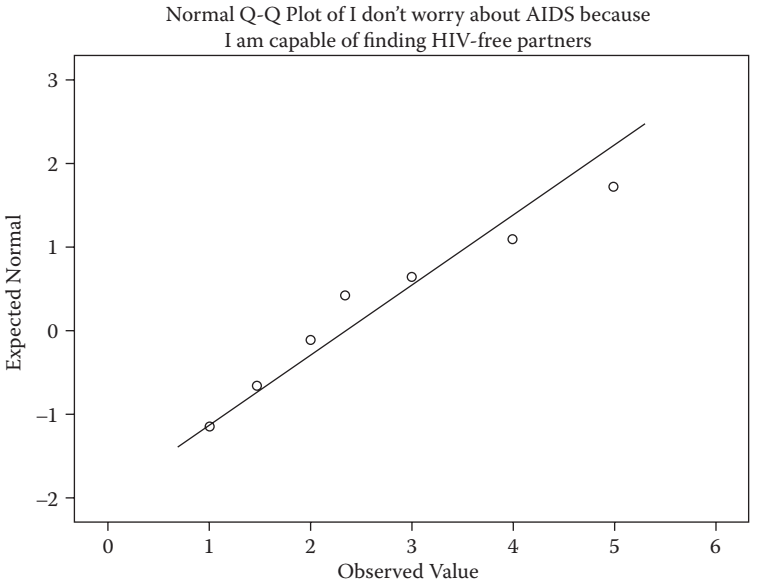


17.5.2.2.3 SPSS Output

Explore Analysis (Selected) Output







Normal Q–Q plots for selected variables

The normal distribution forms a straight diagonal line, and if a variable's distribution is normal, the data distribution will fall more or less along the diagonal. A visual check of the normal probability plots shows very slight departure from normality for the above five variables.

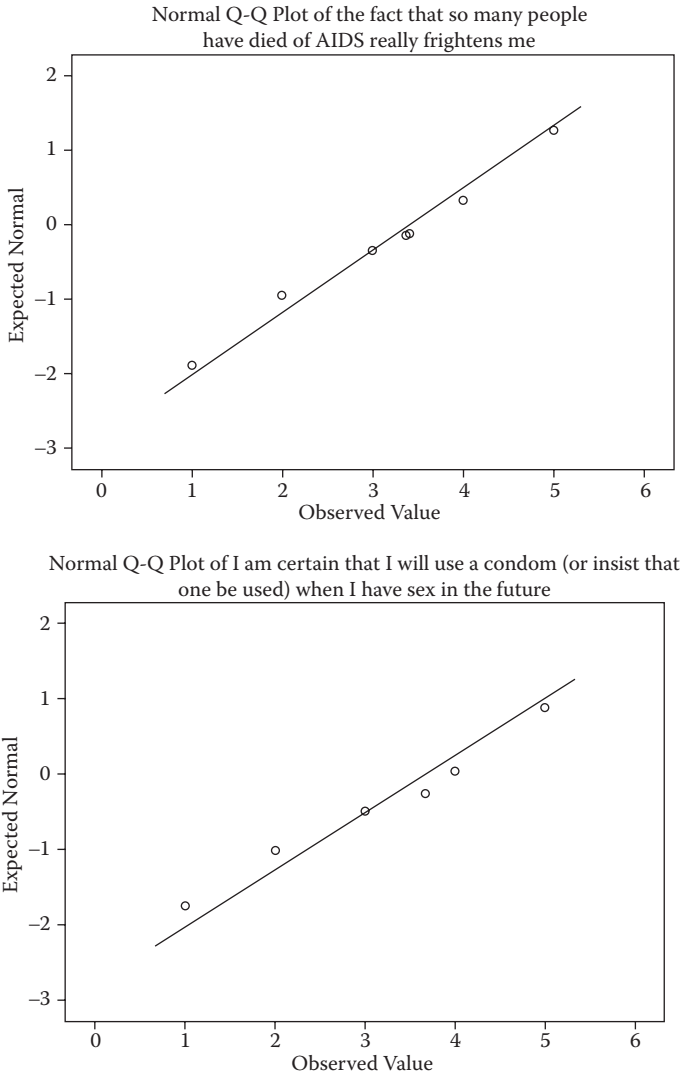


FIGURE 17.2
Normal Q-Q plots for selected variables.

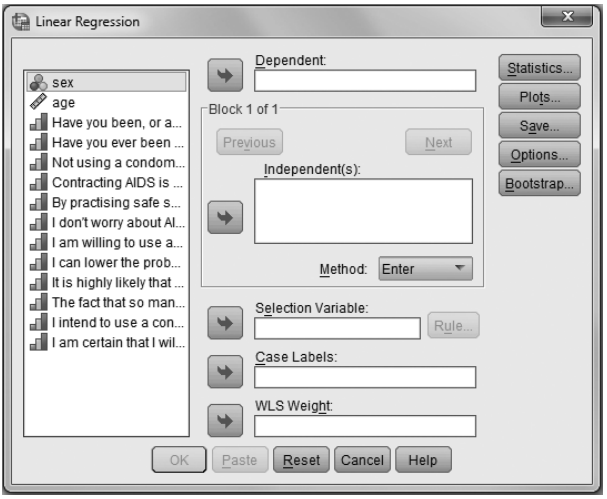
NOTE: In the interests of brevity and page space, the normal probability plots for the other five variables are not presented to save space.

17.5.2.3 Homoscedasticity

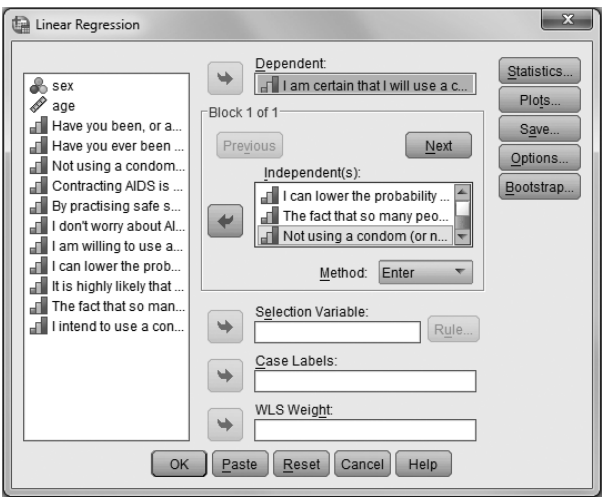
The assumption of homoscedasticity was tested using both residual scatterplots and the normal plot of regression standardized residuals for the dependent variables **WILLING_USE**, **LIKELY_USE**, **INTEND_USE**, and **CERTAIN_USE**.

17.5.2.3.1 Windows Method

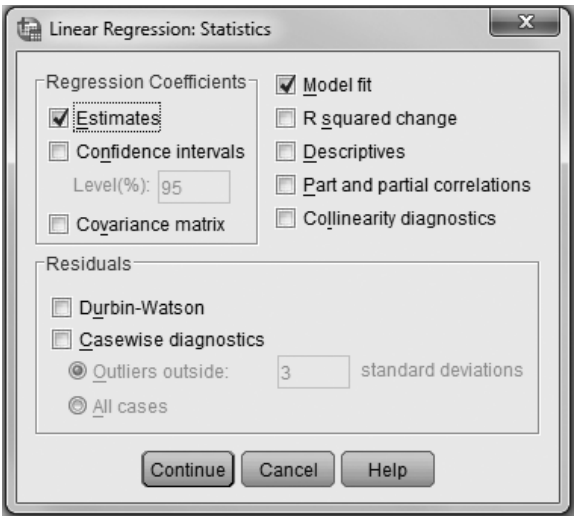
1. From the menu bar, click **Analyze**, then **Regression**, and then **Linear**. The following **Linear Regression** window will open.



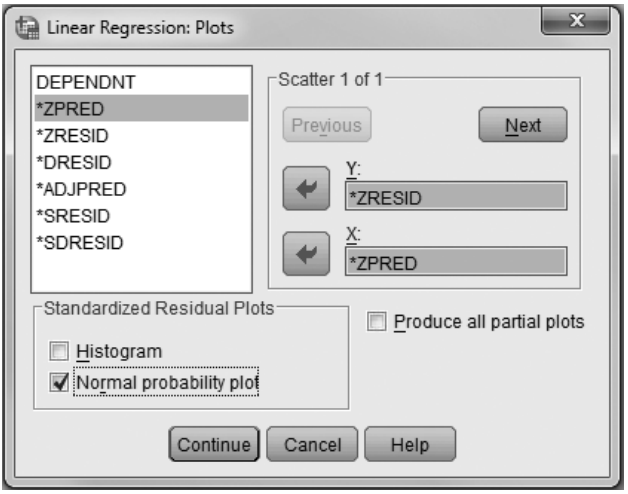
2. Transfer the **CERTAIN_USE** variable to the **Dependent:** field by clicking this variable (highlight) and then clicking . Transfer the **PERCEIVED_RISK**, **PERCEIVED_SEVERITY**, **SELF_EFFICACY**, **RESPONSE_EFFICACY**, **MALADAPTIVE_COPING**, and **FEAR** variables to the **Independent(s):** field by clicking these variables (highlight) and then clicking .



3. Click Statistics... to open the **Linear Regression: Statistics** window. Check the **Estimates** and **Model fit** fields and click Continue to return to the **Linear Regression** window.



4. In the **Linear Regression** window click Plots... to open the **Linear Regression: Plots** window. Transfer ***ZRESID** to the **Y:** field by clicking this variable (highlight) and then clicking ▶. Transfer ***ZPRED** to the **X:** field by clicking this variable (highlight) and then clicking ▶. In the **Standardized Residual Plots** box, check the **Normal probability plot** field.



Click Continue to return to the **Linear Regression** window.

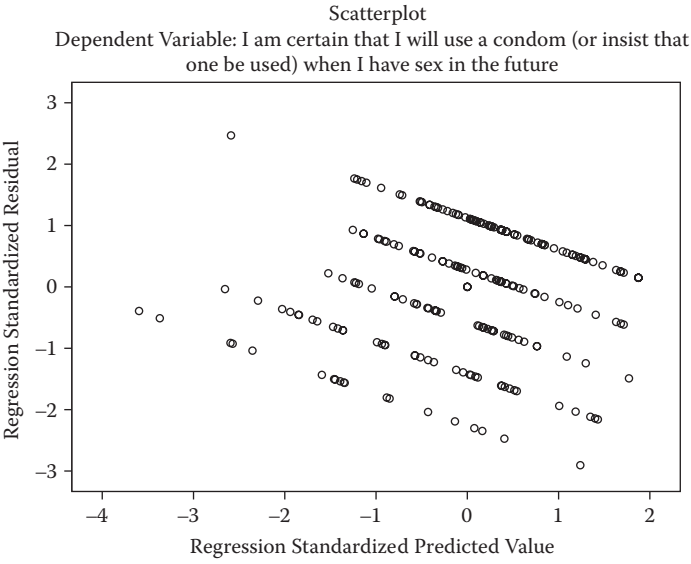
5. When the **Linear Regression** window opens, click OK to complete the analysis.
6. Repeat steps 1 to 5 to obtain residual scatterplots and the normal plot of regression standardized residuals for the other three dependent variables **WILLING_USE**, **LIKELY_USE**, and **INTEND_USE**.

17.5.2.3.2 SPSS Syntax Method

```
REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA
/CRITERIA = PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT CERTAIN_USE WILLING_USE LIKELY_USE INTEND_USE
/METHOD = ENTER PERCEIVED_RISK PERCEIVED_SEVERITY SELF_
EFFICACY MALADAPTIVE_COPING RESPONSE_EFFICACY FEAR
/SCATTERPLOT = (*ZRESID,*ZPRED)
/RESIDUALS NORMPROB(ZRESID).
```

17.5.2.3.3 SPSS Output

Figures 17.3 to Figure 17.6 present the scatterplots of residuals against predicted values, and the normal plot of regression standardized residuals for the dependent variables **CERTAIN_USE**, **WILLING_USE**, **LIKELY_USE**, and **INTEND_USE**.



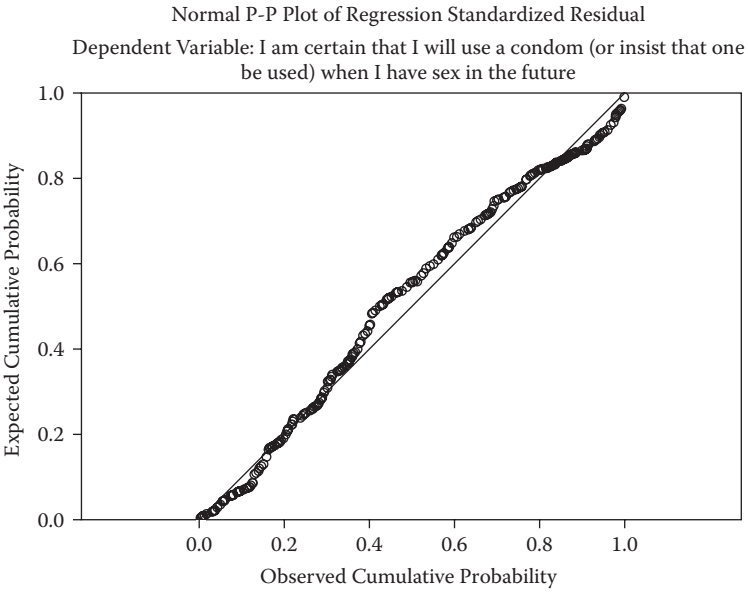
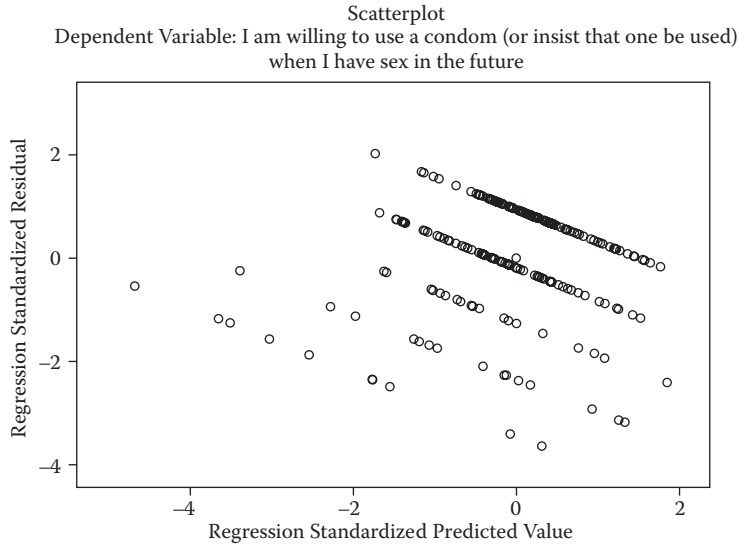


FIGURE 17.3
Scatterplot of residuals against predicted values, and the normal plot of regression standardized residuals for the dependent variable of CERTAIN_USE.

The **scatterplots of residuals against predicted values** show the cluster of points to be approximately the same width all over (the shape of the residual plots is rectangular with a concentration of points along the center) indicating no clear relationship between the residuals and the predicted values. Looking at the **Normal P-P Plot of Regression Standardized**



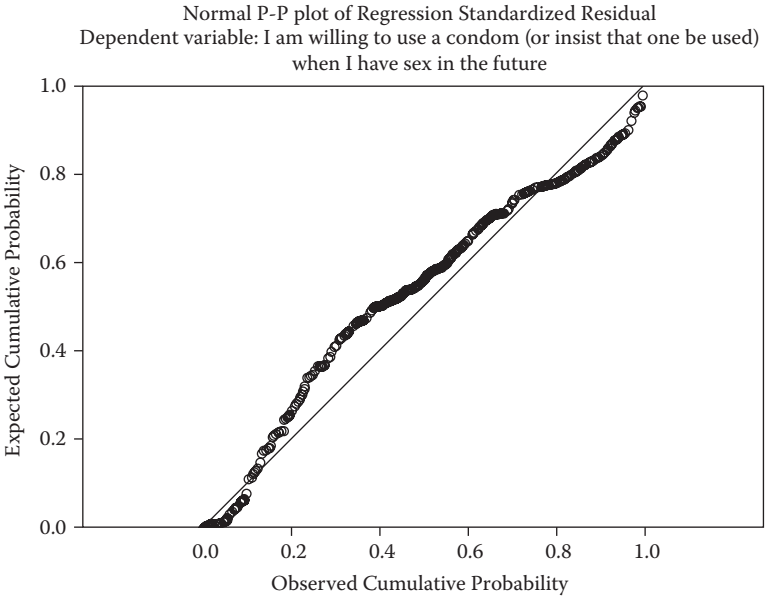
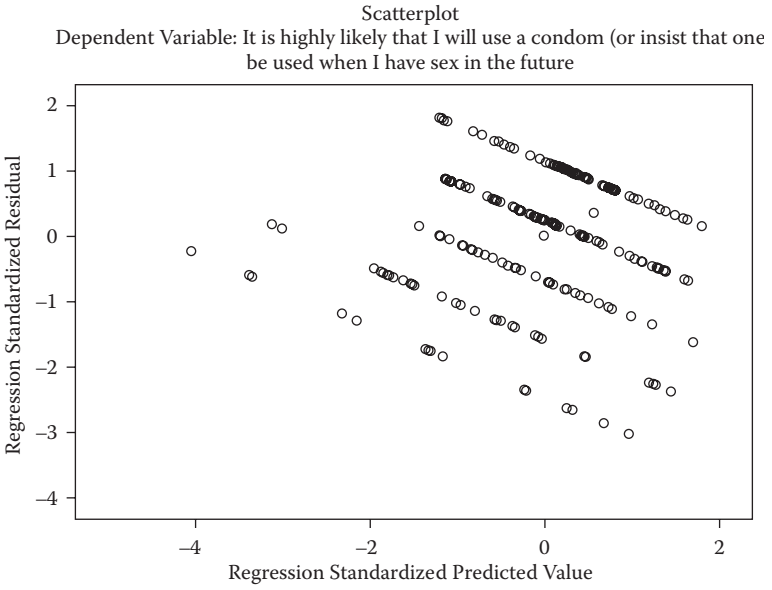


FIGURE 17.4
Scatterplot of residuals against predicted values, and the normal plot of regression standardized residuals for the dependent variable of WILLING_USE.



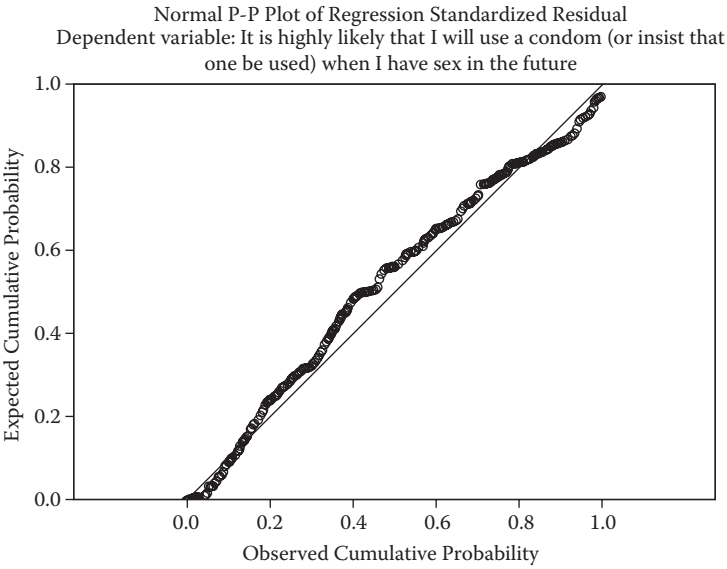
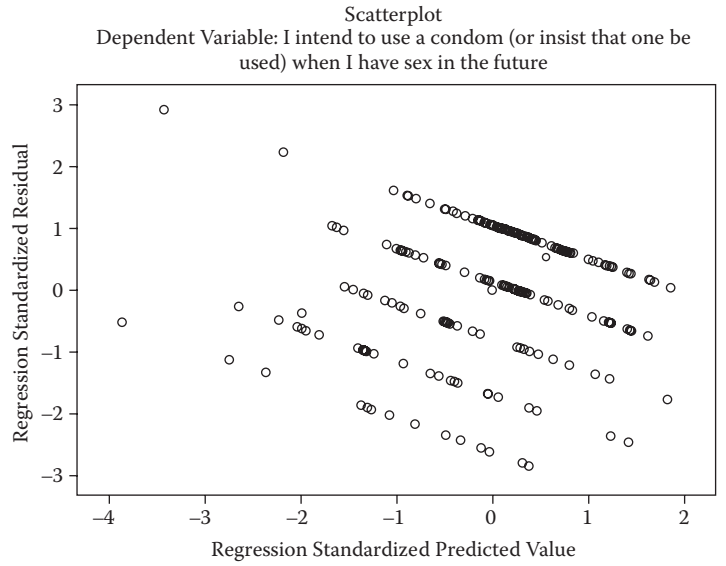


FIGURE 17.5
Scatterplot of residuals against predicted values, and the normal plot of regression standardized residuals for the dependent variable of LIKELY_USE.



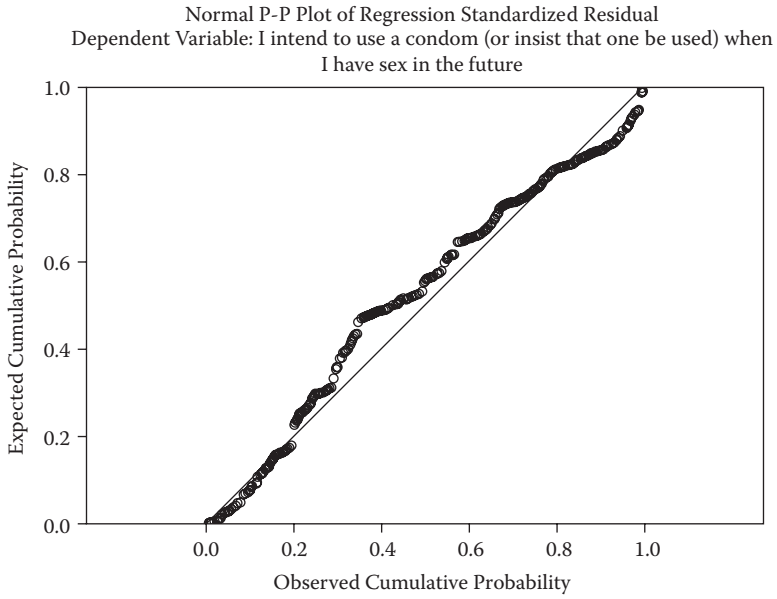


FIGURE 17.6
Scatterplot of residuals against predicted values, and the normal plot of regression standardized residuals for the dependent variable of INTEND_USE.

Residuals for the four dependent variables, it can also be seen that the data are homoscedastic as the residuals plot is the same width for all values of the predicted dependent variable (DV). Heteroscedasticity is usually indicated by a cluster of points that is wider as the values for the predicted DV get larger.

17.5.3 Example of Canonical Correlation Analysis

17.5.3.1 Windows Method

There is no Windows point-and-click option in SPSS for CCA. The analysis will be conducted via the SPSS syntax method.

17.5.3.2 SPSS Syntax Method

```
MANOVA WILLING_USE LIKELY_USE INTEND_USE CERTAIN_USE WITH
PERCEIVED_RISK PERCEIVED_SEVERITY SELF_EFFICACY RESPONSE_
EFFICACY MALADAPTIVE_COPING FEAR
/PRINT = SIGNIF(MULT EIGEN DIMENR)
/DISCRIM = (STAN ESTIM COR ALPHA(.999)).
```

17.5.3.3 SPSS Output

TABLE 17.1

Canonical Correlation Analysis Output

MANOVA

```
The default error term in MANOVA has been changed from WITHIN CELLS to
WITHIN+RESIDUAL. Note that these are the same for all full factorial
 designs.
* * * * * * * * * * * * * * * * * A n a l y s i s o f V a r i a n c e * * *
248 cases accepted.
0 cases rejected because of out-of-range factor values.
0 cases rejected because of missing data.
1 non-empty cell.
1 design will be processed.
```

*** * * * * * * * * * A n a l y s i s o f V a r i a n c e— Design 1 ***

EFFECT.. WITHIN CELLS Regression

Multivariate Tests of Significance (S = 4, M = 1/2, N = 118)

Test Name	Value Approx.	F Hypoth.	DF	Error DF	Sig. of F
Pillais	.42671	4.79656	24.00	964.00	.000
Hotellings	.56586	5.57612	24.00	946.00	.000
Wilks	.61355	5.20746	24.00	831.49	.000
Roys	.30026				

Eigenvalues and Canonical Correlations

Root No.	Eigenvalue	Pct.	Cum. Pct.	Canon Cor.	Sq. Cor
1	.42910	75.83175	75.83175	.54796	.30026
2	.09960	17.60216	93.43392	.30097	.09058
3	.03655	6.45888	99.89280	.18778	.03526
4	.00061	.10720	100.00000	.02462	.00061

Dimension Reduction Analysis

Roots	Wilks	L. F Hypoth.	DF	Error DF	Sig. of F
1 TO 4	.61355	5.20746	24.00	831.49	.000
2 TO 4	.87682	2.14646	15.00	660.18	.007
3 TO 4	.96416	1.10514	8.00	480.00	.358
4 TO 4	.99939	.04873	3.00	241.00	.986

Standardized canonical coefficients for DEPENDENT variables

Function No.

Variable	1	2	3	4
willing_use	.63363	1.05342	-.87571	.38484
likely_use	.29951	.66846	1.55151	-.94464
intend_use	.04925	-1.30269	-1.13959	-1.01083
certain_use	.10571	-.62695	.52111	1.61291

TABLE 17.1 (*Continued*)

Canonical Correlation Analysis Output

MANOVA

Correlations between DEPENDENT and canonical variables

	Function No.			
Variable	**1**	**2**	**3**	**4**
willing_use	.95960	.16800	−.21405	.07157
likely_use	.88372	−.12531	.38690	−.23163
intend_use	.83784	−.49283	−.08312	−.21962
certain_use	.81370	−.42234	.22557	.32962

Variance in dependent variables explained by canonical variables

CAN. VAR.	Pct Var DEP	Cum Pct DEP	Pct Var COV	Cum Pct COV
1	76.64672	76.64672	23.01404	23.01404
2	11.62929	88.27601	1.05340	24.06744
3	6.33251	94.60853	.22328	24.29073
4	5.39147	100.00000	.00327	24.29399

Standardized canonical coefficients for COVARIATES CAN. VAR.

COVARIATE	1	2	3	4
perceived_risk	.28582	.99383	.00864	.08364
perceived_severity	−.05234	−.26458	.57603	−.67543
self_efficacy	.20748	−.08289	−.23799	.10297
response_efficacy	.35028	−.25751	−.66394	−.45017
maladaptive_coping	−.54061	.04031	−.49239	.17546
fear	.37043	−.41352	−.11883	.85545

Correlations between COVARIATES and canonical variables CAN. VAR.

Covariate	1	2	3	4
perceived_risk	.50170	.77993	.05105	−.00114
perceived_severity	.27121	−.18813	.52781	−.40719
self_efficacy	.42535	−.07576	−.36879	−.11279
response_efficacy	.53298	−.14214	−.63337	−.48263
maladaptive_coping	−.72944	.09676	−.43125	.08668
fear	.54397	−.31032	.21129	.58940

Variance in covariates explained by canonical variables

CAN. VAR.	Pct Var DEP	Cum Pct DEP	Pct Var COV	Cum Pct COV
1	8.09822	8.09822	26.97058	26.97058
2	1.17046	9.26868	12.92154	39.89212
3	.61644	9.88512	17.48280	57.37492
4	.00774	9.89286	12.77290	70.14783

17.5.3.4 Results and Interpretation

As stated earlier, the two basic research questions posed by the CCA are:

1. Is there a noteworthy relationship between the predictor and criterion latent variable sets?
2. If an important relationship is found to exist between the predictor and criterion latent variable sets, then what is the relative contribution of each variable to the canonical functions (relationships) extracted?

In answering these two research questions, the interpretation of the CCA results will follow Thompson's (1997) suggestion of a two-stage hierarchical approach. Stage 1 will seek to identify and to isolate notable (significant) relationships, and stage 2 will seek to identify what variables in the model may have contributed to that relationship.

17.5.3.4.1 Stage 1. Is There a Noteworthy Relationship between the Predictor and Criterion Latent Variable Sets? As CCA will generate as many functions (i.e., variates) as there are variables in the smaller latent variable set (WILLING_USE, LIKELY_USE, INTEND_USE, CERTAIN_USE), the analysis derived four canonical functions. The initial consideration will therefore be to determine the number of canonical functions to be included for interpretation.

17.5.3.4.1.1 Multivariate tests of significance The first part of the SPSS output presents four multivariate tests of significance (Pillai's criterion, Hotelling's trace, Wilks' lambda, and Roy's GCR) for the full canonical model, that is, they evaluate the shared variance between the predictor and criterion variables across all four canonical functions. While all four tests returned statistically significant results at the .001 level, the most common test used is Wilks' lambda (λ), as it tends to have the most general applicability (Sherry and Henson, 2005). For the full model, Wilks' λ is .614, $F(24,831.49) = 5.207$, $p < .001$. Thus, the null hypothesis that there is no relationship between the two latent variable sets ($Rc = 0$) can be rejected. That is, there is probably a relationship.

17.5.3.4.1.2 Effect size It should be noted that statistically significant results can be impacted heavily by sample size, such that small, unimportant effects can be found to be statistically significant given large enough sample sizes. Therefore, it is important to interpret effect size indices to determine the model's *practical significance* alongside its statistical significance. Effect size in CCA can be determined by Wilks' λ, which represents an inverse effect size or the amount of variance not shared between the variable sets (Sherry and Henson, 2005). Therefore, by subtracting this amount from 1 provides us with an overall effect size, that is, the amount of variance shared between the variable sets ($1 - .614 = .386$ for the full model). This effect statistic can be interpreted as the proportion of variance

shared between the two variable sets across the four functions. Taken together, the results (so far) indicate that the entire model is both statistically significant and has what may be considered a moderate effect size.

17.5.3.4.1.3 Eigenvalues and canonical correlations This next set of statistics is used to evaluate each part individually. Only those functions that explain a reasonable amount of variation between the two latent variable sets should be interpreted. Looking at the squared canonical correlations (**Sq. Cor**), it is clear that only the first function (.30026) (which explained 30.03% of the variance within its function) should be retained for interpretation. Nevertheless, it was determined to retain the second function (.09058) (which explained only 9.06% of the variance within its function) for interpretation also, in order to compare as well as to demonstrate the explanatory power of Function 1 over Function 2.

17.5.3.4.1.4 Dimension reduction analysis The question can be raised as to why each of the four functions' canonical correlations are not tested for statistical significance in order to ascertain whether the function should be interpreted. The simple answer is that there is no simple way to do this. The tests of significance presented in this table do not test each function separately, but rather in a hierarchical fashion in which the full model (Functions 1 to 4) is tested first, then Functions 2 to 4 are tested second, Functions 3 to 4 are tested third, and then Function 4 is tested by itself. As pointed out by Sherry and Henson (2005), it is a common error to assume that the "1 to 4" test evaluates the first function, the "2 to 4" test evaluates the second function, and the "3 to 4" test evaluates the third function. Only the "4 to 4" test evaluates the fourth function on its own. Looking at the results presented in the **Dimension Reduction Analysis** table, it can be seen that the full model (Functions 1 to 4) is statistically significant, Wilks' λ is .614, $F(24,831.49) = 5.207$, $p < .001$. The cumulative effects of Functions 2 to 4 are also statistically significant, Wilks' λ is .877, $F(15,660.18) = 2.146$, $p < .01$. The cumulative effects of Functions 3 to 4 are not statistically significant nor are the effects of Function 4 in isolation. Although Functions 2 to 4 are cumulatively statistically significant, they only explained 9.06%, 3.53%, and .061%, respectively, of the variance by themselves (see **Sq. Cor.** under the **Eigenvalues and Canonical Correlations** table). Given the small amount of variance accounted for by Function 2 in isolation (9.06%), it would not have been retained for interpretation. However, for reasons stated earlier, it will be retained along with Function 1 for illustrative purposes.

17.5.3.4.1.5 Summary It can be concluded from the analysis so far that there is indeed a noteworthy canonical relationship between the two latent variable sets. This conclusion is based on statistical significance of the relationship, effect size, and the amount of variance explained within each function. Moreover, this relationship is largely captured by the first function in the canonical model. Although the second function could be considered practically nonsignificant,

owing to the minuscule amount of shared variance within the function, it will be included in the interpretation phase for illustrative purposes.

17.5.3.4.2 Stage 2. What Is the Relative Contribution of Each Variable to the Canonical Functions (Relationships) Extracted? In the present case, we are interested in knowing (in terms of magnitude and direction) what protection motivation variables are related to the "intentions for condom use." The four methods for interpretation are (1) standardized canonical coefficients (Coef), (2) structure coefficients (rs), (3) squared structure coefficients (rs^2), and (4) communality coefficients (h^2).

17.5.3.4.2.1 Standardized canonical coefficients versus structure coefficiets **Standardized canonical coefficients** reflect the relative contribution of one predictor variable to the criterion variable given the contribution of other predictors. **Structure coefficients** reflect the direct contribution of one predictor variable to the criterion variable regardless of other predictors. Because standardized canonical coefficients are typically unstable, especially in the presence of multicollinearity (when predictor variables are highly correlated), owing to their calculation solely to optimize the canonical correlation, the structure coefficients are considered more appropriate for interpretation (Hair et al., 1995). Indeed, structure coefficients increase in importance when multicollinearity increases among the observed variables in the canonical model.

17.5.3.4.2.2 Interpretation Only the first two functions will be interpreted. Table 17.2 presents the (1) standardized canonical coefficients (*Coef*),

TABLE 17.2

Canonical Solution for Protection Motivation Predicting Intentions for Condom Use for Functions 1 and 2

	Function 1			Function 2			
Variable	**Coef**	**rs**	**rs^2 (%)**	**Coef**	**rs**	**rs^2 (%)**	**h^2 (%)**
Willing_use	.634	.959	91.97	1.053	.168	2.82	94.79
Likely_use	.299	.883	77.97	.668	−.125	1.56	79.53
Intend_use	.049	.838	70.22	−1.303	−.493	24.30	94.52
Certain_use	.106	.814	66.26	−.627	−.422	17.81	84.07
Perceived_risk	.286	.502	25.20	.994	.780	60.84	86.04
Perceived_severity	−.052	.271	7.34	−.265	−.188	3.53	10.87
Self_efficacy	.207	.425	18.06	−.083	−.076	.01	18.07
Response_efficacy	.350	.533	28.41	−.258	−.142	2.02	30.43
Maladaptive_coping	−.541	−.729	53.29	.040	.097	.09	53.38
Fear	.370	.544	29.59	−.414	−.310	9.61	39.20

Note: Structure coefficients (rs) greater than .45 are underlined. Communality coefficients (h^2) greater than 45% are underlined. Coef = standardized canonical coefficient; rs = structure coefficient; rs^2 = squared structure coefficient; h^2 = communality coefficient.

(2) structure coefficients (*rs*), (3) squared structure coefficients (rs^2), which represent the percentage of shared variance between the observed variable and the latent variable created from the observed variable's set, and (4) communality coefficients (h^2) for all variables across both functions. The communality coefficients represent the amount of variance in the observed variable that is explained by the retained functions. These coefficients are simply the sum of the variables squared structure coefficients (rs^2).

17.5.3.4.3 Function 1

17.5.3.4.3.1 Standardized canonical coefficients (Coef) The magnitude of the standardized coefficients represents their relative contribution to the function. Looking at the independent variables, and based on the magnitude of their coefficients, it can be seen that the order of their contribution to the first function is *maladaptive coping* (–.541), *fear* (.370), *response efficacy* (.350), *perceived risk* (.286), *self-efficacy* (.207), and *perceived severity* (–.052). For the set of dependent variables, their order of contribution to the first function is *willing_use* (.634), *likely_use* (.299), *certain_use* (.106), and *intend_use* (.049). However, as mentioned earlier, due to the instability of these standardized coefficients, particularly in the presence of multicollinearity, interpretation of the structure coefficients is considered more appropriate.

17.5.3.4.3.2 Structure coefficients (rs) and squared structure coefficients (rs^2) Table 17.2 shows that the four independent variables with the highest structure loadings are *maladaptive coping* (–.729), *fear* (.544), *response efficacy* (.533), and *perceived risk* (.502), with *self-efficacy* (.425) and *perceived severity* (.271) having made secondary contributions to the latent predictor (protection motivation) variable. Note that the order of magnitude of these coefficients is similar to that found for the standardized coefficients, with the first four independent variables having the larger coefficients. These variables' contribution to the latent independent variable is further borne out by their squared structure coefficients, which indicate the amount of variance these observed variables contributed to the latent variable. In terms of the coefficients' directionality, notice that except for maladaptive coping, all the other variable structure coefficients have the same sign (positive), suggesting that they are all positively related. Maladaptive coping is inversely related to the other protection motivation variables.

For the latent criterion variable (intention for condom use), Table 17.2 results show that all four observed variables returned high structure coefficients (*willing_use* (.959); *likely_use* (.883); *intend_use* (.838); *certain_use* (.814)), indicating that all four variables are primary contributors to the latent criterion variable of *condom use*. Because these variables structure coefficients are positive, they are positively related to all the protection motivation variables except for maladaptive coping. These relationships can be interpreted as follows. Following HIV/AIDS communication, the higher the threat appraisal

(perceived severity and perceived risk), the higher the coping appraisal (response efficacy and self-efficacy), and the greater the fear experienced, the higher is the intention for condom usage. Inversely, the higher the use of maladaptive coping, the lower is the intention for condom usage. As these findings are consistent with protection motivation theory, Function 1 seems to have captured theoretically meaningful relationships as articulated by this motivation model.

17.5.3.4.3.3 Communality coefficients (h^2) Just as these coefficients can be used to assess how useful the variable was for the CCA solution, they can also be used to identify those variables not (or only somewhat) useful in the model. For example, the independent variables of *perceived severity* (h^2 = 10.87%) and *self-efficacy* (h^2 = 18.07%) only made marginal contributions as predictors, suggesting that they may not have been strongly related to the criterion latent variable of *condom use*.

17.5.3.4.4 Function 2 Looking at Function 2, the structure coefficients presented in Table 17.2 show that the only observed variable of relevance in contributing to the latent independent protection motivation variable is *perceived risk* (.780), with a secondary (albeit less so) contribution by *fear* (–.310). For the latent criterion variable of *condom use*, the only noteworthy contributor is the observed variable of *intend_use* (–.493) with a secondary contribution by *certain_use* (–.422). In terms of the coefficients' directionality, it can be seen that the observed independent variable of *perceived risk* is inversely related to the criterion variables *intend_use* and *certain_use*. Because the structure coefficient for *fear* is negative, it is positively related to the criterion variables *intend_use* and *certain_use*. Therefore, the greater the fear experienced, the higher is the intention for condom usage.

18

Structural Equation Modeling

18.1 What Is Structural Equation Modeling (SEM)?

Structural equation modeling (SEM) is a multivariate technique that can best be described as a combination of both factor analysis and path analysis. It is a statistical technique that allows the analyst to examine a series of dependence relationships between **exogenous** and **endogenous** variables simultaneously. Before moving on to explaining the characteristics of SEM, it is necessary to remark on the distinction between exogenous and endogenous variables when used in causal models. According to Pedhazur (1997), an **exogenous** variable is one whose variability is assumed to be determined by causes outside the causal model under consideration. An **endogenous** variable, on the other hand, is one whose variation is to be explained by exogenous and other endogenous variables in the causal model.

The usefulness of SEM in research is distinguished by three characteristics:

1. **It provides a method of dealing with multiple relationships simultaneously.** As shown in Chapter 14, model testing via path analysis can be implemented with a conventional multiple regression technique. That is, path coefficients can be calculated by regressing the endogenous variable (the dependent variable) onto the exogenous variables (the predictor variables), and then repeating the procedure by treating the exogenous variables as endogenous variables. Such a method, however, is typically piecemeal in nature and does not provide information regarding the hypothesized model's goodness-of-fit. Without information about the model's goodness-of-fit, it is difficult to assess the adequacy of the theory underlying the hypothesized model. SEM, on the other hand, is able to estimate the multiple and interrelated dependence relationships simultaneously. Because it test the model as a whole, rather than in a piecemeal fashion, statistics can be calculated to express the goodness-of-fit of the data to the hypothesized model.

2. **It is able to represent unobserved (latent) concepts in the analysis of dependence relationships**. While multiple regression can be used to examine relationships between variables, its usage is limited to the analysis of those variables that can only be directly observed (or measured). SEM, on the other hand, has the ability to incorporate latent (or unobserved) variables in the analysis. A **latent** variable is a hypothesized or unobserved construct and as such, cannot be measured directly. It can only be approximated by observable or measured variables. For example, a researcher wants to investigate the pattern of relationships between three psychological constructs: *aggression*, *authoritarianism*, *and intelligence*. All three constructs cannot be directly observed, and therefore cannot be directly measured. They are "measured" indirectly using various types of scale items, from questionnaires and inventories. On the basis of the responses to these scale items, the magnitude of these latent variables can be calculated. This is similar to factor analysis where highly interrelated clusters of scale items are identified as latent factors. SEM can then be used to estimate paths between latent factors rather than between variables.
3. **It improves statistical estimation by accounting for measurement error in the estimation procedure.** The univariate and multivariate techniques covered in the previous chapters all assume that there is no error associated with the measurement of variables. That is, these techniques assume that variables in the analyses are error-free. Although this is an unrealistic assumption, many if not most researchers, regrettably, act as if this is indeed the case for their data. It is well known from both theoretical and practical perspectives that concepts can seldom be measured perfectly, either because of inaccurate responses by the respondents, or because of problems associated with the operationalization of the concepts. Consequently, measured variables usually contain at least moderate amounts of error, and when such measures are used in univariate and multivariate models (e.g., ANOVA, ANCOVA, MANOVA, multiple regression), the coefficients obtained will be biased, most often in unknown degree and direction. SEM, on the other hand, uses scores on the measured, or manifest, variables to produce estimates of the individual's scores on the underlying construct, or latent variables. As these estimates are derived on the basis of the common or shared variance among the measured variables, scores on the latent variables are unaffected by the random measurement error. Therefore, when these variables are used in SEM analysis, the potential biasing effects of random measurement error on the results are removed. The net result is that the statistical estimation process is improved, because the structural paths between latent variables are relatively free of the unreliabilities of their measurement indicators.

18.2 The Role of Theory in SEM

SEM is theory driven. That is, when a researcher employs SEM to test a theory, there is a substantial demand for justification for the specification of the dependence relationships. Theory provides this justification. For example, a researcher specifies the following model to represent the hypothesized relationships between the three variables of authoritarianism, racism, and aggression.

$$\text{Authoritarianism} \rightarrow \text{Racism} \rightarrow \text{Aggression}$$

For this model, the unidirectional arrows linking the three variables suggest a "causal" flow of influence, with authoritarianism hypothesized to influence levels of racism, and the level of racism hypothesized to influence levels of aggression. When specifying the pattern of relationships between the three variables, specification of the directional relationships (paths) must be guided by theoretical considerations; that is, the specification does not arrive out of thin air. Without theoretical guidance, the fit of the hypothesized model will at best capitalize on chance, and at worst, is nonsensical.

The need for theory to guide the specification process becomes particularly critical when model modifications are made. Because of the flexibility of SEM, the researcher can set many constraints on the model so as to make it "fit" better. But without theoretical guidance, the researcher may "overfit" the model by setting too many constraints. The result may be a model that fits well, but one which is too restricted and with little generalizability. From a practical perspective, a theory-based approach to SEM is a distinct strength of this technique, as it implies a mode of thinking that forces the researcher to specify the theoretical model employed more exactly, testing the theory more precisely, and yielding a more thorough understanding of the data.

18.3 The Structural Equation Model

In its most general form, SEM consists of two parts: the **measurement model** and the **structural equation model**.

- The **measurement model** specifies the rule governing how the latent variables are measured in terms of the observed variables, and it describes the measurement properties of the observed variables.

That is, measurement models are concerned with the relations between observed and latent variables. Such models specify hypotheses about the relations between a set of observed variables, such as ratings or questionnaire items, and the unobserved variables or constructs they were designed to measure.

- The **structural equation model** is a flexible, comprehensive model that specifies the pattern of relationships among independent and dependent variables, either observed or latent. It integrates the strengths of multiple regression analysis, factor analysis, and multivariate ANOVA (MANOVA) in a single model that can be evaluated statistically. Moreover, it permits directional predictions among a set of independent or a set of dependent variables, and it permits modeling of indirect effects.

Of the two models, the structural model is of greater interest to the researcher, because it provides a direct test of the theory of interest. The measurement model is important as it provides a test for the reliability of the observed variables employed to assess the latent variables. A measurement model that offers a poor "fit" to the data suggests that at least some of the observed indicator variables are unreliable, and precludes the researcher from moving to the analysis of the structural model.

18.4 Goodness-of-Fit Criteria

A number of goodness-of-fit measures are available to assess the overall fit of the hypothesized model. Goodness-of-fit measures the extent to which the actual or observed covariance input matrix corresponds to (or departs from) that predicted from the proposed model. Goodness-of-fit measures can be classified into three types: (1) absolute fit measures, (2) incremental fit measures, and (3) parsimonious fit measures. Examples of these criteria are given below. These examples have been chosen not because they represent the "best" indicators of goodness-of-fit, but because they are probably the easiest to understand.

18.4.1 Absolute Fit Measures

These measures determine the degree to which the proposed model predicts (fits) the observed covariance matrix. Some commonly used measures of absolute fit include the chi-square statistic, the goodness-of-fit statistic, and the root mean square error of approximation.

- **Chi-Square Statistic**. The most fundamental measure of overall fit is the likelihood-ratio chi-square (χ^2) statistic, the only statistically based measure of goodness-of-fit available in SEM (Jöreskog and Sörbom, 1993). In applying the chi-square test, the researcher, customarily, wishes to reject the null hypothesis so as to claim support for its alternative, that is, there is a significant difference between the "observed" and the "expected." When used in this way, the larger the chi-square value the "better." However, when used in SEM, the researcher is looking for *nonsignificant differences* between the actual and predicted matrices. As such, the researcher does not wish to reject the null hypothesis and, consequently, the smaller the chi-square value, the better fit of the model. However, the chi-square statistic is very sensitive to departures from multivariate normality of the observed variables and increases as a direct function of sample size. In the case of large samples, the power of the statistical test underlying the SEM approach is very high. With a great deal of statistical power, almost every reasonable model will be rejected if only the chi-square value and the associated probability are considered. Thus, given departures from multivariate normality or larger samples, a proposed model can easily fail to fit the data statistically, even though the discrepancy between the sample covariance matrix and that reproduced by the parameter estimates of the proposed model may be insignificant from a practical standpoint. Given these limitations, the researcher should complement the chi-square measure with other goodness-of-fit measures.
- **Goodness-of-Fit Index (GFI)**. The GFI measures how much *better* the model fits compared with no model at all (Jöreskog and Sörbom, 1989). It is a non-statistical measure ranging from 0 (poor fit) to 1 (perfect fit). While higher values indicate better fit, no threshold levels of acceptability have been established.
- **Root Mean Square Error of Approximation (RMSEA)**. The RMSEA takes into account the error of approximation in the population. It is a measure of *discrepancy per degree of freedom*, and asks the question, "How well would the model, with unknown but optimally chosen values, fit the population covariance matrix if it were available?" (Browne and Cudeck, 1993, pp. 137–138). The value is representative of the *badness-of-fit* when the proposed model is estimated in the population, in that a value of "0" indicates the best fit and higher values indicate worse fit. Values ranging from 0.05 to 0.08 are deemed acceptable; values ranging from 0.08 to 0.10 indicate mediocre fit, and those greater than 0.10 indicate a poor fit (Browne and Cudeck, 1993; MacCallum, Browne, and Sugawara, 1996).

In reporting the RMSEA, MacCallum et al. (1996) suggested that it is also important to report its 90% confidence interval for the population parameter estimated by RMSEA. This interval reflects the degree of uncertainty associated with RMSEA as a point estimate at the 90% level of statistical confidence. Suppose that the estimated RMSEA is .045 with the 90% confidence interval within the range 0 to .15. Because the lower bound of this interval (0) is less than .05, the null hypothesis of close approximate fit is not ruled out. However, the upper bound of the same confidence interval (. 15) exceeds .10, so the hypothesis of poor approximate fit can also not be ruled out. Thus, the RMSEA value of .045 for this example is subject to a fair amount of sampling error because it is just as consistent with the hypothesis of good approximate fit as it is with the hypothesis of poor approximate fit. This type of "mixed" outcome is more likely in smaller samples.

18.4.2 Incremental Fit Measures

These measures compare the proposed model to some baseline model, most often referred to as the null or independence model. In the independence model, the observed variables are assumed to be uncorrelated with each other. The independence model is so severely and implausibly constrained that it would provide a poor fit to any interesting set of data. A number of incremental fit measures have been proposed: Tucker-Lewis Index (TLI); Normed Fit Index (NFI); Relative Fit Index (RFI); Incremental Fit Index (IFI); and Comparative Fit Index (CFI). While the calculations of these fit indices and their underlying assumptions may be somewhat different, they all represent comparisons between the proposed model and a null or independence model. Specifically, they show the improvement achieved by a proposed model over the null model (i.e., a model assuming independence among the variables); they range from 0 (a fit that is no better than the null model) to 1 (a perfect fit).

18.4.3 Parsimonious Fit Measures

In scientific research, theories should be as simple, or parsimonious, as possible. As Bentler and Mooijaart (1989) put it, "models with fewer unknown parameters may be counted as standing a better chance of being scientifically applicable and explainable" (p. 315). Pursuing this line of thought, parsimonious fit measures relate the goodness-of-fit of the proposed model to the number of estimated coefficients required to achieve this level of fit. Their basic aim is to diagnose whether model fit has been achieved by "overfitting" the data with too many coefficients. Their role is mainly to compare models on the basis of some criteria that take parsimony (in the sense of number of parameters to be estimated) as well as fit into account.

- **Parsimonious Normed Fit Index (PNFI)**—The PNFI takes into account the number of degrees of freedom used to achieve a level of fit. Parsimony is defined as attaining higher degrees of fit per degree of freedom used (one degree of freedom per estimated coefficient). Higher values of PNFI are better, and its primary use is in the comparison of models with differing levels of freedom. When comparing between models, differences of .06 to .09 are proposed to be indicative of substantial model differences (Williams and Holahan, 1994).
- **Akaike Information Criterion (AIC)**—The AIC is a comparative measure between models with differing numbers of constructs. AIC values closer to zero indicate better fit and greater parsimony. A small AIC generally occurs when small chi-square values are achieved with less estimated coefficients. This indicates not only a good fit of observed versus predicted covariances, but also a model not prone to "overfitting." In applying this measure to the comparison decision problem, one estimates all models, ranks them according to the AIC criterion, and chooses the model with the smallest value.

18.4.4 Note of Caution in the Use of Incremental Fit Indices as "Rules of Thumb"

By convention, researchers have used incremental fit indices >0.90 as traditional cutoff values to indicate acceptable levels of model fit. This is based on the logic that if a posited model achieves incremental fit indices >0.90, then the model represents a more than 90% improvement over the null or independence model. Put differently, the only possible improvement to the model is less than 10%. The popularity of these indices as tools to evaluate the fit of models in covariance structure analyses lies with their ability to provide such absolute cutoff values that allow researchers to decide whether or not a model adequately fits the data—that has broad generality across different conditions and sample sizes.

In a number of published articles, Marsh and his colleagues (Marsh, Hau, and Wen, 2004; McDonald and Marsh, 1990) have sounded a note of warning about relying on these traditional cutoff values as "rules of thumb" to assess model fit across different research conditions and sample sizes. Their warning is consistent with Hu and Bentler's (1998, 1999) conclusion that "it is difficult to designate a specific cutoff value for each fit index because it does not work equally well with various types of fit indices, sample sizes, estimators, or distributions" (p. 449). Moreover, they argued that high incremental fit indices (>0.90) are not a sufficient basis to establish the validity of interpretations based on the theory underlying the posited model.

Rather, as pointed out by Hu and Bentler (1998), "consideration of other aspects such as the adequacy and interpretability of parameter estimates, model complexity, and many other issues remains critical in deciding on the validity of a model" (p. 450). In their commentary on the dangers of setting cutoff values for fit indices, Marsh et al. (2004) recommended that interpretations of the model fit "should ultimately have to be evaluated in relation to substantive and theoretical issues that are likely to be idiosyncratic to a particular study" (p. 340).

18.5 Model Assessment

In model testing, the model initially specified by the researcher is not assumed to hold exactly in the population and may only be tentative. Its fit to the data is to be evaluated and assessed in relation to what is known about the substantive area, the quality of the data, and the extent to which various assumptions are satisfied. The goal is to derive a model that not only fits the data well from a statistical standpoint, taking all aspects of error into account, but also has the property that every parameter of the model can be given a substantively meaningful interpretation. Jöreskog and Sörbom (1993) proposed examining three classes of information when evaluating a model's goodness-of-fit.

1. **Examine the parameter estimates to determine if there are any unreasonable values or other anomalies.** Parameter estimates should have the right sign and size according to a theory or *a priori* specifications. Examine the squared multiple correlation (SMC) for each relationship in the model. A SMC is an indicator of the amount of variance in the observed indicator variable accounted for by its latent construct. As such, a SMC is a measure of the strength of the linear relationship. A small SMC indicates a weak relationship and suggests that the model is not good.
2. **Examine the measures of the overall fit of the model**. If any of these measures indicate a poor fit of the data, proceed with the detailed assessment of fit with the next step.
3. **The tools for examining the fit in detail** are the *residuals*, *relative residuals*, and *standardized residuals*; the *modification indices*; and the *expected change.* All this information is presented in the various SEM software program outputs (e.g., LISREL, EQS, AMOS), and can be used to identify the source of misspecification in the model, and to suggest how the model should be modified to fit the data better.

18.6 Improving Model Fit

Because the fit of most initial models is deemed unsatisfactory (Jöreskog and Sörbom, 1989), "model modification … has been an inevitable process in the application of covariance structure analysis" (Chou and Bentler, 1993, p. 97). Broadly, model modification, under such circumstances, consists of freeing fixed parameters with the aim of improving the fit of the model. In most instances, parameters are freed sequentially, one at a time, until the researcher is satisfied with the fit of the revised model (Pedhazur, 1997).

18.6.1 Modification Indices

One useful aid in evaluating the fit of a specified model involves modification indices, which are calculated for each non-estimated (fixed and constrained) parameter. Each such modification index measures how much a chi-square value is expected to decrease if a particular constrained parameter is set free (i.e., estimated) and the model is re-estimated. The largest modification index tells us which parameter to set free to improve the fit maximally (Jöreskog and Sörbom, 1993). Associated with each modification index is an **expected parameter change**, which measures the magnitude and direction of change of each fixed parameter, if it is set free. This parameter differs from the modification index in that it does not indicate the change in overall model fit (χ^2); instead it depicts the change in the actual parameter value.

Although modification indices can be useful in assessing the impact of theoretically based model modifications, they should only be used to relax a parameter (with the largest modification index) if that parameter can be interpreted substantively. Model modification must have a **theoretical justification** before being considered, and even then the researcher should be quite skeptical about the changes (MacCullum, 1986). If model respecification is based solely on the values of the modification indices, the researcher is capitalizing on the uniqueness of these particular data, and the answer will most probably be an atheoretical but statistically significant model that has little generalizability and limited use in testing causal relationships (Hair et al., 1998).

18.6.2 Correlated Errors

One form of model modification is the addition of correlated errors to improve fit. Two kinds of correlated errors can be identified: (1) those between error terms of indicators of latent variables (i.e., measurement errors), and (2) those between error terms of latent variables (i.e., residuals) (Pedhazur, 1997). In cross-sectional studies, an important assumption underlying latent variable analysis is that the error terms between indicator variables are uncorrelated. If the error terms for two or more indicators correlate, this means that these

indicators measure something else or something in addition to the construct they are supposed to measure. If this is the case, the meaning of the construct and its dimensions may be different from what is intended.

There are gains and drawbacks in the use of correlated errors of latent variables to improve model fit. The virtue of adding correlated errors of measurement is that such a move can result in dramatic improvements in the overall fit of a model and, on occasion, can reveal unexpected and possible problematic sources of covariance among ratings or items of a measure. The drawback to adding correlated errors of measurement is that such a move is almost always *post hoc* and rarely eventuates in a satisfactory explanation for the correlation. Thus, the likelihood that the correlation is idiosyncratic to the sample and, therefore, not likely to replicate, is disturbingly high (Hoyle and Smith, 1994). As Gerbing and Anderson (1984) put it, "While the role of correlated errors improves the fit by accounting for … unwanted covariation, it does so at a corresponding loss of the meaning and substantive conclusions which can be drawn from the model" (p. 574). Simply, it is a widespread misuse of structural equation modeling to include correlated error terms (whatever the type) in the model for the sole purpose of obtaining a better fit to the data. **Every correlation between the error terms must be justified and interpreted substantively** (Jöreskog and Sörbom, 1993).

18.7 Problems with Estimation

In estimating the fit of a model, there are a number of estimation errors commonly encountered by the researcher. These errors—*nonpositive definite covariance component matrices* and *Heywood cases*—will result in the estimation solution process to be inadmissible.

- *Nonpositive definite covariance component matrices*—The maximum likelihood and least square fit criteria employed by SEM programs such as LISREL, EQS, and AMOS can return estimates outside the range of admissible parameters. Several modeling aspects are associated with indefinite matrix estimates: small sample sizes, listwise deletion of missing observations, the presence of outliers and nonnormalities, and too many parameters in the structural model (Wothke, 1993). For example, Kline (2005) showed that when individual values in a covariance matrix are based on different numbers of cases (resulting from pairwise deletion of missing observations), it is possible that some of the values are mathematically out of range. Kline employed Pearson correlations (which are part of a covariance) to demonstrate this.

Suppose there are three variables—X, Y, and W. The correlation matrix formed by these three variables will be:

	X	Y	W
X	1		
Y	*xy*	1	
W	*xw*	*yw*	1

If R*xw* = .70 and R*yw* = .50, then the value of R*xy* must be within the range:

$$R_{xw}R_{yw} \pm \sqrt{(1 - R^2_{xw})(1 - R^2_{yw})}$$

$$(.70)(.50) \pm \sqrt{(1 - .49)(1 - .25)} \text{ or } \mathbf{.35 \pm .62 (i.e., -.27 \text{ to } .97)}$$

Any other value will be out of bounds. If an out-of-bounds correlation is part of a covariance matrix, the matrix is nonpositive definite or singular, which implies that certain mathematical operations with the matrix such as division will fail because of problems such as denominators that equal zero.

- *Heywood cases*—These are illogical cases that include negative variance estimates or estimated correlation values greater than 1.0. Since the variance is computed as the mean squared deviation from the mean, all variables must have nonnegative variances. Likewise, since correlations are standardized covariances, correlation coefficients are bounded within the range ±1.00. That is, they cannot exceed ±1.00. Heywood cases can be caused by specification errors, the presence of outliers, a combination of small sample sizes ($N <$ 100) and only two indicators per factor, bad start values, or extremely high or low population correlations (Kline, 2005).

18.8 Checklist of Requirements

- **Sample Size**. As a test of model fit, the use (and validity) of the chi-square test is predicated on the viability of various assumptions, one which specifies that the "sample size is sufficiently large" (Jöreskog and Sörbom, 1993, p. 122). Unfortunately, there is a lack of agreement about the meaning of "sufficiently large." While there is no single criterion that dictates the necessary sample size, Hair et al. (1998) suggested that the absolute minimum sample size must be at least greater than the number of covariances in the input data

matrix. The most appropriate minimum ratio is 10 respondents per parameter, with an increase in the sample size as model complexity increases. Thus, a path model with 20 parameters should have a minimum sample size of 200 cases.

- **Number of Indicator Variables.** A major advantage of using multiple indicators in SEM is that they afford the study of relations among latent variables uncontaminated by errors of measurement in the indicators. This is predicated, among other things, on judicious choices of indicators or manifest variables (Pedhazur, 1997). The minimum number of indicators for a construct is one, and aside from the theoretical basis that should be used to select variables as indicators of a construct, there is no upper limit in terms of the number of indicators. However, as pointed out by Bentler (1980), in practice, too many indicators make it difficult if not impossible to fit a model to data. As a practical matter, three is the preferred minimum number of indicators, and in most cases, five to seven indicators should represent most constructs (Hair et al., 1998). In selecting the number of indicators, researchers should be guided by the axiom that it is preferable to employ a relatively small number of "good" indicators than to delude oneself with a relatively large number of "poor" ones (Pedhazur, 1997).

18.8.1 Item Parcels

Instead of using individual indicators to represent latent variables, a common practice involves creating item parcels on the basis of sums of responses to individual items and then using scores on these parcels in the latent variable analysis. For example, suppose that nine items were written to measure the latent variable of *resilience*. On the basis of a factor analysis of these nine measures, divide the items into three parcels, and then sum the items in each parcel to form three measured variables to operationalize the latent variable. Following the procedure described by Russell et al. (1998), the development of these item parcels involves the following steps:

1. Fit a one-factor model for the nine items assessing resilience.
2. Rank-order items on the basis of their loadings on this factor.
3. Assign items to parcels so as to equate the average loadings of each parcel of items on the factor.

Specifically, assign items ranked 1, 5, and 9 to parcel 1; items ranked 2, 6, and 8 to parcel 2; and items ranked 3, 4, and 7 to parcel 3. With this procedure, the resulting item parcels should reflect the underlying construct of resilience to an equal degree.

The utilization of item parcels in latent variable analysis is supported by a number of reasons (Russell et al., 1998). First, responses to individual items are likely to violate the assumptions of multivariate normality that underlie the maximum likelihood estimation procedure often used in estimating structural equation models with latent variables. Second, analyses using individual items as measured indicators for the latent variables often necessitate estimating a large number of parameters (i.e., factor loadings and error terms) in fitting the model to the data. This, in turn, necessitates the use of a large sample, given the recommendation that approximately 10 cases per parameter be used in testing structural equation models. Finally, by using parcels rather than individual items, the results of the analysis are not likely to be distorted by idiosyncratic characteristics of individual items. One consequence of analyzing parcels rather than individual items is that the overall fit of the model to the data is improved. This effect is due to improvements in the distribution of the measured variables and to fewer parameters being estimated as a consequence of a simpler measurement model.

18.9 Assumptions

- **Independence**—Observations are independent of each other.
- **Random sampling**—Random sampling of respondents.
- **Linearity**—Linearity of relationships between exogenous and endogenous variables.
- **Multivariate normality**—Distribution of observed variables is multivariate normal. A lack of multivariate normality is particularly troublesome because it substantially inflates the chi-square statistic and creates an upward bias in critical values for determining coefficient significance (Muthuen and Kaplan, 1985; Wang, Fan, and Wilson, 1996).

18.10 Examples of Structural Equation Modeling

The following examples will be demonstrated using the software program **AMOS version 20**. These examples assume that the researcher already has some experience with this software program, and in particular, in the use of its powerful graphical interface. For those researchers who have not used AMOS before, suffice it to say that this program is extremely easy to use and to master. Its appeal lies in its approach of

using a path diagram (via a graphical user interface) to specify a model that the researcher wants to test. Drawing path diagrams to represent hypothesized models is a perfectly natural approach to structural equation modeling. For the beginner, reading **AMOS User's Guide version 19** is highly recommended.

A free student version of **AMOS 5.0** and **AMOS User's Guide version 19** can be downloaded from **http://amosdevelopment.com/download/**. This student version contains all the graphical tools and goodness-of-fit indices found in the full version, but is limited to only eight measurement variables.

18.11 Example 1: Linear Regression with Observed Variables

Ho (1998, 1999) investigated euthanasia and the conditions under which people are most likely to support euthanasia. Items were written to represent four factors, three of which related to conditions of suffering, and one related to the extent of support for voluntary euthanasia. These four factors, together with their representative items, are listed below.

- **BODY—The debilitated nature of a person's body.**

  C1—You have lost control of all your bodily functions.

  C7—The significant person in your life has lost control of all his/her bodily functions.

  C13—A person has lost control of all his/her bodily functions.
- **FAMILY—The perceived negative impact that the terminal illness of a family member has on his/her family.**

  C5—Your terminal illness will cause you to be a burden on your family.

  C11—The terminal illness of the significant person in your life will cause him/her to be a burden on his/her family.

  C17—This person's terminal illness will cause him/her to be a burden on his/her family.
- **PAIN—The experience of physical pain.**

  C3—Your terminal illness has given you continuous excruciating pain.

  C9—The terminal illness of the significant person in your life has given him/her continuous excruciating pain.

  C15—This person's terminal illness has given him/her continuous excruciating pain.

- **VOLUNTARY EUTHANASIA—The extent of support for voluntary euthanasia.**

  E1—Doctors have the right to administer medication that will painlessly end the life of a terminally ill person, if he/she requests it.
  E2—Terminally ill patients have the right to decide that life-sustaining drugs or mechanisms be withheld or withdrawn, to hasten their death.
  E5—Terminally ill patients have the right to decide about their own lives and deaths.

18.11.1 Data Entry Format

The variance-covariance matrix is generated from the SPSS file **EUTHAN.SAV.**

Variables	Column(s)	Code
• E1 to E12	1–12	1 = strongly disagree, 5 = strongly agree
• C1 to C18	13–30	1 = strongly disagree, 5 = strongly agree
• GENDER	31	1 = male, 2 = female

18.11.2 Modeling in AMOS Graphics

This example will demonstrate a conventional regression analysis, predicting a single observed variable (**E1**: extent of support for voluntary euthanasia) as a linear combination of three other observed variables (**C3**, **C9**, **C15**: the experience of physical pain).

- **Regression model.** Figure 18.1 presents the regression model to be tested. This model was drawn using the icons displayed in the toolbox of the AMOS 20 Graphics main window.

  This model evaluates how scores on C3, C9, and C15 (the experience of physical pain) are related to support for E1 (voluntary euthanasia). As it is not assumed that E1 will be perfectly predicted by a linear combination of C3, C9, and C15, the model includes an error term (er1). The

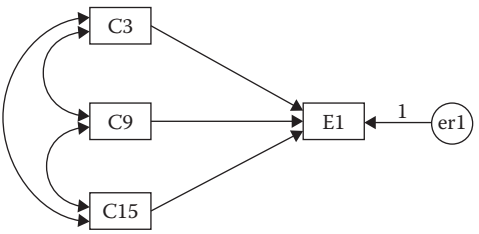
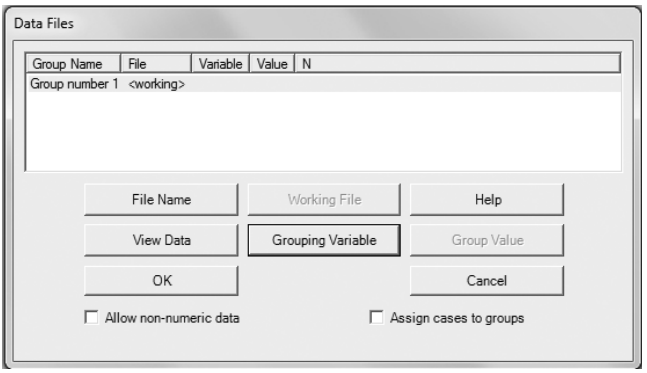
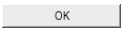


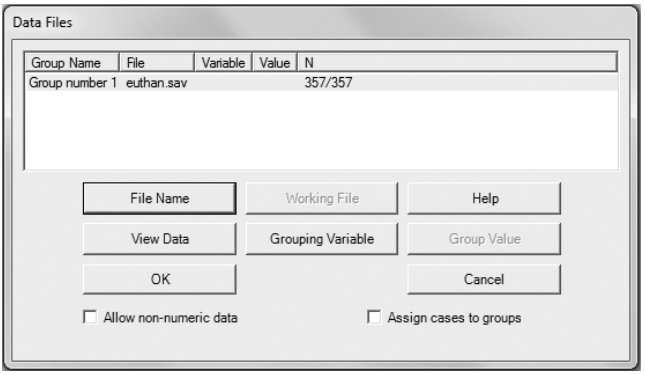
FIGURE 18.1
Conventional linear regression with observed variables.

variable er1 is enclosed in a circle because it is not directly observed. The single-headed arrows represent linear dependencies, and the double-headed arrows connecting the predictor variables suggest that these variables may be correlated with each other. Notice that the path coefficient from er1 to E1 is *fixed* to 1. Fixing the path coefficient to unity is necessary for the model to be identified. Without this constraint, there is not enough information to estimate the regression weight for the regression of E1 on er1 and the variance of er1 at the same time.

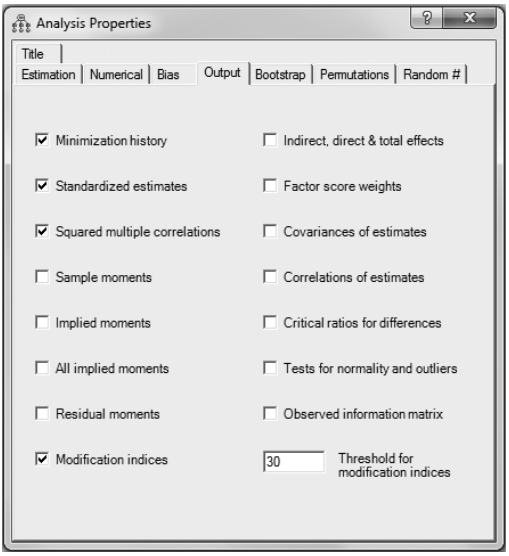
- **Linking the model to the data set.** Once the model has been drawn, the next step is to link the model to the data set (EUTHAN.SAV) to be analyzed.
  1. Click  from the icon toolbox (**File → Data Files…**) to open the **Data Files** window below.



  2. Click  to search for the data file (**EUTHAN.SAV**) in the computer's directories. Once the data file has been located, open it. By opening the data file, AMOS will automatically link the data file to the regression model, as shown in the window below. Click  to exit the **Data Files** window.



- **Results output.** In order to obtain specific results necessary for interpretation of the AMOS output, click the **Analysis Properties** icon from the icon toolbox (from the menu bar, click **View → Analysis Properties…**) to open the **Analysis Properties** window below. Click the **Output** tab to open the **Output** page. Check the statistics boxes to get the statistics required in the results output. In the present instance, check the following boxes: **Minimization history, Standardized estimates, Squared multiple correlations,** and **Modification indices.** Type the number **30** in the box labeled **Threshold for modification indices** (this ensures that only those modification indices with values greater than or equal to 30 will be presented in the AMOS output). Close this window when finished.



- **Conducting the analysis.** From the AMOS icon toolbox, click (from the menu bar, click **Analyze → Calculate Estimates**) to conduct the analysis.
- To view the analysis output, click (from the menu bar, click **View → Text output**).

18.11.3 Results and Interpretation

Recall that in SEM, the goodness-of-fit of the hypothesized model is indicated by the non-significant difference between the observed and predicted covariance matrices. As such, the smaller the chi-square value (non-significant), the better the fit of the model. For this example, the chi-square value is 0, which means that the fit of the model cannot be calculated (see Table 18.1).

TABLE 18.1

Computation of Degrees of Freedom and Chi-Square Statistics for Goodness-of-Fit

Parameter Summary (Group Number 1)						
	Weights	**Covariances**	**Variances**	**Means**	**Intercepts**	**Total**
Fixed	1	0	0	0	0	1
Labeled	0	0	0	0	0	0
Unlabeled	3	3	4	0	0	10
Total	4	3	4	0	0	11

Models
Default model (Default model)
Notes for Model (Default model)
Computation of degrees of freedom (Default model)

Number of distinct sample moments:	10
Number of distinct parameters to be estimated:	10
Degrees of freedom (10 – 10):	0

Result (Default model)
Minimum was achieved.
Chi-square = .000
Degrees of freedom = 0
Probability level cannot be computed.

The reason why the fit of the model cannot be calculated lies with the number of degrees of freedom. In SEM**, degrees of freedom** are the number of non-redundant variance and covariance moments in the input covariance matrix *minus* the number of estimated coefficients. Each estimated coefficient "uses up" a degree of freedom. A model can never estimate more coefficients than the number of covariances, meaning that zero is the lower bound for the degrees of freedom for any model. In the present example, the input covariance matrix generated from the four observed variables (C3, C9, C15, E1) contains four variances and six covariances, for a total of 10 sample moments. For the hypothesized model (Figure 18.1), there are three regression paths, four variances, and three covariances, for a total of 10 parameters that must be estimated. Hence, the model has zero degrees of freedom (10 sample moments minus 10 estimated parameters), meaning there is not enough information in the observed covariance matrix to compute the chi-square goodness-of-fit statistic. Such a model is often called **saturated** or **just identified**.

18.11.3.1 Regression Weights, Standardized Regression Weights, and Squared Multiple Correlations

The **regression weights** are the unstandardized coefficient estimates generated from maximum likelihood procedure (see Table 18.2).

TABLE 18.2

Regression Weights, Standardized Regression Weights, and Squared Multiple Correlations

Regression Weights: (Group Number 1—Default Model)

			Estimate	S.E.	C.R.	*P*	Label
e1	←	c3	.478	.072	6.656	***	
e1	←	c9	–.020	.084	–.240	.810	
e1	←	c15	.229	.082	2.791	.005	

Estimates (Group number 1—Default model)
Scalar estimates (Group number 1—Default model)
Maximum likelihood estimates

Standardized Regression Weights: (Group Number 1—Default Model)

			Estimate
e1	←	c3	.457
e1	←	c9	–.019
e1	←	c15	.212

Covariances: (Group Number 1—Default Model)

			Estimate	S.E.	C.R.	*P*	Label
c9	↔	c15	.955	.080	11.914	***	
c3	↔	c15	.884	.079	11.254	***	
c3	↔	c9	.926	.081	11.472	***	

Correlations: (Group Number 1—Default Model)

			Estimate
c9	↔	c15	.814
c3	↔	c15	.743
c3	↔	c9	.766

Variances: (Group Number 1—Default Model)

	Estimate	S.E.	C.R.	*P*	Label
c3	1.227	.092	13.342	***	
c9	1.191	.089	13.342	***	
c15	1.154	.086	13.342	***	
er1	.836	.063	13.342	***	

Squared Multiple Correlations: (Group Number 1—Default Model)

	Estimate
e1	.378

Associated with each estimated **unstandardized regression** coefficient (in the **Regression Weights** table) are a **standard error** (S.E.) and a **critical ratio** (C.R.) value. The **standard error** of the coefficients represents the expected variation of the estimated coefficients, and is an index of the "efficiency" of the predictor variables in predicting the endogenous variable; the smaller the S.E. the more efficient the predictor variable is. The **critical ratio** is a test of the significance of the path coefficients. Each C.R. value is obtained by dividing that parameter estimate by its respective standard error, and it is distributed approximately as *z*. As such, a critical ratio that is more extreme than ±1.96 indicates a significant path ($p < .05$). On the basis of this criterion, it can be seen that the variables C3 and C15 are highly significant predictors of E1 (C.R. = 6.65, $p < .001$; C.R. = 2.79, $p < .01$, respectively).

The standardized regression weights (β) are standardized coefficient estimates, and are independent of the units in which all variables are measured. These standardized coefficients allow the researcher to compare directly the relative relationship between each independent variable and the dependent variable. From Table 18.2, it can be seen that ratings on the two variables C3 and C15 (both written to measure the evaluation of pain) are both significantly and positively related to E1 (extent of support for voluntary euthanasia) ($\beta = 0.46$; $\beta = 0.21$, respectively). Therefore, it can be concluded that the greater the perception of pain experienced by oneself and by a nondescript person, the greater is one's support for voluntary euthanasia.

The **covariances** (unstandardized correlation coefficients) between the three predictor variables are all highly significant by the C.R. test ($p < .001$). The standardized **correlation** coefficients are all positive and ranged from 0.74 (C3 ↔ C15) to 0.81 (C15 ↔ C9).

The squared multiple correlation is an index of the proportion of the variance of the endogenous variable (E1) that is accounted for by the exogenous or predictor variables. It can be assumed that the higher the value of the squared multiple correlation, the greater the explanatory power of the regression model, and thus the better the prediction of the dependent variable. In the present example, the predictor variables C3, C9, and C15 accounted for 0.378 or 37.8% of the variance of E1. As such the residual or the amount of unexplained variance (er1) for this model (support for active euthanasia) is 0.622 or 62.2% (calculated as 1 – square multiple correlation).

18.12 Example 2: Regression with Unobserved (Latent) Variables

In the previous example, the regression model incorporated only observed (i.e., measured) variables. However, these observed variables (C3, C9, C15, E1) are attitudinal variables and their measurement would, therefore, be unreliable to some degree. Unreliability of predictor variables, in particular,

can be problematic as it can lead to biased regression estimates. The present example demonstrates the use of unobserved (latent) variables in a similar regression analysis. The use of unobserved variables (rather than observed variables) allows the researcher to incorporate the reliabilities of the measurement variables into the regression analysis, resulting in more accurate estimates.

In the previous example, support for voluntary euthanasia was measured by one observed variable (E1), and the experience of pain (the predictor) was measured by three observed variables (C3, C9, C15). In the present example, the latent construct of *support for voluntary euthanasia* will be measured by the three observed variables E1, E2, and E5, and the latent construct of *pain* will be measured by the three observed variables C3, C9, and C15 (see Section 18.11 for a description of these variables). The regression model incorporating these two latent constructs and their respective measurement indicators is presented in Figure 18.2.

This model evaluates how the experience of pain (represented by the latent construct PAIN) predicts the extent of support for voluntary euthanasia (represented by the latent construct VOLUNTARY EUTHANASIA). As it is not assumed that the extent of one's support for voluntary euthanasia will be perfectly predicted by one's experience of pain, this dependent variable includes a residual (z1).

With nine unobserved variables in this model, it is certainly not identified. That is, there is not enough information to estimate all of the model's parameters simultaneously. For the model to be identified, it will be necessary to fix the unit of measurement of each latent variable to unity (e.g., the parameters between C3 and the latent variables PAIN and er1 are fixed to 1). If a latent variable has more than one parameter (single-headed arrow leading out from it), then fixing any one of them to unity will usually suffice. In this example, the latent variable PAIN has three paths leading away from it, and therefore only one of these paths (between C3 and PAIN) has been constrained up to unity. All paths connecting the error components (er1 to er6) to their respective indicators have also been set to unity.

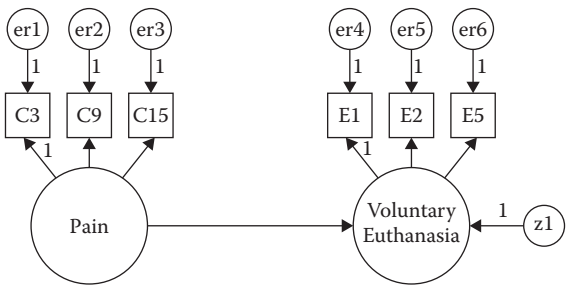


FIGURE 18.2
Regression with unobserved variables predicting support for voluntary euthanasia.

The same variance-covariance matrix used in Example 1 (generated from the SPSS file EUTHAN.SAV) will be used to test the present model.

The procedure involved in (i) **linking the model to the data set,** (ii) obtaining the desired statistics in the **results output**, and (iii) **conducting the analysis**, is identical to that described in Section 18.11.2 (**Modeling in AMOS Graphics**) above.

18.12.1 Results and Interpretation

The input covariance matrix generated from the model's six observed variables contains 21 sample moments. For the hypothesized model (Figure 18.2), there are five regression weights and eight variances, for a total of 13 parameters to be estimated. The model, therefore, has positive degrees of freedom (21 – 13 = 8), and the chi-square goodness-of-fit statistic was computed. The result indicates that the model did not fit the data well by the chi-square test, $\chi^2(N = 357, df = 8) = 36.71$, $p < .05$ (see Table 18.3).

Although the hypothesized model did not fit the observed variance-covariance matrix well by the chi-square test, the baseline comparisons fit indices of NFI, RFI, IFI, TLI, and CFI are all above 0.9 (range: 0.949–0.978) (see Table 18.4).

TABLE 18.3

Computation of Degrees of Freedom and Chi-Square Statistics for Goodness-of-Fit

Parameter Summary (Group Number 1)						
	Weights	**Covariances**	**Variances**	**Means**	**Intercepts**	**Total**
Fixed	9	0	0	0	0	9
Labeled	0	0	0	0	0	0
Unlabeled	5	0	8	0	0	13
Total	14	0	8	0	0	22

Notes for group (Group number 1)
The model is recursive.
Sample size = 357
Notes for model (Default model)
Computation of degrees of freedom (Default model)

Number of distinct sample moments:	21
Number of distinct parameters to be estimated:	13
Degrees of freedom (21 – 13):	8

Result (Default model)
The minimum was achieved.
Chi-square = 36.706
Degrees of freedom = 8
Probability level = .000

TABLE 18.4

Incremental Fit Indices

	Baseline Comparisons				
Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.973	.949	.979	.960	.978
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

These indices compare the fit of the hypothesized model to the null or independence model. Although there are no clearly established rules as to what constitutes a good fit, a widely applied guideline for these incremental fit indices is 0.90 (Bentler, 1980; Bentler and Bonett, 1980). With the incremental fit indices ranging from 0.949 to 0.978, the possible improvement in the fit of the hypothesized model (range: 0.022–0.051) appears so small as to be of little practical significance.

In employing the incremental fit indices to evaluate model fit, it is necessary to take into account the warning offered by Marsh et al. (2004) regarding the use of these indices as "rules of thumb." Specifically, the conventional cutoff value of >.90 may not work equally well with various types of fit indices, sample sizes, estimators, or distributions. Therefore, deciding on the validity of a model must take into consideration other aspects of the model such as the adequacy and interpretability of parameter estimates, model complexity, and most importantly, the substantive and theoretical issues underlying the posited model (Hu and Bentler, 1998; Marsh et al., 2004).

18.12.1.1 Regression Weights, Standardized Regression Weights, and Squared Multiple Correlations

In the **Regression Weights** table (see Table 18.5), the results indicate that the unstandardized regression weights are all significant by the critical ratio test ($>\pm 1.96$, $p < .05$) (except for those parameters fixed to 1). In the **Standardized Regression Weights** table, the results indicate that the experience of pain is significantly and positively related to the support for voluntary euthanasia (standardized regression weight: $\beta = 0.72$, $p < .001$). Thus, the greater the perception of pain experienced, the stronger the reported support for voluntary euthanasia. The results also indicate that the six observed measurement variables are significantly represented by their respective latent constructs ($p < .001$).

The **squared multiple correlations** show that 0.519 or 51.9% of the variance of support for VOLUNTARY EUTHANASIA is accounted for by the variance in PAIN. The remaining 0.481 or 48.1% of the variance of support for VOLUNTARY EUTHANASIA cannot be explained by the model, and is thus attributed to the unique factor z1 (residual).

TABLE 18.5

Regression Weights, Standardized Regression Weights, and Squared Multiple Correlations

			Estimate	S.E.	C.R.	*P*	Label
Voluntary_ euthansia	←	Pain	.722	.058	12.413	***	
C3	←	Pain	1.000				
C9	←	Pain	1.035	.047	21.931	***	
C15	←	Pain	1.012	.047	21.728	***	
E1	←	Voluntary_ euthansia	1.000				
E2	←	Voluntary_ euthansia	.599	.044	13.517	***	
E5	←	Voluntary_ euthansia	.787	.052	15.263	***	

Scalar estimates (Group number 1—Default model)
Maximum likelihood estimates
Regression weights: (Group number 1—Default model)

Standardized Regression Weights: (Group Number 1—Default Model)

			Estimate
Voluntary_euthansia	←	Pain	.720
C3	←	Pain	.855
C9	←	Pain	.898
C15	←	Pain	.892
E1	←	Voluntary_euthansia	.818
E2	←	Voluntary_euthansia	.714
E5	←	Voluntary_euthansia	.813

Variances: (Group Number 1—Default Model)

	Estimate	S.E.	C.R.	*P*	Label
Pain	.896	.091	9.820	***	
z1	.433	.059	7.321	***	
er1	.331	.033	9.978	***	
er2	.231	.029	8.081	***	
er3	.237	.028	8.417	***	
er4	.445	.053	8.438	***	
er5	.309	.028	10.882	***	
er6	.287	.033	8.612	***	

Squared Multiple Correlations: (Group Number 1—Default Model)

	Estimate
Voluntary_euthansia	.519
E5	.660
E2	.510

TABLE 18.5 (*Continued*)

Regression Weights, Standardized Regression Weights, and Squared Multiple Correlations

Squared Multiple Correlations: (Group Number 1—Default Model)	
	Estimate
E1	.669
C15	.795
C9	.806
C3	.730

18.12.1.2 Comparing the Latent-Construct Model (Example 2) with the Observed-Measurement Model (Example 1)

Note that the present example employed latent constructs to test the hypothesis that the experience of pain influences the extent of support for voluntary euthanasia. This model differs from the model tested in Example 1, which employed only observed (measured) variables. Comparing the squared multiple correlation for the latent construct of "support for voluntary euthanasia" in this latent-construct model to the squared multiple correlation for the measurement variable of E1 (support for voluntary euthanasia) in the observed-variable model tested in Example 1, it can be seen that the amount of variation accounted for by this latent-construct model is greater than the amount of variance accounted for in E1 by the observed-variable model (51.9% versus 37.8%). This demonstrates the greater explanatory

TABLE 18.6

Standard Error Estimates for Latent-Construct Model and Observed-Variable Model

Latent-Construct Model			
			S.E.
Voluntary_euthansia	←	Pain	**.058**
C3	←	Pain	
C9	←	Pain	**.047**
C15	←	Pain	**.047**
E1	←	Voluntary_euthansia	
E2	←	Voluntary_euthansia	**.044**
E5	←	Voluntary_euthansia	**.052**

Observed-Variable Model			
			S.E.
e1	←	c3	**.072**
e1	←	c9	**.084**
e1	←	c15	**.082**

power of the latent-construct analysis approach over the observed-variable analysis approach in predicting the dependent variable.

Comparing the standard errors associated with the parameter estimates for these two models (see Table 18.6), it can be determined that the standard error estimates obtained under the present latent-construct model are all smaller than the estimates obtained under the previous observed-variable model.

Hence, the present latent-construct model parameter estimates are more efficient (assuming that the model is correct), and are to be preferred over the ones from the observed-variable model (Example 1).

18.13 Example 3: Multi-Model Path Analysis with Latent Variables

This example demonstrates a path analysis with latent constructs to investigate the **direct** and **indirect** structural relationships between the exogenous and endogenous variables in a hypothesized path model. This model is based on the same variance-covariance matrix (generated from the SPSS file EUTHAN.SAV) employed in Examples 1 and 2.

The present example, based on Ho's (1999) study, will assess the relationships between three predictor variables— (1) *the debilitated nature of one's body*, (2) *the burden placed on one's family*, and (3) *the experience of physical pain*—and the dependent variable of *support for voluntary euthanasia*. The hypothesized direct and indirect relationships between the three predictor variables and the dependent variable are represented in the model shown in Figure 18.3.

The model hypothesizes that assessment of the debilitated nature of one's body (**BODY**), and the extent to which one's illness is perceived to be a burden on one's family (**FAMILY**), will be related to one's decision to support voluntary euthanasia (**VOLUNTARY EUTHANASIA**), both directly and indirectly, being mediated by one's assessment of the physical pain (**PAIN**) experienced. The description of the 12 items written to represent the four factors **BODY** (debilitated nature of one's body—C1, C7, C13), **FAMILY** (burden on one's family—C5, C11, C17), **PAIN** (experience of physical pain—C3, C9, C15), and **VOLUNTARY EUTHANASIA** (E1, E2, E5) are presented in Section 18.11.

18.13.1 Evaluation of the Measurement Model: Confirmatory Factor Analysis (CFA)

Before evaluating the fit of the path model presented in Figure 18.3, it is necessary to specify a measurement model to verify that the 12 measurement variables written to reflect the four unobserved constructs (BODY,

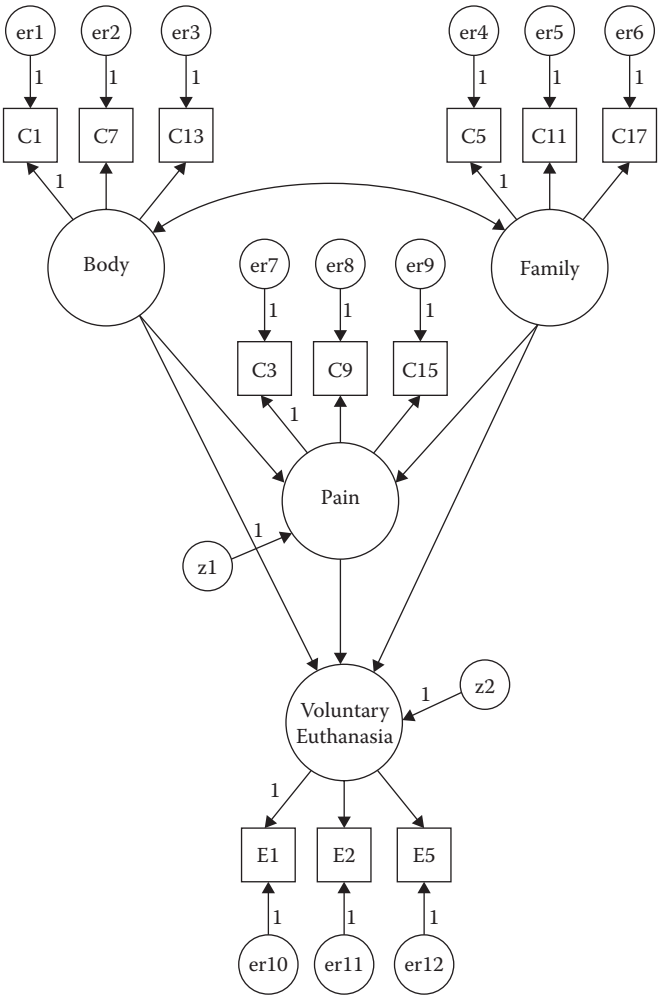


FIGURE 18.3
Path model for the prediction of voluntary euthanasia.

FAMILY, PAIN, VOLUNTARY EUTHANASIA) do so in a reliable manner. The overall fit of a measurement model is determined by a confirmatory factor analysis (CFA). The fit of this model is extremely important in that all possible latent-variable structural models are nested inside it. Obtaining a poor fit at this stage indicates a need for further refinement of the measurement model and precludes moving on to investigate latent-variable structural models (Anderson and Gerbing, 1988).

For this phase of the analysis, CFA is carried out to determine the degree of model fit, the adequacy of the factor loadings, and the standardized residuals

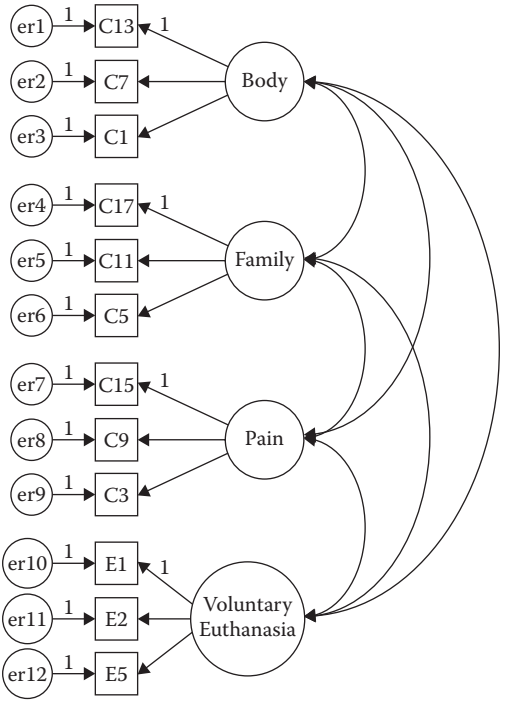


FIGURE 18.4
Measurement model.

and explained variances for the measurement variables. Figure 18.4 presents the measurement model for this example.

For this constructed measurement model, all factor loadings are freed (i.e., estimated); items are allowed to load on only one construct (i.e., no cross-loading); and latent constructs are allowed to correlate (equivalent to oblique rotation in exploratory factor analysis).

The procedure involved in (i) **linking the model to the data set**, (ii) obtaining the desired statistics in the **results output**, and (iii) **conducting the analysis** is identical to that described in Section 18.11.2.

18.13.2 Results and Interpretation

The input covariance matrix generated from the model's 12 measurement variables contains 78 sample moments. For the measurement model, there are 8 regression weights, 6 covariances, and 16 variances, for a total of 30 parameters to be estimated. The model, therefore, has 48 degrees of freedom (78 – 30), and the chi-square goodness-of-fit statistic was computed (see Table 18.7).

The chi-square goodness-of-fit test shows that the model did not fit the data well, $\chi^2(N = 357, df = 48) = 253.472$, $p < .05$. Although the model did not fit well by the chi-square test, the baseline comparisons fit indices NFI, RFI, IFI,

TABLE 18.7

Computation of Degrees of Freedom and Chi-Square Statistics for Goodness-of-Fit

Parameter Summary (Group Number 1)						
	Weights	**Covariances**	**Variances**	**Means**	**Intercepts**	**Total**
Fixed	16	0	0	0	0	16
Labeled	0	0	0	0	0	0
Unlabeled	8	6	16	0	0	30
Total	24	6	16	0	0	46

Notes for group (Group number 1)
The model is recursive.
Sample size = 357
Notes for model (Default model)
Computation of degrees of freedom (Default model)

Number of distinct sample moments:	78
Number of distinct parameters to be estimated:	30
Degrees of freedom (78 – 30):	48

Result (Default model)
The minimum was achieved
Chi-square = 253.472
Degrees of freedom = 48
Probability level = .000

TABLE 18.8

Incremental Fit Indices

	Baseline Comparisons				
Model	**NFI Delta1**	**RFI rho1**	**IFI Delta2**	**TLI rho2**	**CFI**
Default model	.924	.895	.937	.914	.937
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

TLI, and CFI are close to or exceed 0.9 (range: 0.89–0.94) (see Table 18.8). Given the range of the computed baseline comparisons fit indices, the remaining possible improvement in the fit of the hypothesized model (range: 0.06–0.11) appears so small as to be of little practical significance.

18.13.2.1 Regression Weights and Standardized Regression Weights

The unstandardized regression weights are all significant by the critical ratio test ($>\pm 1.96$, $p < .05$) (see Table 18.9). The standardized regression weights range from 0.706 to 0.921. These values indicate that the 12 measurement variables are significantly represented by their respective latent constructs.

TABLE 18.9

Regression Weights and Standardized Regression Weights

Regression Weights: (Group Number 1—Default Model)

			Estimate	S.E.	C.R.	*P*	Label
C1	←	BODY	1.042	.044	23.508	***	
C7	←	BODY	1.040	.039	26.823	***	
C13	←	BODY	1.000				
C5	←	FAMILY	.969	.060	16.034	***	
C11	←	FAMILY	1.025	.049	21.052	***	
C17	←	FAMILY	1.000				
C3	←	PAIN	.999	.045	22.183	***	
C9	←	PAIN	1.021	.043	23.662	***	
C15	←	PAIN	1.000				
E5	←	Voluntary_euthanasia	.781	.050	15.499	***	
E2	←	Voluntary_euthanasia	.588	.044	13.471	***	
E1	←	Voluntary_euthanasia	1.000				

Standardized Regression Weights: (Group Number 1—Default Model)

			Estimate
C1	←	BODY	.862
C7	←	BODY	.921
C13	←	BODY	.907
C5	←	FAMILY	.731
C11	←	FAMILY	.898
C17	←	FAMILY	.879
C3	←	PAIN	.861
C9	←	PAIN	.893
C15	←	PAIN	.889
E5	←	Voluntary_euthanasia	.813
E2	←	Voluntary_euthanasia	.706
E1	←	Voluntary_euthanasia	.824

18.13.2.2 Explained Variances and Residual Variances

The explained variances of the 12 measurement variables are represented by their **squared multiple correlations** (see Table 18.10). The percentage of variance explained ranges from 0.498 or 49.8% (E2) to 0.848 or 84.8% (C7). The residual (unexplained) variances are computed by subtracting each explained variance from 1 (i.e., 1 – squared multiple correlation). Thus, for the 12 measurement variables, the residual variances range from 15.2% to 50.2%.

18.13.2.3 Modification Indices

Examination of the **modification indices** suggests that the fit of the model can be improved substantially by allowing the error terms er1 (associated with the measurement variable **C13**) and er7 (associated with the measurement variable **C15**) to correlate (see Table 18.11).

TABLE 18.10

Explained Variances (Squared Multiple Correlations) for the 12 Measurement Variables

Squared Multiple Correlations: (Group number 1—Default model)	
	Estimate
E1	.678
E2	.498
E5	.660
C15	.791
C9	.798
C3	.742
C17	.773
C11	.807
C5	.534
C13	.823
C7	.848
C1	.743

TABLE 18.11

Modification Indices

Covariances: (Group Number 1—Default Model)				
			M.I.	**Par change**
er9	↔	Voluntary_euthanasia	14.183	.101
er9	↔	Pain	10.368	–.070
er4	↔	er7	18.018	.076
er4	↔	er8	21.856	–.084
er5	↔	er8	18.232	.075
er6	↔	er7	10.964	–.086
er6	↔	er9	28.331	.151
er1	↔	er7	**58.981**	.133
er1	↔	er8	17.942	–.074
er1	↔	er9	10.241	–.061
er1	↔	er4	15.178	.067
er1	↔	er5	12.440	–.059
er2	↔	er7	13.963	–.064
er2	↔	er8	16.948	.072
er2	↔	er4	11.948	–.059
er2	↔	er5	31.138	.093
er3	↔	er7	14.623	–.081
er3	↔	er9	23.425	.112

As can be seen from the table, allowing these two error terms to correlate will reduce the chi-square value of the modified model by at least 58.981. While this is a substantial decrease (for the loss of 1 degree of freedom), the decision to implement or not to implement this modification rests with the researcher, and in particular, on the **theoretical justification** for this modification. As mentioned earlier, without strong theoretical justification, employing the values of the modification indices to improve model fit increases the probability that the researcher is capitalizing on the uniqueness of the particular data set, and the results will most likely be atheoretical.

For the present case, the motivation to include correlated errors in a modified model is twofold. First, allowing the error terms of er1 and er7 to correlate will reduce the chi-square goodness-of-fit value substantially (i.e., improving the model fit). While this reason does not lower the probability that the strategy may improve the fit of the model by capitalizing on chance, it does have a legitimate place in exploratory studies (Arbuckle and Wothke, 1999). Second, and more significantly, the two measurement variables (C13, C15) associated with the error terms of er1 and er7 appear to share something in common, above and beyond the latent constructs they were written to represent. Both items C13 (*A person has lost control of all his/her bodily functions*) and C15 (*This person's terminal illness has given him/her continuous excruciating pain*) appear to reflect the physical pain commonly associated with a debilitated body.

18.13.3 The Modified Model

This modification was carried out and the modified model was re-estimated. The correlation between two error terms is achieved in AMOS Graphics by joining the error terms with a double-pointed arrow (see Figure 18.5).

Table 18.12 presents the chi-square goodness-of-fit value, and the unstandardized and standardized regression weights for the modified model.

The chi-square goodness-of-fit value for this modified model (188.008) is smaller than the chi-square value obtained for the original model (253.472). With a smaller chi-square value, the modified model, therefore, represents a better fit to the data than the original model. However, the question remains as to whether the improvement in fit represents a statistically significant improvement.

18.13.4 Comparing the Original (Default) Model against the Modified Model

In this example, a direct comparison in goodness-of-fit between the original and modified models is possible because both models are based on the same data set, and have different degrees of freedom. A test of the original model against the modified model can be obtained by subtracting the smaller chi-square value from the larger one. In this example, the comparison chi-square

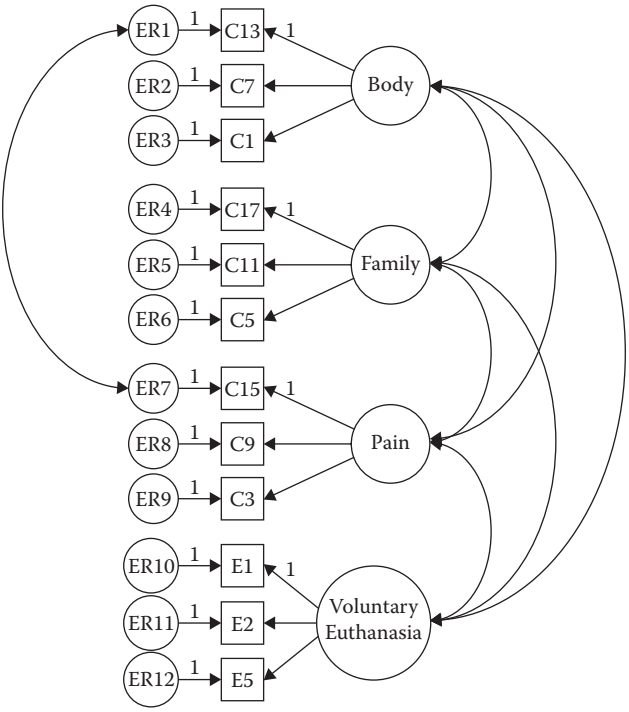


FIGURE 18.5
Modified measurement model.

value is 65.464 (i.e., 253.472 – 188.008). If the original model is correctly specified, this value will have an approximate chi-square distribution with degrees of freedom equal to the difference between the degrees of freedom of the competing models. In this example, the difference in degrees of freedom is 1 (i.e., 48 – 47). Therefore, with 1 degree of freedom, a chi-square value of 65.464 is significant at the 0.05 level. On the basis of this test, it can be concluded that the modified model (with correlated errors) represents a *significantly* better fit to the data than the original model. Please note that the regression weights, obtained under the modified model, remain statistically significant by the critical ratio test, and are highly similar to those obtained under the original model.

18.13.5 Multi-Model Analysis: Evaluation of the Direct Path Model versus the Indirect Path Model

Once the modified measurement model has been confirmed, the fit of the structural path model (with correlated error terms) (Figure 18.6) can be evaluated. The factor structure confirmed in the measurement model will be used as the basis for the path model. That is, the four unobserved factors

TABLE 18.12

Chi-Square Goodness-of-Fit Value, Unstandardized and Standardized Regression Weights for the Modified Measurement Model

Regression Weights: (Group Number 1—Default Model)

			Estimate	S.E.	C.R.	*P*	Label
C1	←	BODY	1.036	.043	24.302	***	
C7	←	BODY	1.030	.037	27.659	***	
C13	←	BODY	1.000				
C5	←	FAMILY	.973	.061	16.014	***	
C11	←	FAMILY	1.032	.049	21.036	***	
C17	←	FAMILY	1.000				
C3	←	PAIN	.986	.044	22.645	***	
C9	←	PAIN	1.022	.041	24.897	***	
C15	←	PAIN	1.000				
E5	←	VOLUNTARY_EUTHANASIA	.781	.051	15.462	***	
E2	←	VOLUNTARY_EUTHANASIA	.587	.044	13.445	***	
E1	←	VOLUNTARY_EUTHANASIA	1.000				

Result (Default model)
Minimum was achieved
Chi-square = 188.008
Degrees of freedom = 47
Probability level = .000
Scalar estimates (Group number 1—Default model)
Maximum likelihood estimates

Standardized Regression Weights: (Group Number 1—Default Model)

			Estimate
C1	←	BODY	.867
C7	←	BODY	.921
C13	←	BODY	.907
C5	←	FAMILY	.731
C11	←	FAMILY	.901
C17	←	FAMILY	.876
C3	←	PAIN	.859
C9	←	PAIN	.903
C15	←	PAIN	.889
E5	←	VOLUNTARY_EUTHANASIA	.812
E2	←	VOLUNTARY_EUTHANASIA	.705
E1	←	VOLUNTARY_EUTHANASIA	.824

BODY, FAMILY, PAIN, and VOLUNTARY EUTHANASIA, together with their respective measurement indicators and the correlated error terms will be incorporated into the structure of the path model to be evaluated.

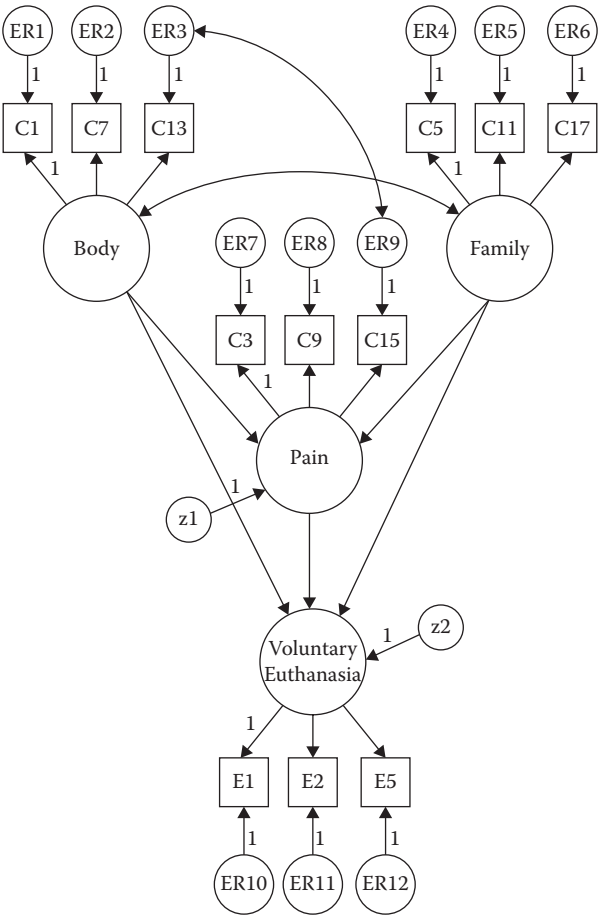


FIGURE 18.6
Path model (with correlated errors) for the prediction of voluntary euthanasia.

The posited model presented in Figure 18.6 contains two models—(1) the full **direct model** that incorporates all identified paths linking the four factors, and (2) the **indirect model**, in which the two direct paths linking BODY to VOLUNTARY EUTHANASIA (**Body → Voluntary Euthanasia**) and FAMILY to VOLUNTARY EUTHANASIA (**Family → Voluntary Euthanasia**) will not be estimated. As both these models are nested (i.e., they are hierarchical models based on the same data set) and have different degrees of freedom, their goodness-of-fit can be directly compared via multi-model analysis.

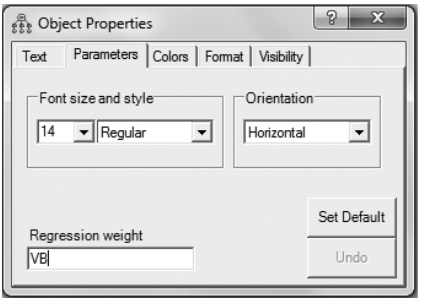
In conducting a multi-model analysis using AMOS Graphics, the procedure involves (1) defining the full direct model as presented in Figure 18.6, and (2) defining the indirect model in which the two direct paths linking BODY and FAMILY to VOLUNTARY EUTHANASIA (**Body → Voluntary**

Euthanasia; **Family → Voluntary Euthanasia**) are constrained to zero. Constraining paths to zero is equivalent to those paths not being estimated.

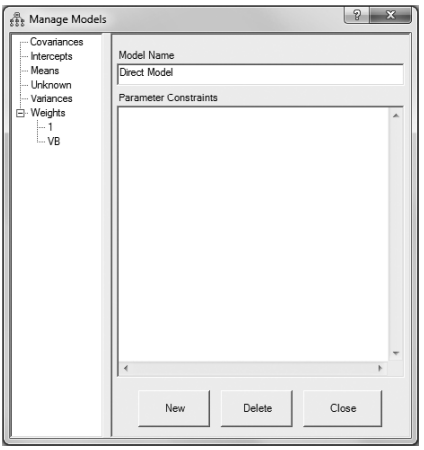
18.13.5.1 Defining the Direct and Indirect Models

Follow these steps to define the **direct** and **indirect** models:

1. Retrieve Figure 18.6. Double-click on the **Body → Voluntary Euthanasia** path. The following **Object Properties** window will open. Label this path **VB** by clicking the **Parameters** tab, and in the **Regression weight** field, type the label **VB**.



2. Double-click on the **Family → Voluntary Euthanasia** path. When the **Object Properties** window opens, click the **Parameters** tab, and in the **Regression weight** field, type the label **VF** (see Figure 18.7).
3. In order to define the first full **direct model**, click **Analyze** and then **Manage Models** from the drop-down menu bar. When the **Manage Models** window opens, type the label **Direct Model** in the **Model Name** field. As this model imposes no constraints on any of the model's parameters, leave the **Parameter Constraints** field blank.



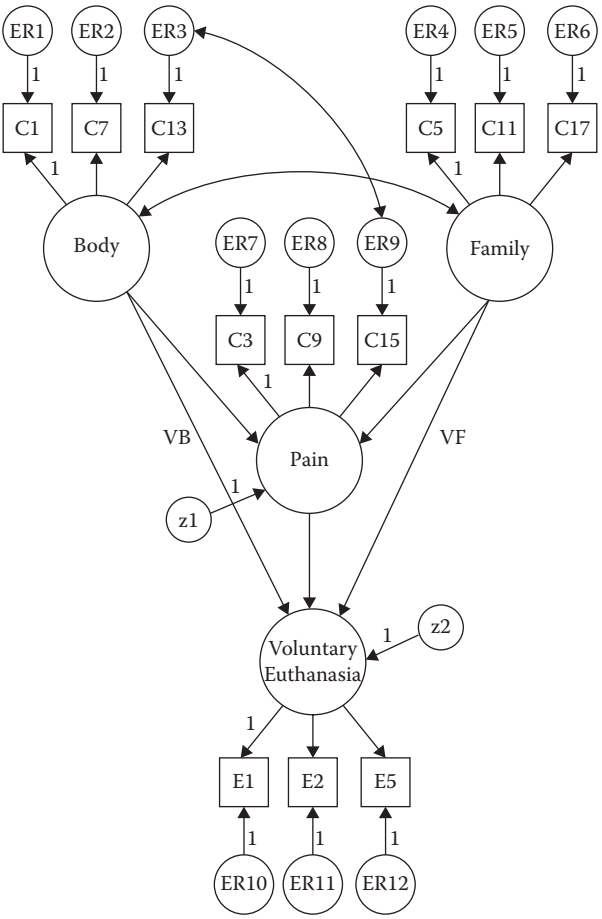
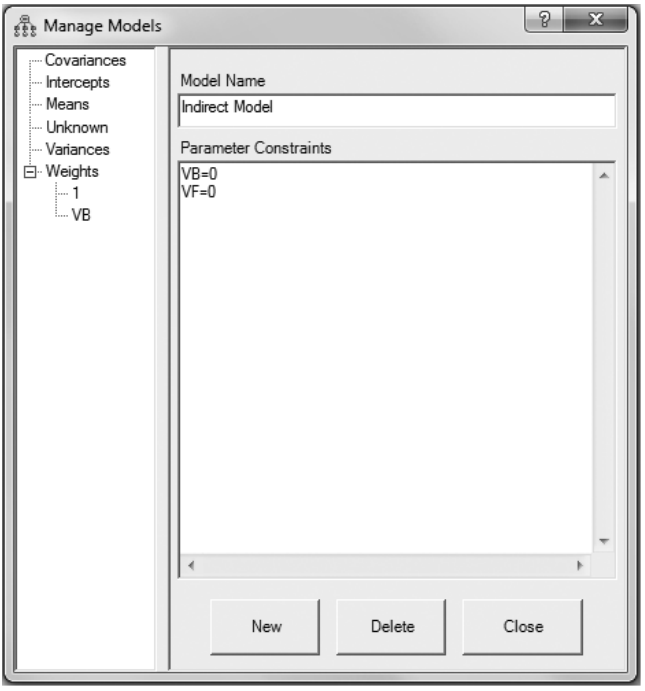


FIGURE 18.7
Direct path model versus indirect path model.

4. In order to define the **indirect model**, click New in the **Manage Models** window. This will open a new **Manage Models** window. Type the label **Indirect Model** in the **Model Name** field. As this indirect model requires constraining the two paths **VB** (**Body → Voluntary Euthanasia**) and **VF** (**Family → Voluntary Euthanasia**) to 0 (i.e., they will not be estimated), impose these constraints by typing **VB = 0** and **VF = 0** in the **Parameter Constraints** field. Click Close when finished.
5. The procedure involved in (i) linking the two nested models to the data set, (ii) obtaining the desired statistics in the results output, and (iii) conducting the multi-model analysis is identical to that described



in Section 18.11.2. The multi-model analysis allows for the testing of the two nested models simultaneously, and presents a comparison of the goodness-of-fit of the two competing direct and indirect models.

18.13.5.2 Results and Interpretion

18.13.5.2.1 Summary of Models

Table 18.13 presents the direct and indirect models' chi-square goodness-of-fit statistics, their baseline comparisons fit indices, and the model comparison statistics.

The direct model has 47 degrees of freedom, two less than the indirect model. This is because the direct model estimated two additional direct paths linking BODY and FAMILY to VOLUNTARY EUTHANASIA. The estimation of these two paths has "used up" two additional degrees of freedom. While the chi-square values for both models are significant (indirect model: $\chi^2(N = 357, df = 49) = 195.766$, $p < .05$; direct model: $\chi^2(N = 357, df = 47) = 188.008$, $p < .05$), the baseline comparisons fit indices NFI, RFI, IFI, TLI, and CFI for both models are above 0.90 (range: 0.921–0.957). These values indicate that both the hypothesized direct and indirect models fit the observed variance-covariance matrix well relative to the null or independence model. Indeed, the only possible improvement in fit for these two models ranges from 0.043 to 0.079.

TABLE 18.13

Direct and Indirect Models' Chi-Square Goodness-of-Fit Indices, Baseline Comparisons Indices, and Model Comparison Statistics

CMIN

Model	NPAR	CMIN	d*f*	P	CMIN/d*f*
Direct model	31	188.008	47	.000	4.000
Indirect model	29	195.766	49	.000	3.995
Saturated model	78	.000	0		
Independence model	12	3332.993	66	.000	50.500

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Direct model	.944	.921	.957	.939	.957
Indirect model	.941	.921	.955	.939	.955
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

Nested Model Comparisons

Assuming Direct Model to Be Correct

Model	d*f*	CMIN	*P*	NFI Delta1	IFI Delta2	RFI rho1	TLI rho2
Indirect model	2	7.758	.021	.002	.002	.000	.000

18.13.5.2.2 Goodness-of-Fit Comparison: Direct Model versus Indirect Model

While both models fit the data well (relative to the null model), the question arises as to which model fits better. Since both models are nested (i.e., hierarchical) and have different degrees of freedom, their goodness-of-fit can be directly compared. Looking at the **Nested Model Comparison** statistics in Table 18.13, it can be seen that subtracting the direct model's chi-square value from the indirect model's chi-square value (195.766 – 188.008) yields a chi-square difference value of 7.758. With two degrees of freedom (49 – 47), this statistic is significant at the 0.05 level. Thus, although both models fit the data relatively well, the direct model represents a significantly better fit than the indirect model, and is to be preferred. This conclusion is further supported by the **Akaike Criterion Information** (**AIC**) comparison statistics. The direct model yielded a lower AIC value (250.008) than the indirect model (253.766), which indicates that the direct model is both better fitting and more parsimonious than the indirect model.

18.13.5.2.3 Regression Weights, Standardized Regression Weights, and Squared Multiple Correlations

Table 18.14 shows the regression weights, standardized regression weights, and squared multiple correlations for the hypothesized direct path model.

TABLE 18.14

Regression Weights, Standardized Regression Weights, and Squared Multiple Correlations for Direct Path Model

Regression Weights: (Group Number 1—Direct Model)

			Estimate	S.E.	C.R.	*P*	Label
PAIN	←	BODY	.380	.056	6.835	***	
PAIN	←	FAMILY	.388	.071	5.467	***	
VOLUNTARY_EUTHANASIA	←	PAIN	.586	.073	8.066	***	
VOLUNTARY_EUTHANASIA	←	BODY	.135	.062	2.157	.031	VB
VOLUNTARY_EUTHANASIA	←	FAMILY	.059	.078	.759	.448	VF
C1	←	BODY	1.000				
C7	←	BODY	.994	.040	24.705	***	
C13	←	BODY	.966	.040	24.302	***	
C5	←	FAMILY	1.000				
C11	←	FAMILY	1.061	.065	16.282	***	
C17	←	FAMILY	1.028	.064	16.014	***	
C3	←	PAIN	1.000				
C9	←	PAIN	1.037	.045	22.935	***	
C15	←	PAIN	1.015	.045	22.645	***	
E5	←	VOLUNTARY_EUTHANASIA	.781	.051	15.462	***	
E2	←	VOLUNTARY_EUTHANASIA	.587	.044	13.445	***	
E1	←	VOLUNTARY_EUTHANASIA	1.000				

Scalar estimates (Group number 1—Direct Model)
Maximum likelihood estimates

Standardized Regression Weights: (Group Number 1—Direct Model)

			Estimate
PAIN	←	BODY	.421
PAIN	←	FAMILY	.347
VOLUNTARY_EUTHANASIA	←	PAIN	.584
VOLUNTARY_EUTHANASIA	←	BODY	.149
VOLUNTARY_EUTHANASIA	←	FAMILY	.053
C1	←	BODY	.867

TABLE 18.14 (*Continued*)

Regression Weights, Standardized Regression Weights, and Squared Multiple Correlations for Direct Path Model

Standardized Regression Weights: (Group Number 1—Direct Model)

			Estimate
C7	←	BODY	.921
C13	←	BODY	.907
C5	←	FAMILY	.731
C11	←	FAMILY	.901
C17	←	FAMILY	.876
C3	←	PAIN	.859
C9	←	PAIN	.903
C15	←	PAIN	.889
E5	←	VOLUNTARY_EUTHANASIA	.812
E2	←	VOLUNTARY_EUTHANASIA	.705
E1	←	VOLUNTARY_EUTHANASIA	.824

Covariances: (Group Number 1—Direct Model)

			Estimate	S.E.	C.R.	*P*	Label
BODY	↔	FAMILY	.570	.068	8.341	***	
ER3	↔	ER9	.137	.019	7.105	***	

Correlations: (Group Number 1—Direct Model)

			Estimate
BODY	↔	FAMILY	.636
ER3	↔	ER9	.583

Squared Multiple Correlations: (Group Number 1—Direct Model)

	Estimate
PAIN	.483
VOLUNTARY_EUTHANASIA	.525
E1	.679
E2	.497
E5	.660
C15	.791
C9	.816
C3	.737
C17	.767
C11	.812
C5	.535
C13	.823
C7	.849
C1	.751

Of the five coefficients associated with the paths linking the model's exogenous and endogenous variables, four are significant by the critical ratio test (>±1.96, $p < .05$). The non-significant coefficient is associated with the path linking FAMILY to VOLUNTARY EUTHANASIA. The standardized path coefficients have been incorporated into the final direct model presented in Figure 18.8.

The results can be interpreted as follows. The perception of the debilitated nature of a person's body (BODY) and the negative impact that a person's terminal illness has on his/her family (FAMILY) is related indirectly to the support for voluntary euthanasia, being mediated by the assessment of the PAIN experienced. Therefore, the more debilitated a person's body is perceived to be, and the greater the perceived negative impact on the person's family, the greater the assessment of the physical pain experienced ($\beta = 0.42$ and $\beta = 0.35$, respectively). The greater the assessment of the physical pain experienced, the greater is the support for voluntary euthanasia ($\beta = 0.58$). The perception of the debilitated nature of the body is also related directly to the support for voluntary euthanasia. Thus, the more debilitated a person's body is perceived to be, the greater is the support for voluntary euthanasia ($\beta = 0.15$). The FAMILY factor was not significantly related to the support for voluntary euthanasia ($p > .05$).

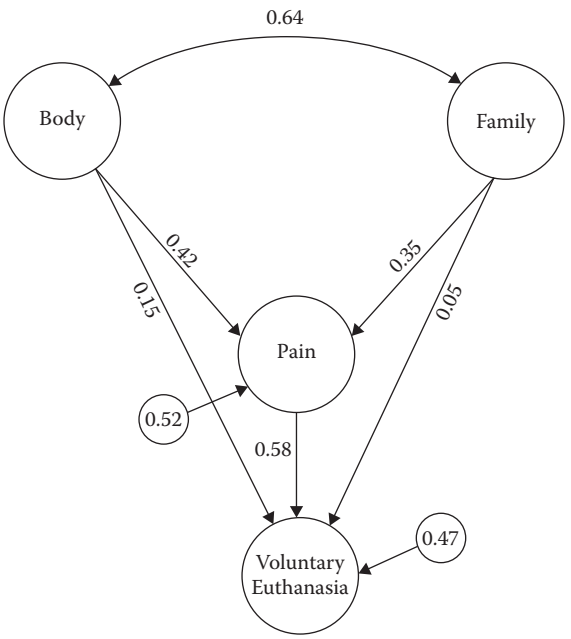


FIGURE 18.8
Direct model predicting support for voluntary euthanasia.

18.13.5.2.4 Explained Variances and Residual Variances

The unidirectional arrows pointing to the latent factors PAIN and VOLUNTARY EUTHANASIA represent unexplained (residual) variances for these two factors. The residual variances are calculated by subtracting the factors' squared multiple correlations (explained variances) (see Table 18.14) from 1. Therefore, for this hypothesized model, 52% of the variation in PAIN is unexplained; alternatively, 48% of the variance is accounted for by the joint influence of the BODY and FAMILY predictors. Similarly, 47% of the variation in the support for voluntary euthanasia is unexplained; alternatively, 53% of the variance is accounted for by the joint influences of the predictors BODY, FAMILY, and PAIN.

18.14 Example 4: Multi-Group Analysis

This example demonstrates a multi-group analysis on the path model tested in Example 3. The path analysis carried out in Example 3 investigated the sequential relationships between three predictor variables—BODY, FAMILY, PAIN—and the dependent variable of support for VOLUNTARY EUTHANASIA. The present example reconsiders this path model (without correlated errors) and attempts to apply it simultaneously to a sample of 136 males and a sample of 221 females. The question to be examined is whether the pattern of structural relationships hypothesized in the path model follows the same dynamics for males and females. This example is based on the same variance-covariance matrix (generated from the SPSS file EUTHAN.SAV) employed in Examples 1, 2, and 3.

18.14.1 Multi-Group Confirmatory Factor Analysis

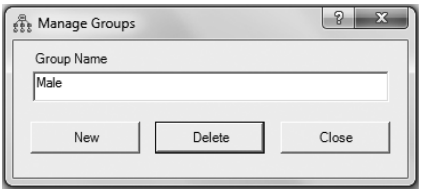
Prior to executing the multi-group analysis for the path model presented in Figure 18.3, it is necessary to perform a multi-group analysis for the measurement model presented in Figure 18.4. Recall that the purpose of the measurement model is to verify that the 12 measurement variables written to reflect the four latent constructs (BODY, FAMILY, PAIN, VOLUNTARY EUTHANASIA) do so in a reliable manner. Thus, in investigating sex differences in the path model, it is necessary to first test whether the factor structure represented by the posited measurement model is the same for both males and females. *If the analysis shows no significant differences in regression weights (i.e., factor loadings) between males and females, then the same regression weights can be used for both groups*. This, in turn, will allow the regression weights themselves to be estimated more efficiently, as well as simplifying the estimation of model fit. However, if the analysis shows significant differences in the regression weights between males and females, then these differences must be incorporated into the structural path model to be estimated.

18.14.1.1 Conducting Multi-Group Modeling for Males and Females: The Measurement Model

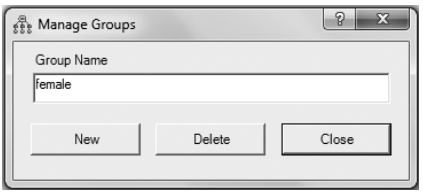
Figure 18.4 presents the measurement model for the combined samples of males and females ($N = 357$). To test for sex (group) differences in the regression weights (factor loadings) for this measurement model, it will be necessary to (1) set up separate but identical measurement models for the male and female samples, (2) link the male and female models to their respective data sets, (3) set up an invariant model (in which males and females are hypothesized to share the same regression weights) and a variant model (in which males and females are hypothesized to have different regression weights) that can be directly compared as to their model fit, and (4) employ the **critical ratio** test to test for sex differences in the regression weights.

1. **Setting up separate but identical measurement models for the male and female samples.**

   To do this, retrieve the diagram shown in Figure 18.4. From the drop-down menu bar, click **Analyze** and then **Manage Groups** to open the **Manage Groups** window below. To create the male measurement model, change the default name (**Group number 1**) in the **Group Name** field to **Male**.

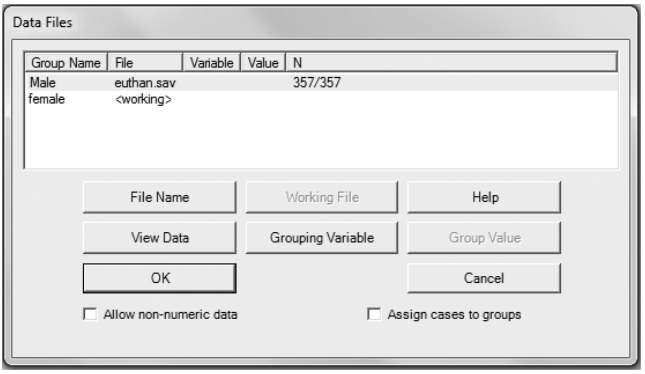


   To create the female measurement model, click **New** to open a new **Manage Groups** window. Change the name in the **Group Name** field to **Female**. Click the **Close** button. This will create two identical measurement models for the male and female samples.

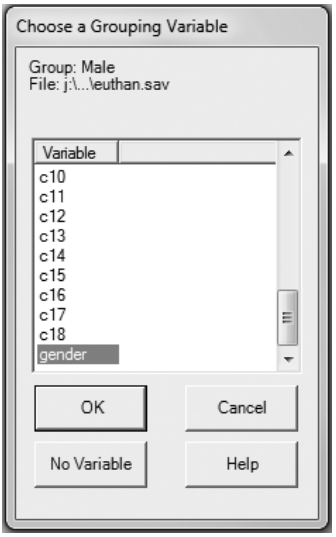


2. **Linking the male and female measurement models of their respective data sets**.
   - Click from the icon toolbox (**File → Data Files…**) to open the following **Data Files** window. Click File Name to search for the

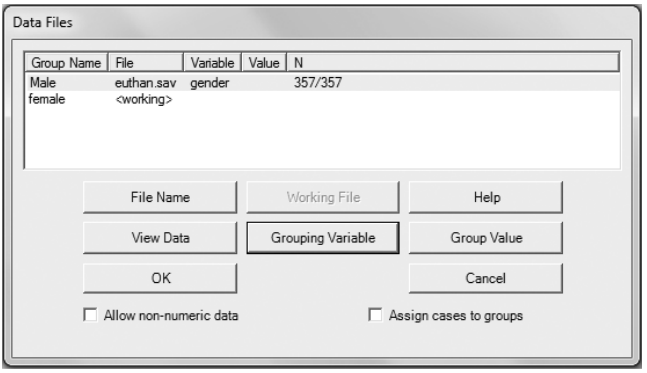
data file (**EUTHAN.SAV**) in the computer's directories. Once the data file has been located, open it. By opening the data file, AMOS will automatically link the data file to the measurement model, as indicated in the window below.



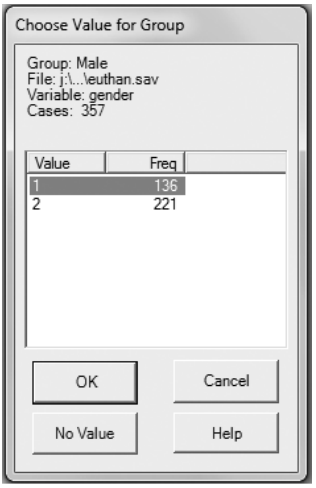
- To link the male model to the male sub-sample within the EUTHAN.SAV data set, click Grouping Variable to open the **Choose a Grouping Variable** window below.



- Select **GENDER** as the grouping variable by clicking it (highlight), and then click OK to return to the **Data Files** window. Note that under the **Variable** heading, GENDER has been selected as the grouping variable.

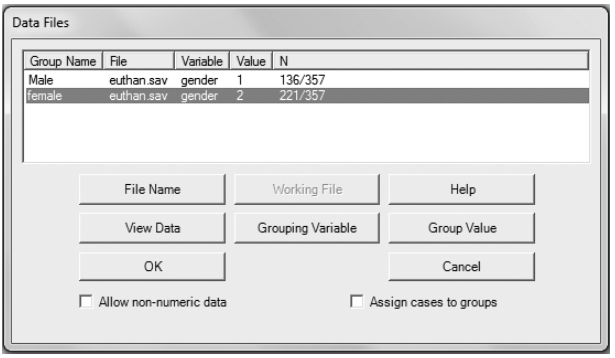


- To link the male model to the male sub-sample (coded **1** within the **GENDER** variable), click Group Value to open the **Choose Value for Group** window below. Select **1** (males) under the **Value** heading by clicking it (highlight).



- Click OK to return to the **Data Files** window. Note that under the **Value** heading, the number 1 (males) has been selected. Notice also that under the **N** (sample size) heading, the number **136/357** has been listed. This indicates that a sub-sample of 136 males (out of a total sample of 357 subjects) has been linked to the male measurement model.
- To link the female model to the female sub-sample within the EUTHAN.SAV data set, repeat the above procedure, but use the Group Value of 2 to select the female sub-sample. Successful operation of this procedure will yield the following final Data Files window below. As can be seen, a sub-sample of 221 females (out

of a total sample of 357 subjects) has been linked to the female measurement model. Click OK to exit this window.

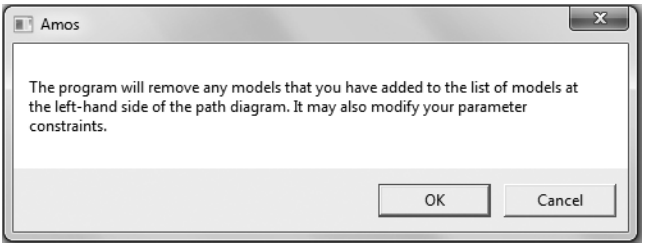


3. **Setting up group invariant and group variant measurement models.**

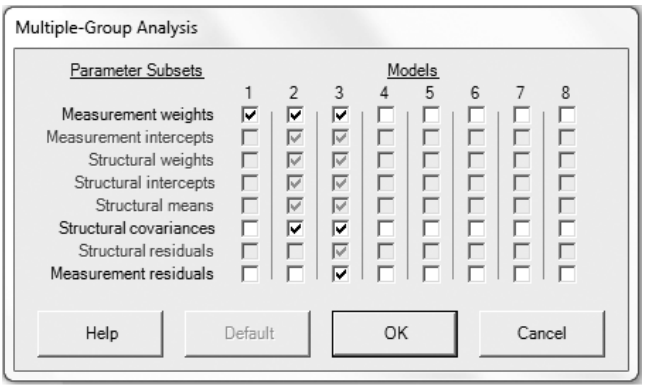
   The hypothesis to be tested is that the measurement model in Figure 18.4 holds for males as well as females. This hypothesis requires that the factor pattern (i.e., the regression weights) be the same for both groups (group invariant model). That is, it requires that every regression weight for the male sample be *equal* to the corresponding regression weight for the female sample. However, it does not require that the unique variances for males and females be grouped invariant. The common factor variances and covariances may also differ in the two groups. The rationale underlying the hypothesis of group invariant regression weights is that, while it is probably reasonable to assume that the observed and unobserved variables have different variances and covariances among males and females, the two groups may share the same regression weights.

   To set up the **group invariant** model (in which males and females are hypothesized to share the same regression weights), it is necessary to constrain 16 paths (8 regression weights for males, 8 regression weights for females) to equality.

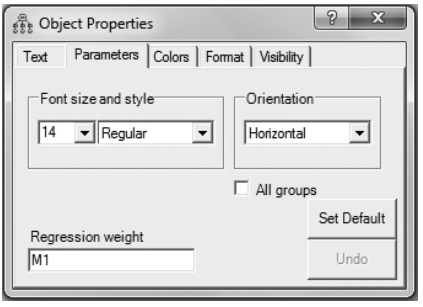
   - Select the male measurement model by clicking the *male* label Male Female in the AMOS graphic page. Click the multi-group icon: . AMOS will provide the following information.



- Click OK to open the **Multiple-Group Analysis** window below.



- Click OK to bring up the measurement model as presented in Figure 18.9.
- As can be seen from Figure 18.9, AMOS has provided default names for all parameters (e.g., paths, variances, covariances, etc.) in the model.
- In setting up the **male** measurement model, delete all parameter names. Rename those parameters associated with the measurement variables C1, C7, C5, C11, C3, C9, E2, and E5 to **M1**, **M2**, **M3**, **M4**, **M5**, **M6**, **M7**, and **M8**. For example, to label the regression weight between the latent construct **BODY** and the measurement variable **C7** (BODY → C7), double-click on this parameter. This will open the **Object Properties** window below. Under the **Parameters** tab, enter the label **M1** in the **Regression weight** field. Continue this process until all eight regression weights (M1–M8) have been labeled (see Figure 18.10).



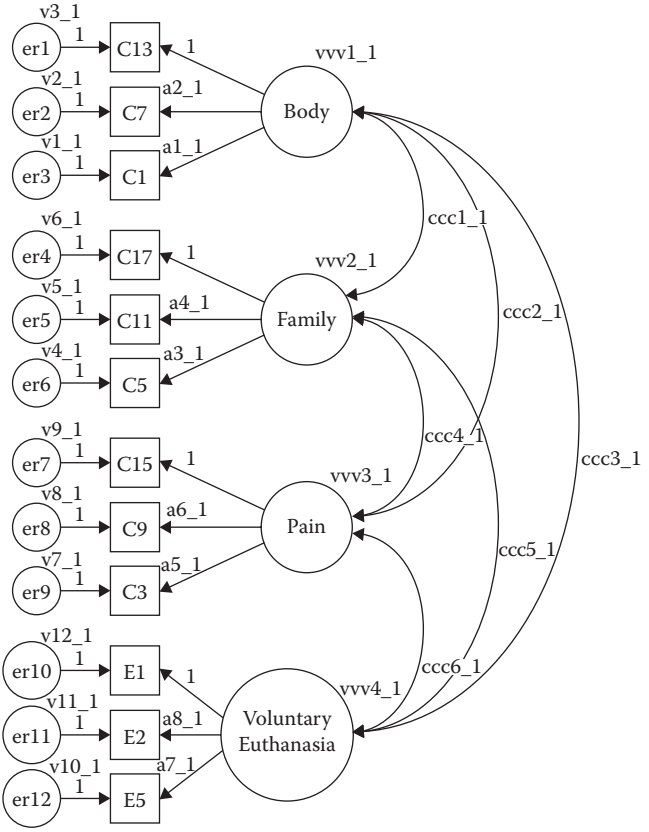


FIGURE 18.9
Measurement model with default names for parameters.

- Select the female measurement model by clicking the *female* label Male Female in the AMOS graphic page. Repeat the above procedure for labeling the eight regression weights for this female group. For this female measurement model, label the eight regression weights **F1–F8** (see Figure 18.11). That is, the labels for the regression weights for the male and female models must be different, as shown in Figures 18.10 and 18.11.
- In order to set up the **group invariant** model, click **Analyze** and then **Manage Model** (from the drop-down menu bar) to open the **Manage Models** window. Set up the group invariant model by typing in the name **group invariant** in the **Model Name** field, and then entering its constraints in the **Parameter Constraints** field (i.e., M1 = F1, M2 = F2, M3 = F3, M4 = F4, M5 = F5, M6 = F6, M7 = F7, M8 = F8). *This procedure constrains the 16 regression weights for males and females to be equivalent*.

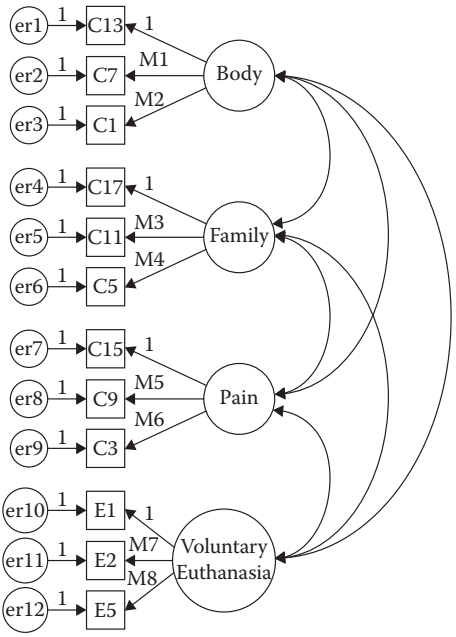


FIGURE 18.10
Male measurement model.

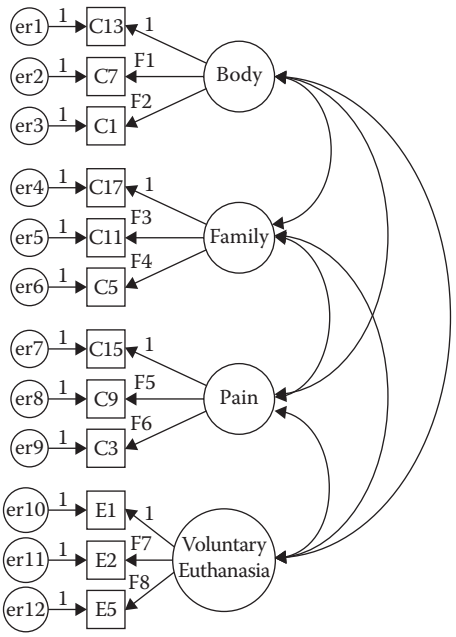
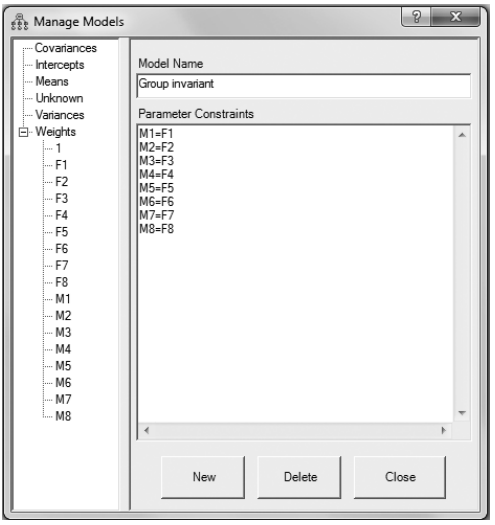
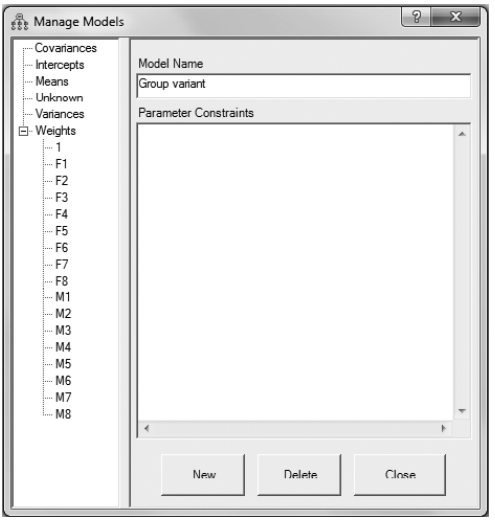


FIGURE 18.11
Female measurement model.



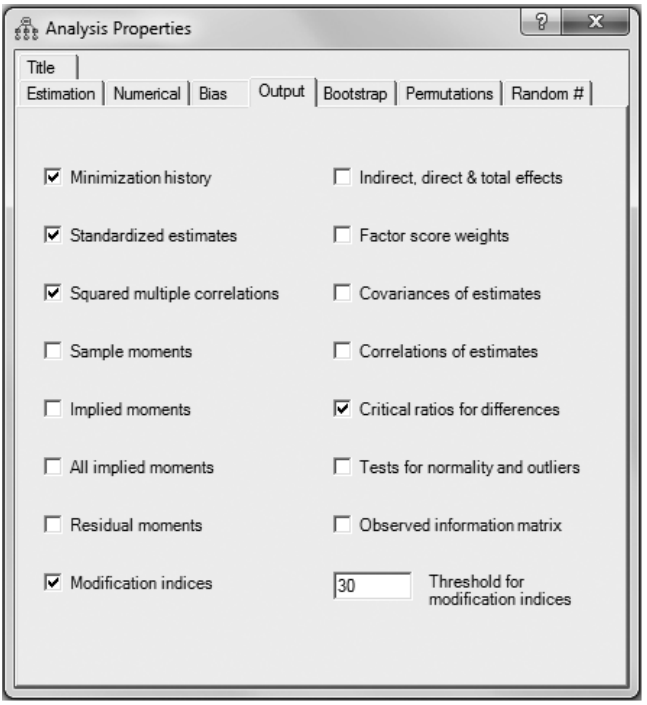
- To set up the **group variant** model (in which males and females are hypothesized to have different regression weights), click New in the **Manage Models** window above. This will open a new **Manage Models** window. In the **Model Name** field, type the name **group variant**. As this model allows the regression weights for males and females to be estimated separately (i.e., there are no constraints), the **Parameter Constraints** field is left empty. Click Close to complete this procedure.



4. **Testing for sex differences among the regression weights.**

   In order to test for sex differences among the regression weights, the **critical ratio** (C.R.) test can be employed to obtain the critical ratio statistics for the differences among male and female subjects' regression weights. The critical ratio of a pair of estimates provides a test of the hypothesis that the two parameters are equal (Arbuckle and Wothke, 1999).

   Click the icon (**View → Analysis Properties…**) to open the **Analysis Properties** dialog box. Under the **Output** tab, ensure that the **Critical ratios for differences** field is checked. This option will compare all model parameters posited for the male and female models.



After closing the **Analysis Properties** dialog box, click (**Analyze → Calculate Estimates**) to perform the multi-group analysis.

18.14.1.2 Results and Interpretation

18.14.1.2.1 Notes for Models

For this multi-group analysis, there are two data sets (for males and females), each containing 12 measurement variables. The two covariance matrices generated from the two data sets contain 156 sample moments.

TABLE 18.15

Computation of Degrees of Freedom and Chi-Square Goodness-of-Fit Statistics for Group Invariant and Group Variant Models

Notes for Model (Group Invariant)	
Computation of Degrees of Freedom (Group Invariant)	
Number of distinct sample moments:	156
Number of distinct parameters to be estimated:	52
Degrees of freedom (156 – 52):	104

Result (group invariant)
Minimum was achieved
Chi-square = 327.336
Degrees of freedom = 104
Probability level = .000

Notes for Model (Group Variant)	
Computation of Degrees of Freedom (Group Variant)	
Number of distinct sample moments:	156
Number of distinct parameters to be estimated:	60
Degrees of freedom (156 – 60):	96

Result (group variant)
Minimum was achieved
Chi-square = 312.254
Degrees of freedom = 96
Probability level = .000

For the **group invariant** model, there are 52 parameters to be estimated. This model, therefore, has 104 (156 – 52) degrees of freedom, and yielded a significant chi-square value, $\chi^2(N = 357, df = 104) = 327.336$, $p < .05$ (see Table 18.15).

For the **group variant** model, there are 60 parameters to be estimated. This model, therefore, has 96 (156 – 60) degrees of freedom, and also yielded a significant chi-square value, $\chi^2(N = 357, df = 96) = 312.254$, $p < .05$ (see Table 18.15).

18.14.1.2.2 Summary of Models

Table 18.16 presents the chi-square goodness-of-fit statistics, baseline comparisons fit indices, and model comparison statistics for the group-invariant and group-variant measurement models.

Although the chi-square values for both models are statistically significant (i.e., both models yielded poor fit by the chi-square goodness-of-fit test), the baseline comparison fit indices NFI, RFI, IFI, TLI, and CFI for both models are close to or are above 0.9 (range: 0.875–0.935). These values indicate the improvement in fit of both models relative to the null model. Indeed, the only possible improvement in fit for these two models ranges from 0.065 to 0.125.

TABLE 18.16

Chi-Square Goodness-of-Fit Statistics, Baseline Comparisons Fit Indices, and Model Comparison Statistics

Model Fit Summary

CMIN

Model	NPAR	CMIN	d*f*	*P*	CMIN/d*f*
Group invariant	52	327.336	104	.000	3.147
Group variant	60	312.254	96	.000	3.253
Saturated model	156	.000	0		
Independence model	24	3442.653	132	.000	26.081

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Group invariant	.905	.879	.933	.914	.933
Group variant	.909	.875	.935	.910	.935
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Group invariant	.078	.068	.087	.000
Group variant	.080	.070	.090	.000
Independence model	.266	.258	.274	.000

AIC

Model	AIC	BCC	BIC	CAIC
Group invariant	431.336	440.184		
Group variant	432.254	442.464		
Saturated model	312.000	338.545		
Independence model	3490.653	3494.737		

Nested Model Comparisons

Assuming Model Group Variant to be Correct

Model	d*f*	CMIN	*P*	NFI Delta1	IFI Delta2	RFI rho1	TLI rho2
Group invariant	8	15.082	.058	.004	.005	–.004	–.004

Another very useful fit index is the *root mean square error of approximation* (**RMSEA**), which takes into account the error of approximation in the population. It is a measure of *discrepancy per degree of freedom* and is representative of the goodness-of-fit when the proposed model is estimated in the population. Values ranging from 0.05 to 0.08 are deemed acceptable (Browne and Cudeck, 1993; MacCallum et al., 1996). The RMSEA values for the **group-invariant** and **group-variant** measurement models are 0.078 and 0.080, respectively. These values suggest that the fit of these two models is adequate.

The fit of the two competing models can be directly compared. From the **Nested Model Comparisons** statistics, it can be seen that the chi-square difference value for the two models is 15.082 (327.336 – 312.254). With eight degrees of freedom (104 – 96), this value is not significant at the 0.05 level ($p > .05$). Thus, the two models do not differ significantly in their goodness-of-fit.

The fit of the two models can also be compared using the **AIC** (Akaike Information Criterion) measure (Akaike, 1973, 1987). In evaluating the hypothesized model, this measure takes into account both model parsimony and model fit. Simple models that fit well receive low scores, whereas poorly fitting models get high scores. The AIC measure for the group invariant model (431.336) is slightly smaller than that for the group variant model (432.254), indicating that the group invariant model is both more parsimonious and better fitting than the group variant model. On the basis of the model comparisons findings, and assuming that the group invariant model is correct, the group invariant model's estimates are preferable over the group variant model's estimates.

18.14.1.2.3 Unstandardized Regression Weights and Standardized Regression Weights

Table 18.17 presents the unstandardized regression weights and the standardized regression weights for males and females (**group invariant model**).

The unstandardized regression weights for males and females are all significant by the critical ratio test ($>\pm 1.96$, $p < .05$). For male subjects, the standardized regression weights range from 0.669 to 0.931, and for female subjects they range from 0.732 to 0.917. These values indicate that, for both males and females, the 12 measurement variables are significantly represented by their respective unobserved constructs.

18.14.1.2.4 Explained Variances and Residual Variances

For males and females, the explained variances for the 12 measurement variables are represented by their squared multiple correlations (see Table 18.18). For the male subjects, the percentage of variance explained ranged from 44.8% (E2) to 86.6% (C13); for the female subjects, the percentage of variance explained ranged from 53.6% (E2) to 84.0% (C7). The residual (unexplained) variances are calculated by subtracting each explained variance from 1. Thus, for the 12 measurement variables, the residual variances ranged from 13.4% to 55.2% for male subjects and from 16% to 46.4% for female subjects.

TABLE 18.17

Unstandardized Regression Weights and Standardized Regression Weights for Males and Females

Regression Weights: (Male—Group Invariant)

			Estimate	S.E.	C.R.	*P*	Label
C1	←	BODY	1.039	.044	23.683	***	m2
C7	←	BODY	1.035	.038	27.239	***	m1
C13	←	BODY	1.000				
C5	←	FAMILY	.964	.060	15.987	***	m4
C11	←	FAMILY	1.031	.049	21.035	***	m3
C17	←	FAMILY	1.000				
C3	←	PAIN	1.015	.044	23.133	***	m6
C9	←	PAIN	1.041	.043	24.216	***	m5
C15	←	PAIN	1.000				
E5	←	Voluntary_euthanasia	.799	.051	15.655	***	m8
E2	←	Voluntary_euthanasia	.596	.044	13.604	***	m7
E1	←	Voluntary_euthanasia	1.000				

Scalar estimates (Male—Group invariant)
Maximum likelihood estimates

Standardized Regression Weights: (Male—Group Invariant)

			Estimate
C1	←	BODY	.855
C7	←	BODY	.928
C13	←	BODY	.931
C5	←	FAMILY	.699
C11	←	FAMILY	.899
C17	←	FAMILY	.878
C3	←	PAIN	.805
C9	←	PAIN	.877
C15	←	PAIN	.877
E5	←	Voluntary_euthanasia	.801
E2	←	Voluntary_euthanasia	.669
E1	←	Voluntary_euthanasia	.791

Scalar estimates (Female—Group invariant)
Maximum likelihood estimates

Regression Weights: (Female—Group Invariant)

			Estimate	S.E.	C.R.	*P*	Label
C1	←	BODY	1.039	.044	23.683	***	m2
C7	←	BODY	1.035	.038	27.239	***	m1
C13	←	BODY	1.000				

TABLE 18.17 (*Continued*)

Unstandardized Regression Weights and Standardized Regression Weights for Males and Females

Regression Weights: (Female—Group Invariant)

			Estimate	S.E.	C.R.	*P*	Label
C5	←	FAMILY	.964	.060	15.987	***	m4
C11	←	FAMILY	1.031	.049	21.035	***	m3
C17	←	FAMILY	1.000				
C3	←	PAIN	1.015	.044	23.133	***	m6
C9	←	PAIN	1.041	.043	24.216	***	m5
C15	←	PAIN	1.000				
E5	←	Voluntary_euthanasia	.799	.051	15.655	***	m8
E2	←	Voluntary_euthanasia	.596	.044	13.604	***	m7
E1	←	voluntary_euthanasia	1.000				

Standardized Regression Weights: (Female—Group Invariant)

			Estimate
C1	←	BODY	.868
C7	←	BODY	.917
C13	←	BODY	.894
C5	←	FAMILY	.748
C11	←	FAMILY	.901
C17	←	FAMILY	.876
C3	←	PAIN	.906
C9	←	PAIN	.913
C15	←	PAIN	.890
E5	←	Voluntary_euthanasia	.830
E2	←	Voluntary_euthanasia	.732
E1	←	Voluntary_euthanasia	.832

18.14.1.2.5 The Critical Ratio Test for Sex Differences between the Regression Weights

Please note that the pairwise comparison C.R. The test is carried out on the regression weights obtained from the variant group model. This is because the regression weights from the invariant group model are set to equality and therefore cannot be compared.

From Table 18.19, it can be seen that two of the pairwise comparisons (males versus females) for regression weights (M5-F5, M8-F8) are significant (C.R. > ±1.96, $p < .05$). Hence, these two sex differences in regression weights (associated with the measurement variables of C9 and E5) will be incorporated into the multi-group analysis of the structural path model (see Figure 18.3).

TABLE 18.18

Explained Variances (Squared Multiple Correlations) for the 12 Measurement Variables

Squared Multiple Correlations: (Male—Group Invariant)	
	Estimate
E1	.626
E2	.448
E5	.641
C15	.769
C9	.769
C3	.648
C17	.772
C11	.808
C5	.489
C13	.866
C7	.861
C1	.731

Squared Multiple Correlations: (Female—Group Invariant)	
	Estimate
E1	.693
E2	.536
E5	.689
C15	.792
C9	.834
C3	.821
C17	.767
C11	.811
C5	.559
C13	.800
C7	.840
C1	.754

18.14.2 Multi-Group Path Analysis

Once the measurement model for both males and females has been confirmed, the fit of the structural path model posited for these two groups can be evaluated and compared. The factor structure confirmed in the measurement model will be used as the basis for the path model. That is, the four latent constructs BODY, FAMILY, PAIN, and VOLUNTARY EUTHANASIA, together with their respective measurement indicators, will be incorporated into the structure of the path model to be evaluated. Multi-group analysis will then be used to apply this model simultaneously to the male and female samples. The question to be examined is whether the pattern of structural relationships hypothesized in the path model follows the same dynamics for males and females.

TABLE 18.19

Critical Ratios (C.R.) for Differences between Regression Weights For Males and Females (Group Variant Model)

	m3	m7	m5	m4	m2	m6	m8	m1
f3	1.600	6.179	1.996	0.817	1.246	1.855	4.470	1.244
f7	–3.388	1.441	–3.197	–3.222	–4.076	–2.888	–0.126	–4.680
f5	1.942	6.891	**2.392**	1.034	1.587	2.193	4.994	1.636
f4	0.196	4.430	0.508	–0.306	–0.202	0.489	2.930	–0.329
f2	1.645	6.414	2.063	0.824	1.282	1.905	4.616	1.289
f6	1.425	6.395	1.853	0.616	1.035	1.701	4.514	1.021
f8	–0.790	3.761	–0.492	–1.140	–1.273	–0.436	**2.172**	–1.517
f1	1.601	6.571	2.038	0.757	1.223	1.869	4.681	1.231

Pairwise parameter comparisons (group variant)

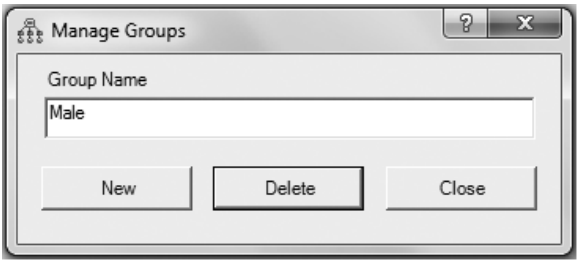
Critical ratios for differences between parameters (group variant)

18.14.2.1 Conducting Multi-Group Modeling for Males and Females: The Path Model

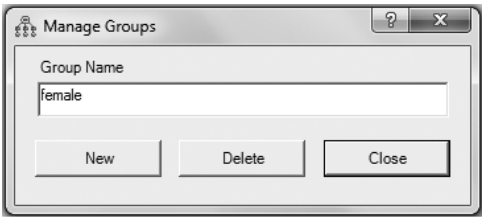
Figure 18.3 presents the original path model for the combined samples of males and females ($N = 357$). To test for sex differences for this path model, it will be necessary to (1) set up separate but identical path models for the male and female samples, (2) link the male and female models to their respective data sets, (3) set up an invariant path model (in which males and females are hypothesized to share the same path coefficients) and a variant path model (in which males and females are hypothesized to have different path coefficients) that can be directly compared as to their model fit, and (4) employ the critical ratio test to test for sex differences in the path coefficients.

1. **Setting up separate but identical path models for the male and female samples.**

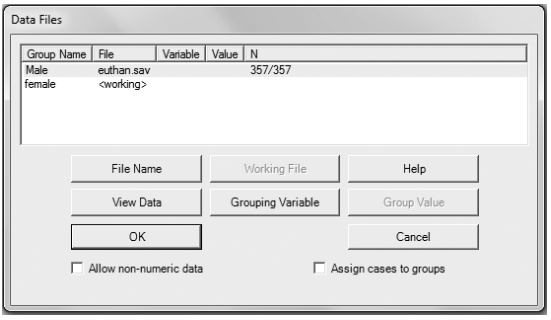
   To do this, retrieve the diagram shown in Figure 18.3. From the drop-down menu bar, click **Analyze** and then **Manage Groups** to open the **Manage Groups** window below. Change the name in the **Group Name** field to **Male**.



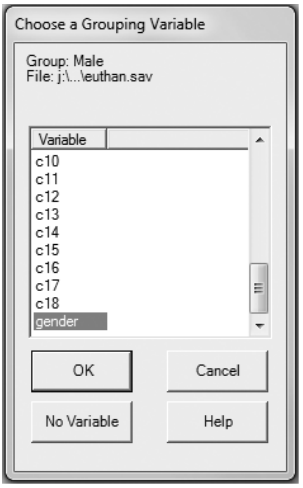
Next, click **New** to open a new **Manage Groups** window. Change the name in the **Group Name** field to **Female**. Click the **Close** button. This will create two identical path models for the male and female samples.

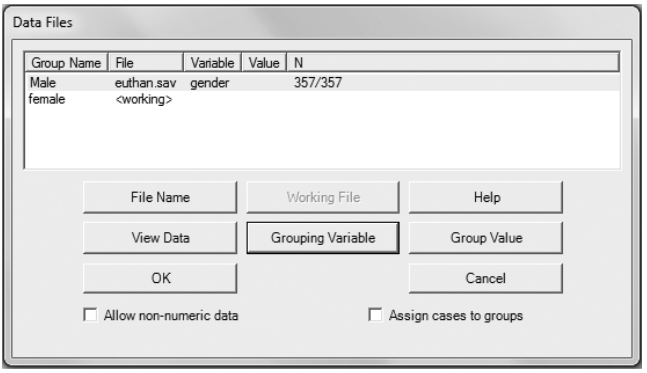


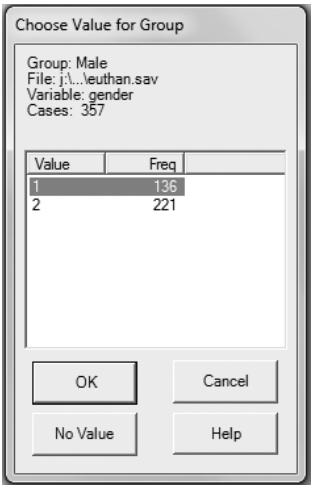
2. **Linking the male and female path models to their respective data sets.**
   - Click from the icon toolbox (**File → Data Files**) to open the **Data Files** window below. Click File Name to search for the data file (**EUTHAN.SAV**) in the computer's directories. Once the data file has been located, open it. By opening the data file, AMOS will automatically link the data file to the path model, as indicated in the window below.



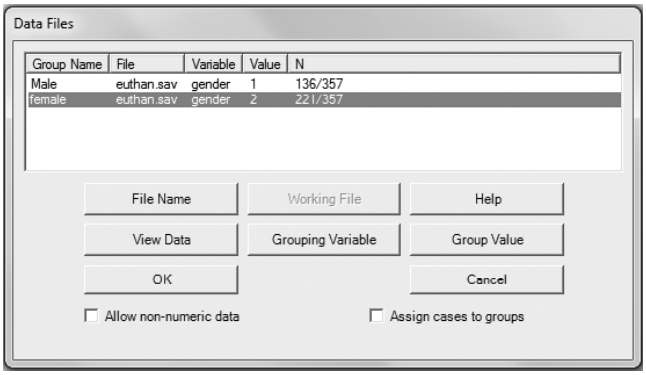
   - To link the male model to the male sub-sample within the EUTHAN.SAV data set, click Grouping Variable to open the following **Choose a Grouping Variable** window.
   - Select **GENDER** as the grouping variable by clicking it (highlight), and then click OK to return to the **Data Files** window. Notice that under the **Variable** heading, **GENDER** has been selected as the grouping variable.
   - To link the male model of the male sub-sample (coded **1** within the **GENDER** variable), click Group Value to open the **Choose Value for Group** window. Select **1** (males) under the **Value** heading by clicking it (highlight).


Choose a Grouping Variable
Group: Male
File: j:\...\euthan.sav
Variable
c10
c11
c12
c13
c14
c15
c16
c17
c18
gender
OK
Cancel
No Variable
Help


Data Files
Group Name
File
Variable
Value
N
Male
euthan.sav
gender
357/357
female
<working>
File Name
Working File
Help
View Data
Grouping Variable
Group Value
OK
Cancel
Allow non-numeric data
Assign cases to groups


Choose Value for Group
Group: Male
File: j:\...\euthan.sav
Variable: gender
Cases: 357
Value
Freq
1
136
2
221
OK
Cancel
No Value
Help

- Click OK to return to the **Data Files** window below. Notice that under the **Value** heading, the number **1** (males) has been selected. Notice also that under the **N** (sample size) heading, the number **136/357** has been listed. This indicates that a sub-sample of 136 males (out of a total sample of 357 subjects) has been linked to the male path model.
- To link the female model to the female sub-sample within the EUTHAN.SAV data set, repeat the above procedure, but use the **Group Value** of **2** to select the female sub-sample. Successful operation of this procedure will yield the final **Data Files** window below. As can be seen, a sub-sample of 221 females (out of a total sample of 357 subjects) has been linked to the female path model. Click OK to exit this window.



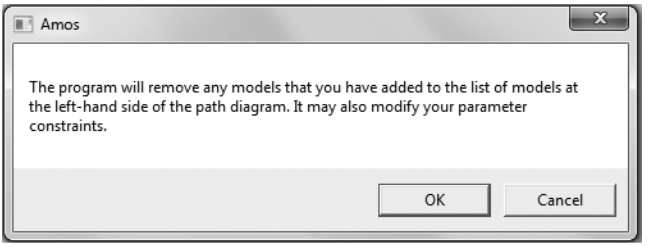
3. **Setting up group-invariant and group-variant path models.**

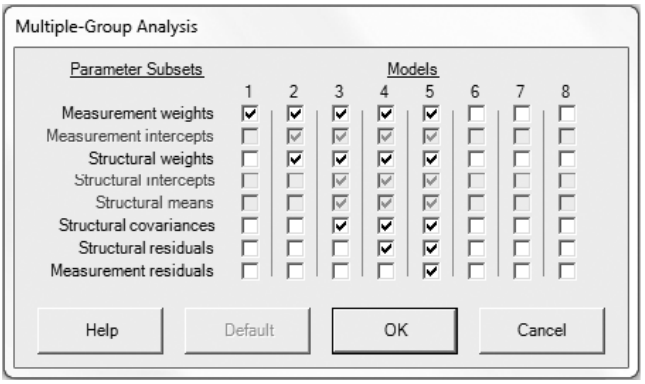
   The hypothesis to be tested is that the path model in Figure 18.3 holds for both males and females. This hypothesis requires that the pattern of relationships (i.e., the path coefficients) is the same for both groups. That is, it requires that every path coefficient for the male sample be *equal* to the corresponding path coefficient for the female sample. However, it does not require that the unique variances for males and females be grouped invariant. The common factor variances and covariances may also differ in the two groups. The rationale underlying the hypothesis of group-invariant path coefficients is that, while it is probably reasonable to assume that the observed and unobserved variables have different variances, covariances, and regression weights among males and females, the process by which the two groups arrived at their decision about voluntary euthanasia may be similar. If the path coefficients are the same for males and females, then the same path coefficients can be used for both groups, which simplifies the prediction of the endogenous variables from the model's exogenous variables.

To setup the **group-invariant** path model (in which males and females are hypothesized to share the same path coefficients), it is necessary to constrain 10 paths (5 path coefficients for males, 5 path coefficients for females) to equality.

Recall that in the prior multi-group confirmatory factor analysis (measurement) model, the critical ratio test for sex differences between the regression weights (factor loadings) yielded significant sex differences for 4 of the 16 regression weights (males and females: **C9** and **E5**) (see Table 18.19). The other 12 regression weights showed no significant gender differences. As the present path model incorporates the factor structure confirmed in the confirmatory factor analysis (measurement) model, it will therefore also be necessary to constrain these 12 regression weights to equality. That is, the two regression weights associated with the measurement variables C9 and E5 (that showed significant gender differences) will be permitted to vary in both the group-invariant and group-variant path models (i.e., they will be estimated separately).

- First, select the male path model by clicking the *male* label Male Female in the AMOS graphic page. Click the multi-group icon: . AMOS will provide the following information.
- Click OK to open the **Multiple-Group Analysis** window below.





- Click OK to bring up the path model as presented in Figure 18.12.

  As can see from Figure 18.12, AMOS has provided default names for all parameters (e.g., paths, variances, covariances, etc.) in the model.

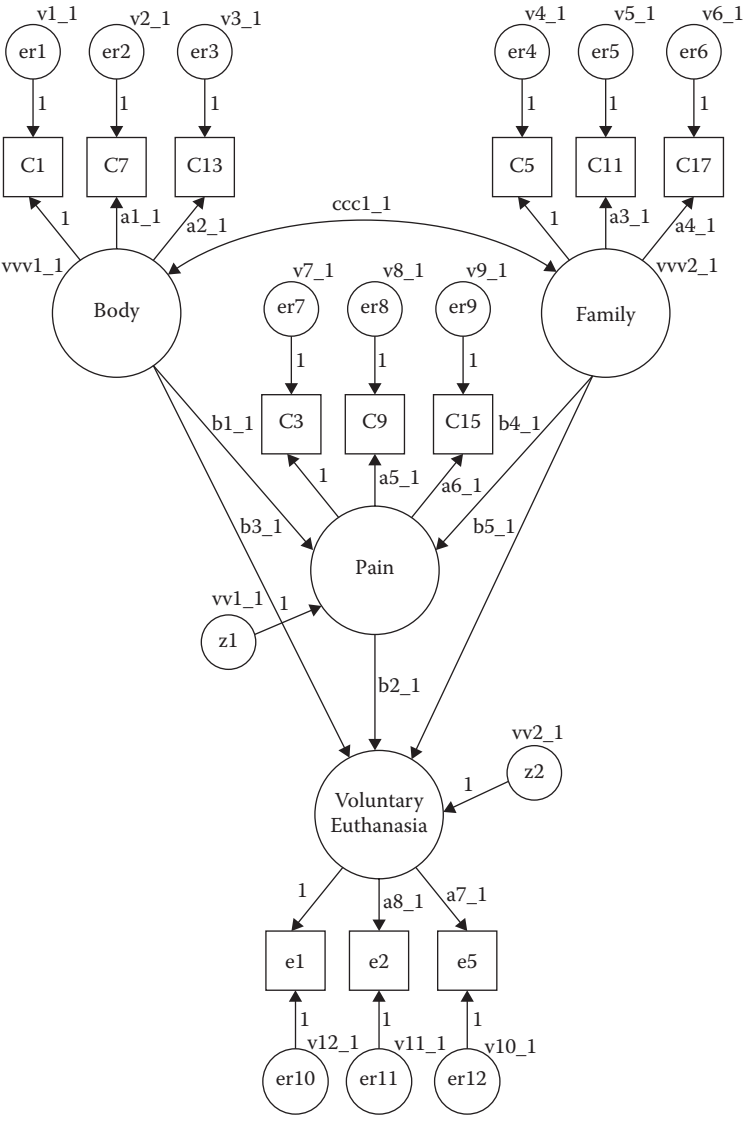
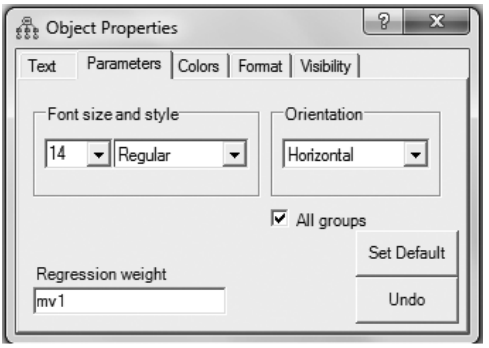
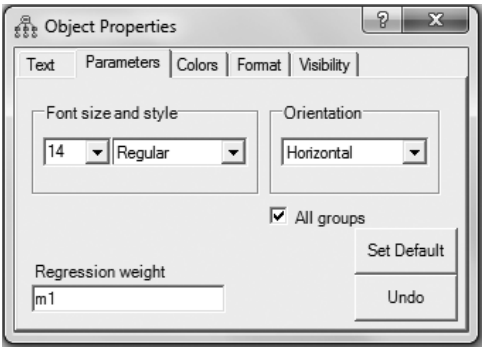


FIGURE 18.12
Path model with default names for parameters.

- In setting up the male path model, delete all parameter names. Rename those parameters associated with the measurement variables C7, C13, C11, C17, C9, C15, E2, and E5 to **mv1**, **mv2**, **mv3**, **mv4**, **mv5**, **mv6**, **mv7**, and **mv8**. For example, to label the regression weight between the latent construct **BODY** and the measurement variable **C7** (BODY → C7), double-click on this parameter. This will open the **Object Properties** window below. Under the **Parameters** tab, enter the label **mv1** in the **Regression weight** field. Continue this procedure until all eight regression weights (**mv1** to **mv8**) have been labeled.



- Label the paths for this male model. For example, to label the path between the latent construct **BODY** and the latent construct **VOLUNTARY EUTHANASIA** (**BODY → VOLUNTARY EUTHAN**), double-click on this parameter. This will open the **Object Properties** window below. Under the **Parameters** tab, enter the label **m1** in the **Regression weight** field. Continue this process until all five path coefficients have been labeled (**m1** to **m5**) (see Figure 18.13).



- Select the female measurement model by clicking the *female* label Male Female in the AMOS graphic page. Repeat the above procedure for labeling the eight regression weights and five paths for this group. For this female path model, label the five regression weights **fv1** to **fv8** and the path coefficients **f1** to **f5** (see Figure 18.14). That is, the labels for the regression weights and the path coefficients for the male and female models must be different, as shown in Figures 18.13 and 18.14.
- In order to set up the **group invariant** model, click **Analyze** and then **Manage Model** (from the drop-down menu bar) to open the **Manage Models** window. Set up the group invariant model by

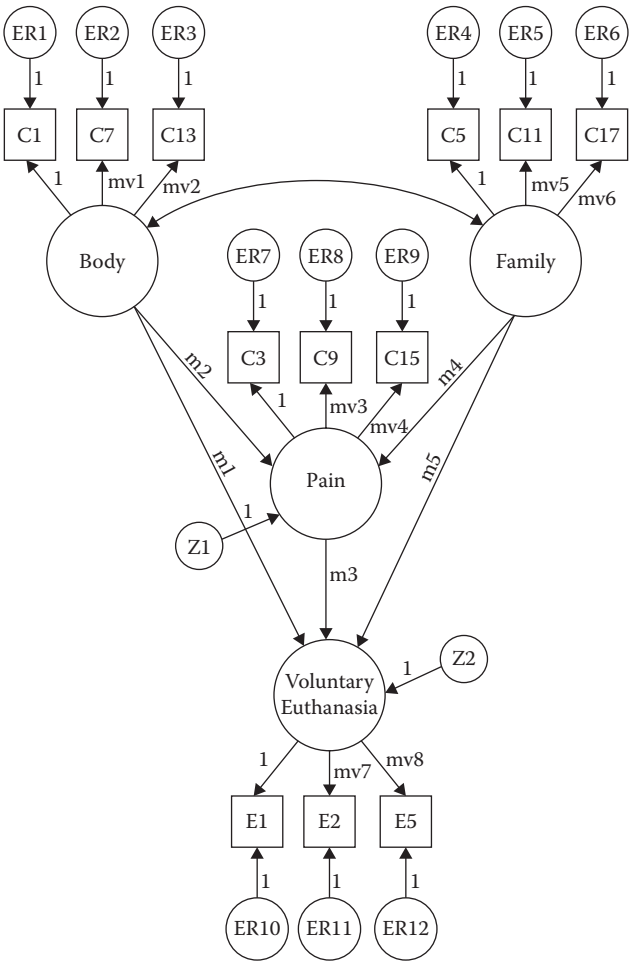


FIGURE 18.13
Male structural path model.

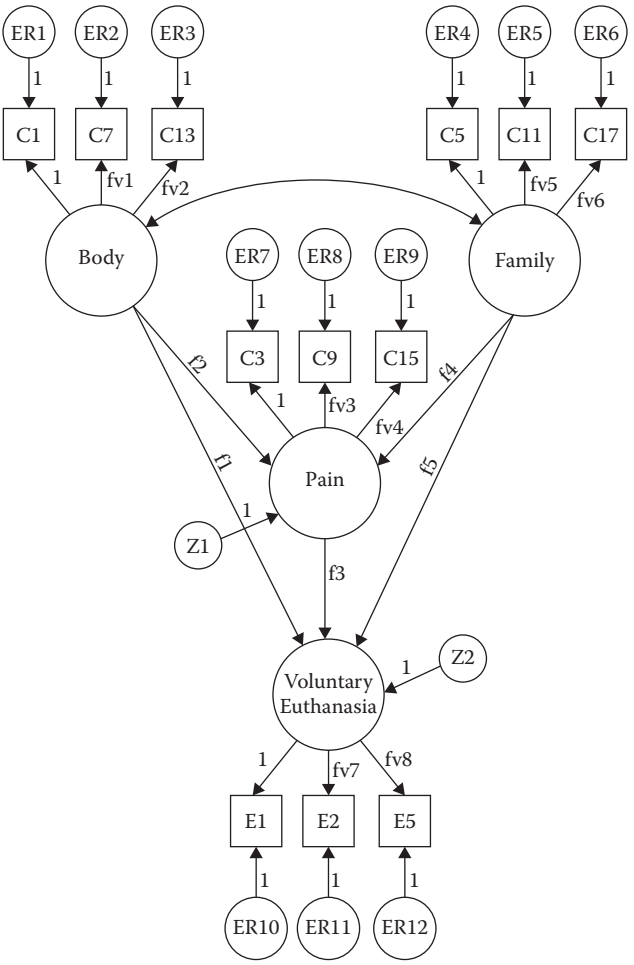
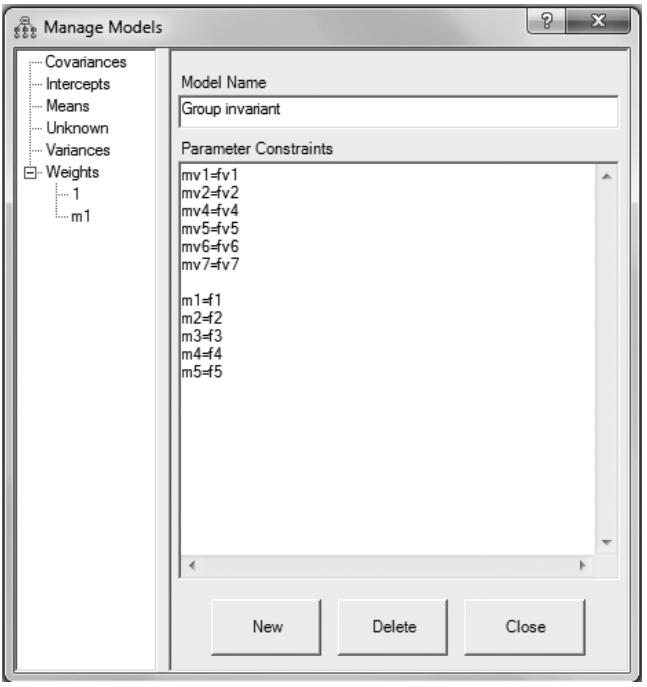


FIGURE 18.14
Female structural path model.

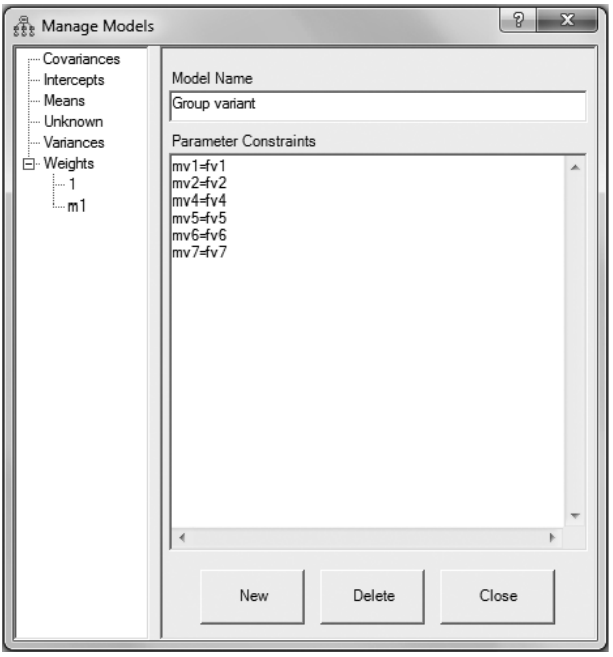
typing in the name **group invariant** in the **Model Name** field, and then entering its constraints in the **Parameter Constraints** field. For the regression weights (factor loadings), the following 12 regression weights will be constrained to equality: **mv1 = fv1**, **mv2 = fv2**, **mv4 = fv4**, **mv5 = fv5**, **mv6 = fv6**, **mv7 = fv7** (the other four regression weights **mv3**, **fv3**, **mv8** and **fv8** will be allowed to vary as they showed significant sex differences in the prior measurement model analysis). For the path coefficients, the following 10 paths will be constrained to equality: **m1 = f1**, **m2 = f2**, **m3 = f3**, **m4 = f4**, **m5 = f5**.

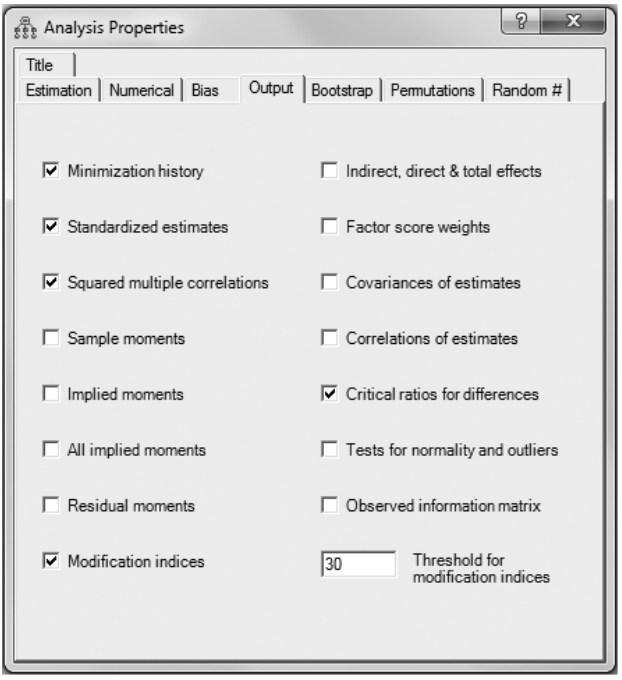


- To set up the **group variant** model (in which males and females are hypothesized to have different regression weights), click New in the **Manage Models** window above. This will open a new **Manage Models** window. In the **Model Name** field, type the name **group variant**. As with the **group invariant** model, this **group variant** model requires the same 12 regression weights to be constrained to equality (**mv1 = fv1**, **mv2 = fv2**, **mv4 = fv4**, **mv5 = fv5**, **mv6 = fv6**, **mv7 = fv7**). However, this model allows the path coefficients for males and females to be estimated separately (i.e., there are no constraints). After typing the regression weight constraints in the **Parameter Constraints** field, click Close to complete this procedure.

4. **Testing for sex differences among the path coefficients.**

   In order to test for sex differences among the path coefficients, the **critical ratio** (C.R.) test can be used to obtain the critical ratio statistics for the differences among male and female subjects' path coefficients. Click the icon (**View → Analysis Properties…**) to open the **Analysis Properties** dialog box. Under the **Output** tab, ensure that the **Critical ratios for differences** field is checked. This option will compare all model parameters posited for the male and female models.


Manage Models
Covariances
Intercepts
Means
Unknown
Variances
Weights
1
m1
Model Name
Group variant
Parameter Constraints
mv1=fv1
mv2=fv2
mv4=fv4
mv5=fv5
mv6=fv6
mv7=fv7
New
Delete
Close


Analysis Properties
Title
Estimation
Numerical
Bias
Output
Bootstrap
Permutations
Random #
Minimization history
Standardized estimates
Squared multiple correlations
Sample moments
Implied moments
All implied moments
Residual moments
Modification indices
Indirect, direct & total effects
Factor score weights
Covariances of estimates
Correlations of estimates
Critical ratios for differences
Tests for normality and outliers
Observed information matrix
30
Threshold for modification indices

After closing the **Analysis Properties** dialog box, click (**Analyze → Calculate Estimates**) to perform the multi-group path analysis.

18.14.2.2 Results and Interpretation

18.14.2.2.1 Notes for Models

For this multi-group path analysis, there are two data sets (for males and females), each containing 12 measurement variables. The two covariance matrices generated from the two data sets contain 156 sample moments.

For the **group invariant** model, there are 49 parameters to be estimated. This model therefore has 107 (156 – 49) degrees of freedom, and yielded a significant chi-square value, $\chi^2(N = 357, df = 107) = 327.318$, $p < .05$ (see Table 18.20).

For the **group variant** model, there are 54 parameters to be estimated. This model, therefore, has 102 (156 – 54) degrees of freedom, and also yielded a significant chi-square value, $\chi^2(N = 357, df = 102) = 322.149$, $p < .05$ (see Table 18.20).

18.14.2.2.2 Summary of Models

Table 18.21 presents the chi-square goodness-of-fit statistics, baseline comparisons fit indices, and model comparison statistics for the group invariant and group variant path models.

TABLE 18.20

Computation of Degrees of Freedom and Chi-Square Goodness-of-Fit Statistics for Group Invariant and Group Variant Path Models

Notes for Model (Group Invariant Model)	
Computation of Degrees of Freedom (Group Invariant Model)	
Number of distinct sample moments:	156
Number of distinct parameters to be estimated:	49
Degrees of freedom (156 – 49):	107

Result (Group invariant model)
Minimum was achieved
Chi-square = 327.318
Degrees of freedom = 107
Probability level = .000

Notes for Model (Group Variant Model)	
Computation of Degrees of Freedom (Group Variant Model)	
Number of distinct sample moments:	156
Number of distinct parameters to be estimated:	54
Degrees of freedom (156 – 54):	102

Result (Group variant model)
Minimum was achieved
Chi-square = 322.149
Degrees of freedom = 102
Probability level = .000

TABLE 18.21

Chi-Square Goodness-of-Fit Statistics, Baseline Comparisons Fit Indices, and Model Comparison Statistics

Model Fit Summary

CMIN

Model	NPAR	CMIN	d*f*	*P*	CMIN/d*f*
Invariant model	49	327.318	107	.000	3.059
Variant model	54	322.149	102	.000	3.158
Saturated model	156	.000	0		
Independence model	24	3442.653	132	.000	26.081

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Invariant model	.905	.883	.934	.918	.933
Variant model	.906	.879	.934	.914	.934
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000
Independence model	3310.653	3123.071	3505.537		

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Invariant model	.076	.067	.086	.000
Variant model	.078	.068	.088	.000
Independence model	.266	.258	.274	.000

AIC

Model	AIC	BCC	BIC	CAIC
Invariant model	425.318	433.656		
Variant model	430.149	439.338		
Saturated model	312.000	338.545		
Independence model	3490.653	3494.737		

(*Continued*)

TABLE 18.21 (*Continued*)

Chi-Square Goodness-of-Fit Statistics, Baseline Comparisons Fit Indices, and Model Comparison Statistics

Nested Model Comparisons							
Assuming Model Variant Model to be Correct							
Model	**d*f***	**CMIN**	**P**	**NFI Delta1**	**IFI Delta2**	**RFI rho1**	**TLI rho2**
Invariant model	5	5.169	.396	.002	.002	–.004	–.004

Although the chi-square values for both path models are statistically significant (i.e., both models yielded poor fit by the chi-square goodness-of-fit test), the baseline comparison fit indices NFI, RFI, IFI, TLI, and CFI for both models are close to or are above 0.90 (range: 0.879–0.934). These values indicate the improvement in fit of both models relative to the null model. Indeed, the only possible improvement in fit for these two models ranges from 0.066 to 0.121.

The *root mean square error of approximation* (**RMSEA**) fit index, which takes into account the error of approximation in the population, yielded values for the group invariant and group variant path models of 0.078 and 0.076, respectively. Values ranging from 0.05 to 0.08 are deemed acceptable (Browne and Cudeck, 1993; MacCallum et al., 1996). Hence, the RMSEA values for the group invariant and group variant path models suggest that the fit of these two models is adequate.

The fit of the two competing models can be directly compared. From the **Nested Model Comparisons** statistics, it can be seen that the chi-square difference value for the two models is 5.169 (327.318 – 322.149). With five degrees of freedom (107 – 102), this value is not significant at the 0.05 level ($p > .05$). Thus, the two models do not differ significantly in their goodness-of-fit.

The fit of the two models can also be compared using the **AIC** (Akaike Information Criterion) measure (Akaike, 1973, 1987). In evaluating the hypothesized model, this measure takes into account both model parsimony and model fit. Simple models that fit well receive low scores, whereas poorly fitting models get high scores. The AIC measure for the group invariant model (425.318) is lower than that for the group variant model (430.149), indicating that the group invariant model is both more parsimonious and better fitting than the group variant model. On the basis of the model comparisons findings, and assuming that the group invariant model is correct, the group invariant model's estimates are preferable over the group variant model's estimates.

18.14.2.2.3 Unstandardized Regression Weights, Standardized Regression Weights, and Squared Multiple Correlations

Table 18.22 presents the unstandardized regression weights, standardized regression weights, and squared multiple correlations for males and females (group invariant model).

TABLE 18.22

Unstandardized Regression Weights, Standardized Regression Weights, and Squared Multiple Correlations for Males and Females

Regression Weights: (Male—Group Invariant)

			Estimate	S.E.	C.R.	*P*	Label
PAIN	←	BODY	.426	.056	7.561	***	m2
PAIN	←	FAMILY	.379	.072	5.297	***	m4
VOLUNTARY_ EUTHANASIA	←	BODY	.114	.064	1.789	.074	m1
VOLUNTARY_ EUTHANASIA	←	PAIN	.612	.073	8.335	***	m3
VOLUNTARY_ EUTHANASIA	←	FAMILY	.049	.076	.646	.518	m5
C1	←	BODY	1.000				
C7	←	BODY	.996	.041	24.154	***	mv1
C13	←	BODY	.961	.041	23.661	***	mv2
C5	←	FAMILY	1.000				
C11	←	FAMILY	1.069	.066	16.214	***	mv5
C17	←	FAMILY	1.037	.065	15.965	***	mv6
C3	←	PAIN	1.000				
C9	←	PAIN	.942	.071	13.314	***	mv3
C15	←	PAIN	.982	.042	23.333	***	mv4
E5	←	VOLUNTARY_ EUTHANASIA	.679	.073	9.319	***	mv8
E2	←	VOLUNTARY_ EUTHANASIA	.588	.043	13.624	***	mv7
E1	←	VOLUNTARY_ EUTHANASIA	1.000				

Scalar estimates (Male—Group invariant)
Maximum likelihood estimates

Standardized Regression Weights: (Male—Group Invariant)

			Estimate
PAIN	←	BODY	.489
PAIN	←	FAMILY	.340
VOLUNTARY_ EUTHANASIA	←	BODY	.123
VOLUNTARY_ EUTHANASIA	←	PAIN	.572
VOLUNTARY_ EUTHANASIA	←	FAMILY	.041
C1	←	BODY	.856
C7	←	BODY	.929
C13	←	BODY	.930
C5	←	FAMILY	.699
C11	←	FAMILY	.901

(*Continued*)

TABLE 18.22

Unstandardized Regression Weights, Standardized Regression Weights, and Squared Multiple Correlations for Males and Females

C17	←	FAMILY	.877
C3	←	PAIN	.807
C9	←	PAIN	.851
C15	←	PAIN	.876
E5	←	VOLUNTARY_ EUTHANASIA	.759
E2	←	VOLUNTARY_ EUTHANASIA	.695
E1	←	VOLUNTARY_ EUTHANASIA	.836

Covariances: (Male—Group Invariant)

			Estimate	S.E.	C.R.	*P*	Label
BODY	↔	FAMILY	.617	.108	5.711	***	par_16

Correlations: (Male—Group Invariant)

			Estimate
BODY	↔	FAMILY	.666

Squared Multiple Correlations: (Male—Group Invariant)

	Estimate
PAIN	.576
VOLUNTARY_EUTHANASIA	.482
E1	.699
E2	.483
E5	.576
C15	.768
C9	.724
C3	.651
C17	.770
C11	.812
C5	.489
C13	.866
C7	.863
C1	.733

Scalar estimates (Female—Group invariant)
Maximum likelihood estimates

Regression Weights: (Female—Group Invariant)

			Estimate	S.E.	C.R.	*P*	Label
PAIN	←	BODY	.426	.056	7.561	***	m2
PAIN	←	FAMILY	.379	.072	5.297	***	m4

TABLE 18.22 (*Continued*)

Unstandardized Regression Weights, Standardized Regression Weights, and Squared Multiple Correlations for Males and Females

Regression Weights: (Female—Group Invariant)

			Estimate	S.E.	C.R.	*P*	Label
VOLUNTARY_ EUTHANASIA	←	BODY	.114	.064	1.789	.074	m1
VOLUNTARY_ EUTHANASIA	←	PAIN	.612	.073	8.335	***	m3
VOLUNTARY_ EUTHANASIA	←	FAMILY	.049	.076	.646	.518	m5
C1	←	BODY	1.000				
C7	←	BODY	.996	.041	24.154	***	mv1
C13	←	BODY	.961	.041	23.661	***	mv2
C5	←	FAMILY	1.000				
C11	←	FAMILY	1.069	.066	16.214	***	mv5
C17	←	FAMILY	1.037	.065	15.965	***	mv6
C3	←	PAIN	1.000				
C9	←	PAIN	1.056	.047	22.656	***	fv3
C15	←	PAIN	.982	.042	23.333	***	mv4
E5	←	VOLUNTARY_ EUTHANASIA	.845	.061	13.932	***	fv8
E2	←	VOLUNTARY_ EUTHANASIA	.588	.043	13.624	***	mv7
E1	←	VOLUNTARY_ EUTHANASIA	1.000				

Standardized Regression Weights: (Female—Group Invariant)

			Estimate
PAIN	←	BODY	.448
PAIN	←	FAMILY	.324
VOLUNTARY_ EUTHANASIA	←	BODY	.127
VOLUNTARY_ EUTHANASIA	←	PAIN	.645
VOLUNTARY_ EUTHANASIA	←	FAMILY	.044
C1	←	BODY	.868
C7	←	BODY	.917
C13	←	BODY	.893
C5	←	FAMILY	.747
C11	←	FAMILY	.899
C17	←	FAMILY	.876
C3	←	PAIN	.908
C9	←	PAIN	.922
C15	←	PAIN	.892
E5	←	VOLUNTARY_ EUTHANASIA	.844
E2	←	VOLUNTARY_ EUTHANASIA	.717
E1	←	VOLUNTARY_ EUTHANASIA	.823

(*Continued*)

TABLE 18.22 (*Continued*)

Unstandardized Regression Weights, Standardized Regression Weights, and Squared Multiple Correlations for Males and Females

Covariances: (Female—Group Invariant)

			Estimate	S.E.	C.R.	*P*	Label
BODY	↔	FAMILY	.529	.079	6.710	***	par_17

Correlations: (Female—Group invariant)

			Estimate
BODY	↔	FAMILY	.615

Squared Multiple Correlations: (Female—Group Invariant)

	Estimate
PAIN	.485
VOLUNTARY_EUTHANASIA	.581
E1	.677
E2	.514
E5	.713
C15	.796
C9	.851
C3	.824
C17	.767
C11	.809
C5	.558
C13	.798
C7	.840
C1	.753

Of the five coefficients associated with the paths linking each sex-based model's exogenous and endogenous variables, three are significant by the critical ratio test (>±1.96, $p < .05$). The two non-significant coefficients for both males and females are associated with the direct paths linking BODY→VOLUNTARY EUTHANASIA and FAMILY→VOLUNTARY EUTHANASIA. These path coefficients for males and females have been incorporated into the models presented in Figures 18.15 and 18.16, respectively.

The **critical ratios for differences** test for sex differences among the path coefficients (see Table 18.23). Please note that the pairwise comparison C.R. The test is carried out on the path coefficients obtained from the group variant model. This is because the path coefficients from the invariant groupmodel or are set to equality and therefore cannot be compared.

From Table 18.23, it can be seen that none of the pairwise comparisons between path coefficients for males and females is significant (<±1. 96, $p > .05$). Hence, the hypothesized structural relationships between the three predictor

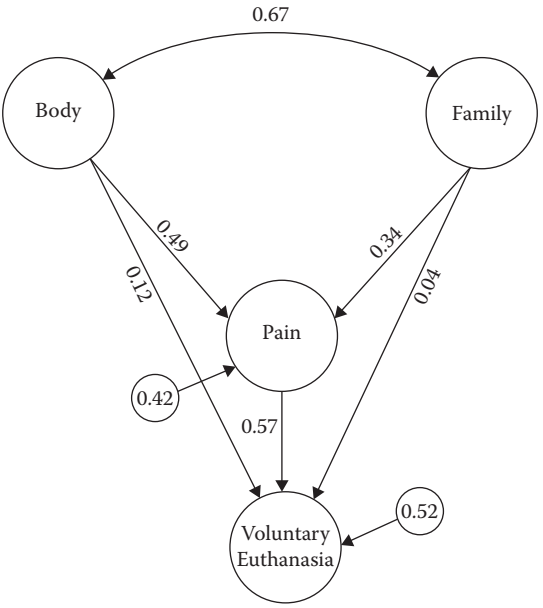


FIGURE 18.15
Male structural path model with standardized path coefficients.

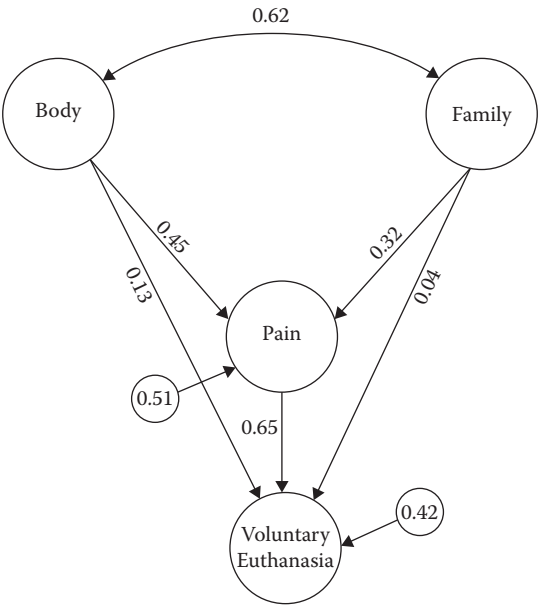


FIGURE 18.16
Female structural path model with standardized path coefficients.

TABLE 18.23

Critical Ratios for Differences between Path Coefficients for Males and Females

	m1	m2	m3	m4	m5	m6
f1	**1.202**	–2.922	–2.230	–1.329	–0.478	–3.377
f2	2.638	**–1.152**	–0.969	0.204	0.804	–1.808
f3	4.283	1.125	**0.649**	2.073	2.380	0.225
f4	2.577	–0.894	–0.824	**0.308**	0.865	–1.525
f5	–0.109	–4.190	–3.249	–2.578	**–1.590**	–4.548
f6	3.616	0.243	–0.007	1.303	1.719	**–0.628**

Pairwise parameter comparisons (Variant model)
Critical ratios for differences between parameters (Variant model)

variables—BODY, FAMILY, PAIN—and support for voluntary euthanasia operated similarly for the male and female samples. Based on these findings, it can be concluded that for males and females, their perception of the debilitated nature of a person's body (BODY) and the negative impact that a person's terminal illness has on his/her family (FAMILY) are related indirectly to their support for voluntary euthanasia, being mediated by their assessment of the PAIN factor. Therefore, the more debilitated a person's body is perceived to be, and the greater the perceived negative impact on the person's family, the greater the physical pain that is perceived to be experienced (Male: $\beta = 0.49$ and $\beta = 0.34$, respectively; Female: $\beta = 0.45$ and $\beta = 0.33$, respectively). The greater the physical pain that is perceived to be experienced, the greater is the support for voluntary euthanasia (Male: $\beta = 0.57$; Female: $\beta = 0.65$).

The **squared multiple correlations** (**SMC**) (see Table 18.22) present the amount of variance in the endogenous variables accounted for by the exogenous variables. For males, the squared multiple correlations show that (1) 57.6% of the variance of PAIN is accounted for by the joint influence of BODY and FAMILY, and (2) 48.2% of the variance of VOLUNTARY EUTHANASIA is accounted for by the joint influence of BODY, FAMILY, and PAIN. For females, the squared multiple correlations show that (1) 48.5% of the variance of PAIN is accounted for by the joint influence of BODY and FAMILY, and (2) 58.1% of the variance of VOLUNTARY EUTHANASIA is accounted for by the joint influence of BODY, FAMILY, and PAIN.

Subtracting the above SMC values from 1 provides the **standardized residuals**. These coefficients provide an estimate of the proportion of variance in each endogenous variable not predicted by its respective model. Thus, these coefficients indicate that for the male path model, 42.4% of the variance in PAIN is not accounted by the joint influence of BODY and FAMILY, and 51.8% of the variance in VOLUNTARY EUTHANASIA is not accounted for by the joint influence of BODY, FAMILY, and PAIN. For the female path model, 51.5% of the variance in PAIN is not accounted for by the joint influence of BODY and FAMILY, and 41.9% of the variance in VOLUNTARY EUTHANASIA is not accounted for by the joint influence of BODY, FAMILY, and PAIN.

18.15 Example 5: Second-Order Confirmatory Factor (CFA) Analysis

This example demonstrates how to conduct a second order CFA and is based on the same variance-covariance matrix (generated from the SPSS file **EUTHAN.SAV**) employed in the previous examples.

The present example will test the hypothesis that the three conditions of suffering (based on Ho's 1999 study)—(1) BODY—*the debilitated nature of one's body*, (2) FAMILY—*the burden placed on one's family*, and (3) PAIN—*the experience of physical pain*—can be most adequately represented by a hierarchical factorial structure. That is, the first-order factors of BODY, FAMILY, and PAIN are explained by a higher order structure that is a single second-order factor of general suffering—SUFFER. Figure 18.17 presents this second-order CFA model.

For this model, the measurement variables representing the three latent constructs **BODY** (C3, C7, C1), **FAMILY** (C17, C11, C5), and **PAIN** (C15, C9, C3) are the same as those presented in the measurement model in Figure 18.4. As these three latent constructs are treated as endogenous variables in the analysis, residual terms (z1, z2, z3) are attached to each construct. For the model to be identified, the variance for the second-order factor of **SUFFER** must be set to 1. CFA is carried out to determine the degree of model fit, the adequacy of the second-order factor loadings, and the standardized residuals and explained variances for the three first-order factors BODY, FAMILY, and PAIN.

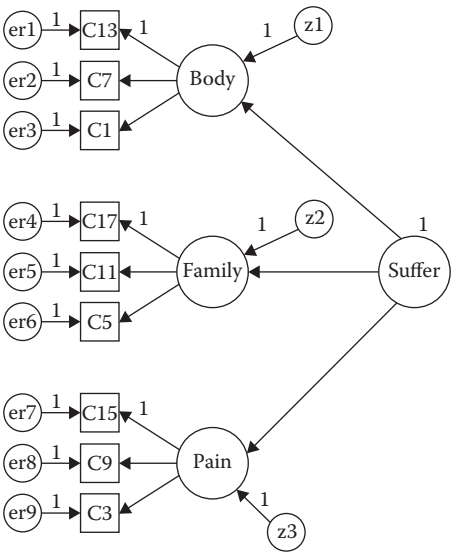


FIGURE 18.17
Second-order confirmatory factor analysis.

The procedure involved in (i) **linking the model to the data set,** (ii) obtaining the desired statistics in the **results output,** and (iii) **conducting the analysis** is identical to that described in Section 18.11.2.

18.15.1 Results and Interpretation

The input covariance matrix generated from the model's 9 measurement variables contains 45 sample moments. For this second-order CFA model, there are 9 regression weights and 12 variances, for a total of 21 parameters to be estimated. The model, therefore, has 24 degrees of freedom (45 – 21), and the chi-square goodness-of-fit statistic was computed (see Table 18.24)

The chi-square goodness-of-fit test shows that the model did not fit the data well, $\chi^2(N = 357, df = 24) = 200.627$, $p < .05$. Although the model did not fit well by the chi-square test, the baseline comparisons fit indices NFI, RFI, IFI, TLI, and CFI are close to or exceed 0.9 (range: 0.89–0.93) (see Table 18.25). Given the range of the computed baseline comparisons fit indices, the remaining possible improvement in the fit of the hypothesized model (range: 0.07–0.11) appears so small as to be of little practical significance.

18.15.1.1 Regression Weights and Standardized Regression Weights

The unstandardized regression weights for the paths linking the second-order factor **SUFFER** with the first-order factors **BODY**, **FAMILY**, and **PAIN** are all significant by the critical ratio test (>±1.96, $p < .05$) (see Table 18.26). The standardized regression weights range from 0.766 to 0.829. These values indicate that the three first-order factors are significantly represented by second-order factor.

TABLE 18.24

Computation of Degrees of Freedom and Chi-Square Statistics for Goodness-of-Fit

Parameter Summary (Group Number 1)						
	Weights	**Covariances**	**Variances**	**Means**	**Intercepts**	**Total**
Fixed	15	0	1	0	0	16
Labeled	0	0	0	0	0	0
Unlabeled	9	0	12	0	0	21
Total	24	0	13	0	0	37

Notes for Model (Default Model)	
Computation of Degree of Freedom (Degree Model)	
Number of distinct sample moments:	45
Number of distinct parameters to be estimated:	21
Degrees of freedom (45 – 21):	24

Result (Default model)
Minimum was achieved
Chi-square = 200.627
Degrees of freedom = 24
Probability level = .000

TABLE 18.25

Incremental Fit Indices

Baseline Comparisons					
Model	**NFI Delta1**	**RFI rho1**	**IFI Delta2**	**TLI rho2**	**CFI**
Default model	.926	.888	.934	.900	.934
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

TABLE 18.26

Regression Weights and Standardized Regression Weights

Regression Weights: (Group Number 1—Default Model)							
			Estimate	**S.E.**	**C.R.**	***P***	**Label**
BODY	←	SUFFER	.835	.056	14.830	***	
FAMILY	←	SUFFER	.669	.051	13.175	***	
PAIN	←	SUFFER	.768	.054	14.112	***	
c1	←	BODY	1.042	.045	23.401	***	
c7	←	BODY	1.044	.039	26.834	***	
c13	←	BODY	1.000				
c5	←	FAMILY	.969	.061	15.948	***	
c11	←	FAMILY	1.033	.049	21.040	***	
c17	←	FAMILY	1.000				
c3	←	PAIN	.982	.046	21.511	***	
c9	←	PAIN	1.032	.043	23.950	***	
c15	←	PAIN	1.000				

Standardized Regression Weights: (Group Number 1—Default Model)			
			Estimate
BODY	←	SUFFER	.829
FAMILY	←	SUFFER	.766
PAIN	←	SUFFER	.802
c1	←	BODY	.861
c7	←	BODY	.923
c13	←	BODY	.906
c5	←	FAMILY	.729
c11	←	FAMILY	.902
c17	←	FAMILY	.876
c3	←	PAIN	.848
c9	←	PAIN	.905
c15	←	PAIN	.891

18.15.1.2 Explained Variances and Residual Variances

The explained variances for the three first-order factors are represented by their squared multiple correlations (see Table 18.27). The percentage of variance explained ranges from 0.587 or 58.7% (FAMILY) to 0.688 or 68.8% (BODY). The residual (unexplained) variances are computed by subtracting each explained variance from 1 (i.e., 1 – squared multiple correlation). Thus, for the three first-order factors, the residual variances range from 31.2% to 41.3%

18.15.1.3 Modification Indices

Examination of the modification indices suggests that the fit of the model can be improved substantially by allowing the error terms er1 (associated with the measurement variable **C13**) and er7 (associated with the measurement variable **C15**) to correlate (see Table 18.28).

As can be seen from the table, allowing these two error terms to correlate will reduce the chi-square value of the modified model by at least 63.09. While this is a substantial decrease (for the loss of 1 degree of freedom), the decision to implement or not to implement this modification rests with the researcher, and in particular, on the **theoretical justification** for this

TABLE 18.27

Explained Variances (Squared Multiple Correlations) for the 12 Measurement Variables

Squared Multiple Correlations: (Group Number 1—Default Model)	
	Estimate
PAIN	.644
FAMILY	.587
BODY	.688
c15	.793
c9	.818
c3	.719
c17	.768
c11	.814
c5	.531
c13	.821
c7	.851
c1	.742

TABLE 18.28

Modification Indices

Covariances: (Group Number 1—Default Model)				
			M.I.	**Par change**
er1	↔	er7	63.090	.138

modification. As mentioned earlier, without strong theoretical justification, employing the values of the modification indices to improve model fit increases the probability that the researcher is capitalizing on the uniqueness of the particular data set, and the results will most likely be atheoretical.

Similar to the arguments presented for the first-order confirmatory factor analysis model (Example 3) presented in Section 18.13.2, the motivation to include correlated errors in a modified model is twofold. First, allowing the error terms er1 and er7 to correlate will reduce the chi-square goodness-of-fit value substantially (i.e., improving the model fit). While this reason does not lower the probability that the strategy may improve the fit of the model by capitalizing on chance, it does have a legitimate place in exploratory studies (Arbuckle and Wothke, 1999). Second, and more significantly, the two measurement variables (C13, C15) associated with the error terms er1 and er7 appear to share something in common, above and beyond the latent constructs they were written to represent. Both items C13 (*A person has lost control of all his/her bodily functions*) and C15 (*This person's terminal illness has given him/her continuous excruciating pain*) appear to reflect the physical pain commonly associated with a debilitated body.

18.15.1.3.1 The Modified Model

This modification was carried out and the modified model was re-estimated. The correlation between two error terms is achieved in AMOS Graphics by joining the error terms with a double-pointed arrow (see Figure 18.18).

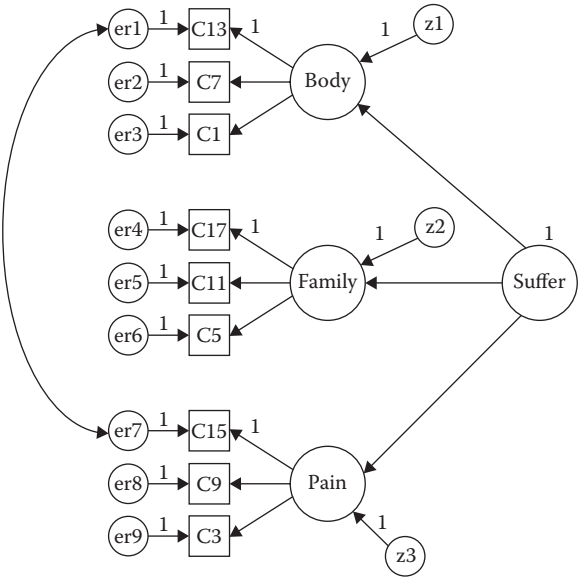


FIGURE 18.18
Modified second-order CFA model.

Table 18.29 presents the chi-square goodness-of-fit value, and the unstandardized and standardized regression weights for the modified model.

The chi-square goodness-of-fit value for this modified model (130.909) is smaller than the chi-square value obtained for the original model (200.627).

TABLE 18.29

Chi-Square Goodness-of-Fit Value, Unstandardized and Standardized Regression Weights for the Modified Second-Order CFA Model

Regression Weights: (Group Number 1—Default Model)							
			Estimate	**S.E.**	**C.R.**	***P***	**Label**
BODY	←	SUFFER	.829	.057	14.536	***	
FAMILY	←	SUFFER	.678	.051	13.235	***	
PAIN	←	SUFFER	.756	.055	13.737	***	
c1	←	BODY	1.036	.043	24.106	***	
c7	←	BODY	1.037	.037	27.765	***	
c13	←	BODY	1.000				
c5	←	FAMILY	.973	.061	15.912	***	
c11	←	FAMILY	1.041	.050	21.021	***	
c17	←	FAMILY	1.000				
c3	←	PAIN	.971	.044	21.916	***	
c9	←	PAIN	1.034	.041	25.050	***	
c15	←	PAIN	1.000				

Result (Default model)
Minimum was achieved
Chi-square = 130.909
Degrees of freedom = 23
Probability level = .000
Scalar estimates (Group number 1—Default model)
Maximum likelihood estimates

Standardized Regression Weights: (Group Number 1—Default Model)			
			Estimate
BODY	←	SUFFER	.816
FAMILY	←	SUFFER	.780
PAIN	←	SUFFER	.783
c1	←	BODY	.865
c7	←	BODY	.925
c13	←	BODY	.905
c5	←	FAMILY	.729
c11	←	FAMILY	.905
c17	←	FAMILY	.872
c3	←	PAIN	.846
c9	←	PAIN	.914
c15	←	PAIN	.889

With a smaller chi-square value, the modified model, therefore, represents a better fit to the data than the original model. Still, the question remains as to whether the improvement in fit represents a statistically significant improvement.

18.15.1.3.2 Comparing the Original (Default) Model against the Modified Model

In this example, a direct comparison in goodness-of-fit between the original and modified models is possible because both models are based on the same data set, and have different degrees of freedom. A test of the original model against the modified model can be obtained by subtracting the smaller chi-square value from the larger one. In this example, the comparison chi-square value is 69.718 (i.e., 200.627 – 130.909). If the original model is correctly specified, this value will have an approximate chi-square distribution with degrees of freedom equal to the difference between the degrees of freedom of the competing models. In this example, the difference in degrees of freedom is 1 (i.e., 24 – 23). Therefore, with 1 degree of freedom, a chi-square value of 69.718 is significant at the 0.05 level. On the basis of this test, it can be concluded that the modified model (with correlated errors) represents a *significantly* better fit to the data than the original model. Please note that the regression weights obtained under the modified model remain statistically significant by the critical ratio test, and are highly similar to those obtained under the original model.

Finally, it should be noted that the difference between the first-order CFA model (see Figure 18.4) and the present second-order CFA model is that the second-order model is a special case of the first-order model in which structure is imposed on the correlation pattern among the first-order factors (Rindskopf and Rose, 1988). However, as pointed out by Byrne (2001), the decision to model a particular scale as a first-order or as a second-order structure ultimately depends on the substantive meaningfulness of the underlying theory.

19

Nonparametric Tests

19.1 Aim

Most of the tests covered in this book (e.g., *t* tests, analysis of variance) are **parametric** tests in that they depend considerably on population characteristics, or parameters, for their use. The *t* test, for instance, uses the sample's mean and standard deviation statistics to estimate the values of the population parameters. Parametric tests also assume that the scores being analyzed come from populations that are *normally distributed* and have *equal variances*. In practice however, the information collected may violate one or both of these assumptions. However, because parametric inference tests are thought to be robust with respect to violations of underlying assumptions, researchers may use these tests even if the assumptions are not met.

When the data collected flagrantly violate the above assumptions, the researcher must select an appropriate nonparametric test. Nonparametric inference tests have fewer requirements or assumptions about population characteristics. For example, to use these tests, it is not necessary to know the means, standard deviations, or shape of the population scores. Because nonparametric tests make no assumptions about the form of the populations from which the test samples were drawn, they are often referred to as *distribution-free* tests.

The following nonparametric techniques are presented in this chapter:

- Chi-square (χ^2) test for single variable experiments
- Chi-square (χ^2) test of independence between two variables
- Mann-Whitney *U* test for two independent samples
- Kruskal-Wallis test for several independent samples
- Wilcoxon signed rank test for two related samples
- Friedman test for several related samples

19.2 Chi-Square (χ^2) Test for Single Variable Experiments

The chi-square inference test is most often used with nominal data, where observations are grouped into several discrete, mutually exclusive categories, and where one counts the frequency of occurrence in each category. The single variable chi-square test compares the observed frequencies of categories to frequencies that would be expected if the null hypothesis was true. The chi-square statistic is computed by comparing the observed values against the expected values for each of the categories and examining the differences between them.

19.2.1 Assumptions

- The data is presumed to be a random sample.
- Independence between each observation recorded in each category. That is, each subject can only have one entry in the chi-square table.
- The expected frequency for each category should be at least 5.

19.2.2 Example 1: Equal Expected Frequencies

Suppose a researcher is interested in determining whether there is a difference among 18-year-olds living in Bangkok, Thailand in their preference for three different brands of cola. The researcher decides to conduct an experiment in which he randomly samples 42 18-year-olds and lets them try out the three different brands. The information recorded in each cell of the table below are the number or frequency of subjects appropriate to that cell. Therefore, 10 subjects preferred Brand A; 10 subjects preferred Brand B; and 22 subjects preferred Brand C. Can the researcher conclude from these data that there is a difference in preference in the population?

Brand A	Brand B	Brand C	Total
10	10	22	42

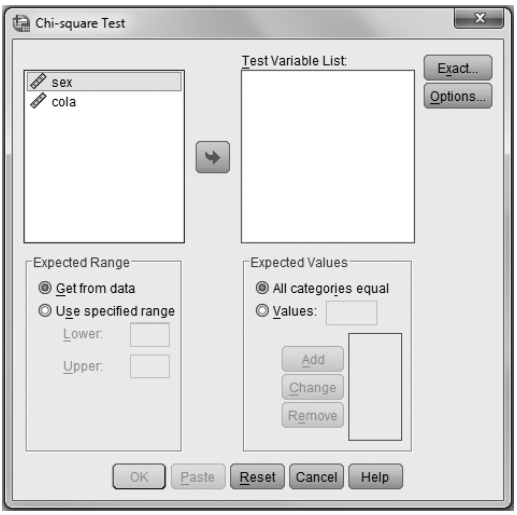
19.2.2.1 Data Entry Format

The data set has been saved under the name **EX19a.SAV**.

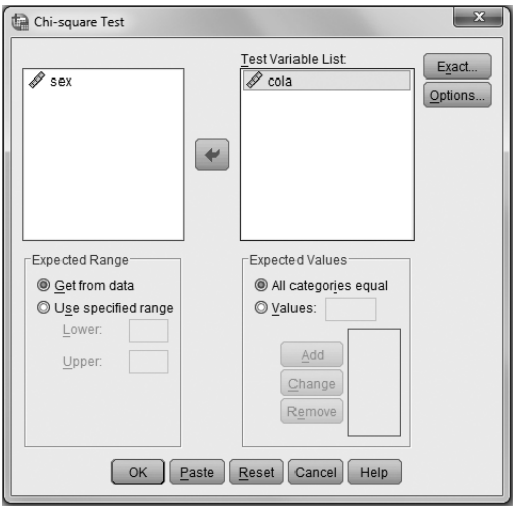
Variables	Column(s)	Code
• SEX	1	1 = male, 2 = female
• COLA	2	1 = Brand A, 2 = Brand B, 3 = Brand C

19.2.2.2 Windows Method

1. From the menu bar, click **Analyze**, **Nonparametric Tests**, **Legacy Dialogs**, and then **Chi-square**. The following **Chi-square Test** window will open.



2. In the field containing the study's variables (**SEX**, **COLA**), click (highlight) the **COLA** variable, and then click to transfer this variable to the **Test Variable List:** field. Ensure that the **All categories equal** cell is checked (this is the default).



3. Click OK to run the chi-square analysis. See Table 19.1 for the results.

19.2.2.3 SPSS Syntax Method

```
NPAR TESTS CHISQUARE = COLA
/EXPECTED = EQUAL.
```

19.2.2.4 SPSS Output

TABLE 19.1

Chi-Square Output for Single Variable (Equal Expected Frequencies)

NPar Tests

Chi-Square Test

Frequencies

Cola

	Observed *N*	**Expected *N***	**Residual**
Brand A	12	14.0	−2.0
Brand B	13	14.0	−1.0
Brand C	17	14.0	3.0
Total	42		

Test Statistics

	Cola
Chi-square	1.000[a]
D*f*	2
Asymptotic significance	.607

[a] 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 14.0.

19.2.2.5 Results and Interpretation

From the **Test Statistics** table, it can be seen that the chi-square value is not significant, χ^2 ($df = 2$) = 1.00, $p >$ .05. There is no difference in the population regarding preference for the three brands of cola.

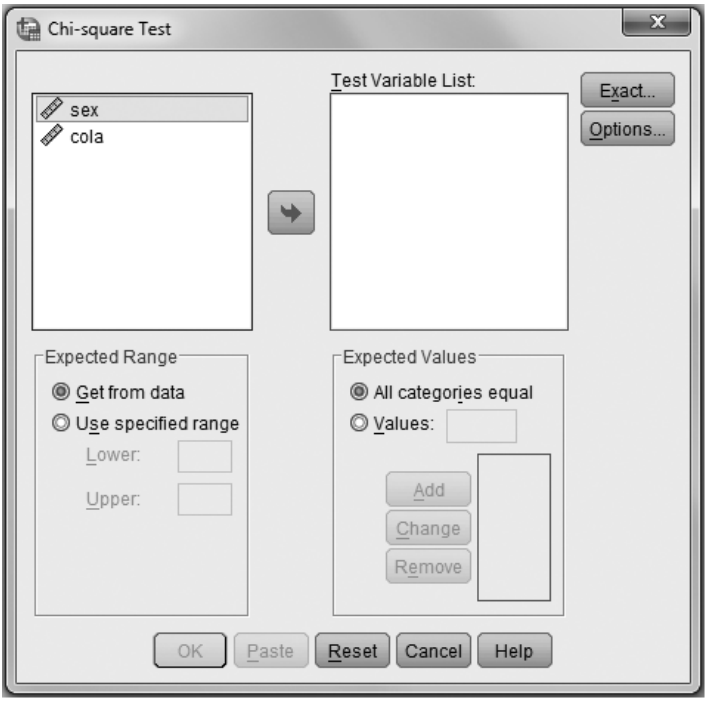
19.2.3 Example 2: Unequal Expected Frequencies

In the above example, the expected frequencies for the three categories were set to be equal (14 subjects per category). That is, the null hypothesis states that, in the population, the proportion of individuals favoring Brand A is equal to the proportion favoring Brand B, which is equal to the proportion favoring Brand C. This is the default null hypothesis tested. However, the researcher can change the expected frequencies in the

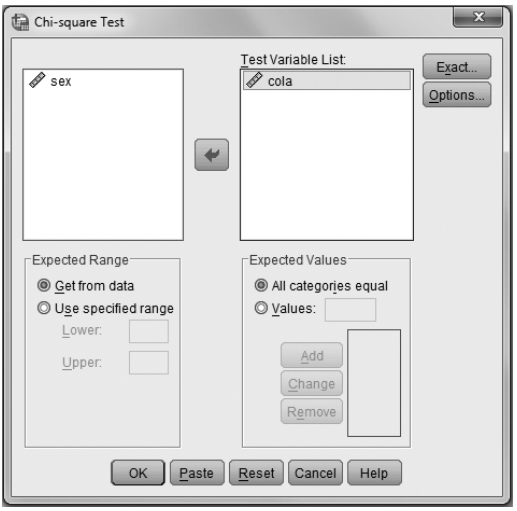
categories to reflect certain expectations about the frequency distribution in the population. Suppose that the researcher expects that the frequency distribution of the 42 18-year-olds across the three brands of cola is not the same. More specifically, the researcher expects that 9 subjects prefer Brand A, 5 prefer Brand B, and 28 prefer Brand C. The research question then is whether there is a difference between the observations and what the researcher expects.

19.2.3.1 Windows Method

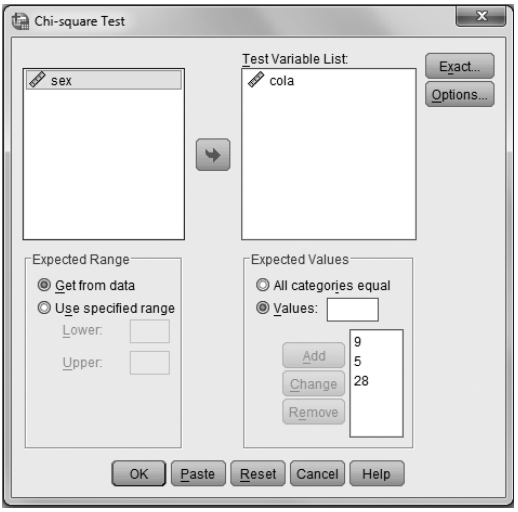
1. From the menu bar, click **Analyze**, **Nonparametric Tests**, **Legacy Dialogs**, and then **Chi-square**. The following **Chi-square Test** window will open.



2. In the field containing the study's variables (**SEX**, **COLA**), click (highlight) the **COLA** variable, and then click to transfer this variable to the **Test Variable List:** field. Ensure that the **All categories equal** cell is checked (this is the default).



3. Under **Expected Values**, check the **Values:** cell. Type 9 (the expected frequency for Brand A) in the **Values:** field. Next, type 5 (the expected frequency for Brand B) in the **Values:** field. Finally, type 28 (the expected frequency for Brand C) in the **Values:** field.



4. Click OK to run the chi-square analysis. See Table 19.2 for the results.

19.2.3.2 SPSS Syntax Method

```
NPAR TESTS CHISQUARE = COLA
/EXPECTED = 9 5 28.
```

19.2.3.3 SPSS Output

TABLE 19.2

Chi-Square Output for Single Variable (Unequal Expected Frequencies)

NPar Tests

Chi-Square Test

Frequencies

Cola

	Observed *N*	Expected *N*	Residual
Brand A	12	9.0	3.0
Brand B	13	5.0	8.0
Brand C	17	28.0	−11.0
Total	42		

Test Statistics

	Cola
Chi-square	18.121[a]
d*f*	2
Asymptotic significance	.000

[a] 0 cells (0.0%) have expected frequencies less than 5. The minimum expected cell frequency is 5.0.

19.2.3.4 Results and Interpretation

From the **Test Statistics** table, it can be seen that chi-square value is significant, $\chi^2(df = 2) = 18.12$, $p < .001$. The observed distribution of preferences for the three brands of cola differs from what the researcher expected.

19.3 Chi-Square (χ^2) Test of Independence between Two Variables

The primary function of the chi-square test of independence is to determine whether two categorical variables are independent or are related. To illustrate, let's suppose that the researcher in Example 1 above is also interested in finding out whether there is a relationship between preference for the three brands of cola and the gender of the 18-year-old subjects. The results are shown in the 2×3 contingency table below. This example will employ the data set **EX19a.SAV**.

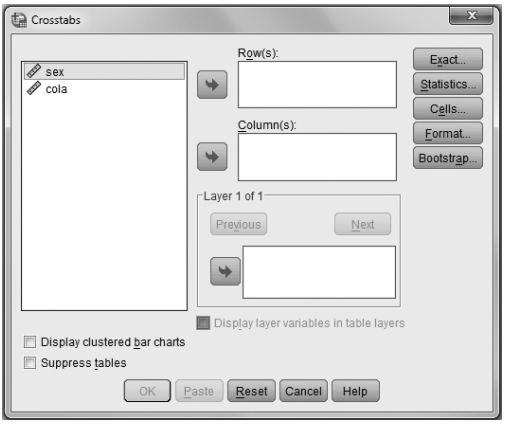
	Brand A	Brand B	Brand C
Male	4	2	15
Female	8	11	2

19.3.1 Assumptions

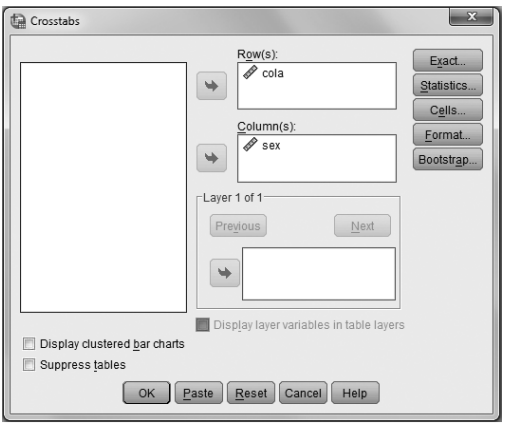
- Data are assumed to be a **random sample**.
- **Independence between each observation** recorded in the contingency table. That is, each subject can only have one entry in the chi-square table.
- The expected frequency for each category should be at least 5.

19.3.2 Windows Method

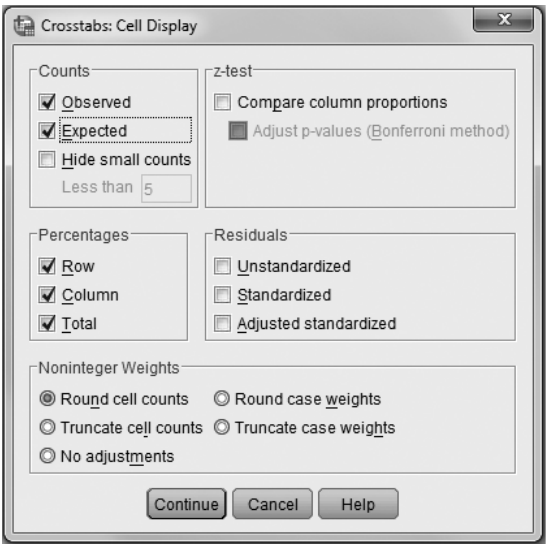
1. From the menu bar, click **Analyze**, then **Descriptive Statistics**, and then **Crosstabs**. The following **Crosstabs** window will open.



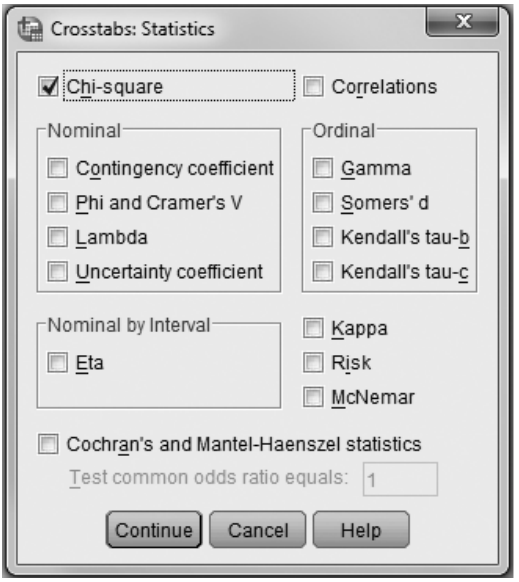
2. In the field containing the study's variables, click (highlight) the **COLA** variable, and then click ▶ to transfer this variable to the **Row(s):** field. Next, click (highlight) the **SEX** variable, and then click ▶ to transfer this variable to the **Column(s):** field.



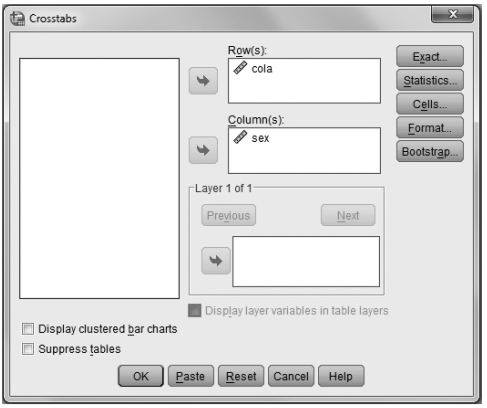
3. Click Cells... to open the **Crosstabs: Cell Display** window below. Under **Counts**, check the **Observed** and **Expected** cells. Under **Percentages**, check the **Row**, **Column**, and **Total** cells. Click Continue to return to the **Crosstabs** window.



4. When the **Crosstabs** window opens, click Statistics... to open the **Crosstabs: Statistics** window below. Check the **Chi-square** cell, and then click Continue to return to the **Crosstabs** window.



5. When the **Crosstabs** window opens, click OK to run the analysis. See Table 19.3 for the results.



19.3.3 SPSS Syntax Method

```
CROSSTABS TABLES = COLA BY SEX
/CELLS = COUNT ROW COLUMN TOTAL EXPECTED
/STATISTICS = CHISQ.
```

19.3.4 SPSS Output

TABLE 19.3

Chi-Square Output for Test of Independence between Two Variables

Cola * Sex Crosstabulation					
			Sex		
			Male	Female	Total
Cola	Brand A	Count	4	8	12
		Expected count	6.0	6.0	12.0
		% Within cola	33.3%	66.7%	100.0%
		% Within sex	19.0%	38.1%	28.6%
		% of Total	9.5%	19.0%	28.6%
	Brand B	Count	2	11	13
		Expected count	6.5	6.5	13.0
		% Within cola	15.4%	84.6%	100.0%
		% Within sex	9.5%	52.4%	31.0%
		% of Total	4.8%	26.2%	31.0%
	Brand C	Count	15	2	17
		Expected count	8.5	8.5	17.0
		% Within cola	88.2%	11.8%	100.0%
		% Within sex	71.4%	9.5%	40.5%
		% of Total	35.7%	4.8%	40.5%

TABLE 19.3 (*Continued*)

Chi-Square Output for Test of Independence between Two Variables

Cola * Sex Crosstabulation

		Sex		
		Male	**Female**	**Total**
Total	Count	21	21	42
	Expected count	21.0	21.0	42.0
	% Within cola	50.0%	50.0%	100.0%
	% Within sex	100.0%	100.0%	100.0%
	% of Total	50.0%	50.0%	100.0%

Chi-Square Tests

	Value	**df**	**Asymp. Sig. (2-Sided)**
Pearson Chi-square	17.505[a]	2	.000
Likelihood ratio	19.470	2	.000
Linear-by-linear association	9.932	1	.002
N of valid cases	42		

[a] 0 cells (.0%) have expected count less than 5. The minimum expected count is 6.00.

19.3.5 Results and Interpretation

The results show that the **expected Count frequency** in each of the six cells generated by the factorial combination of SEX and COLA is greater than 5. This means that the analysis has not violated a primary assumption underlying the chi-square test.

The **Pearson Chi-Square** statistic is used to determine whether there is a relationship between preference for the three brands of cola and the gender of the 18-year-old subjects. The **Pearson Chi-Square** value is statistically significant, $\chi^2(df = 2) = 17.51$, $p < .001$. This means that preference for the three brands of cola varied as a function of the subject's gender. Looking at the **Cola*Sex Crosstabulation** table, it can be seen that the majority of the male subjects prefer Brand C (**Count** = 15; **% Within sex** = 71.4%) over Brand A (**Count** = 4; % **Within sex** = 19%) and Brand B (**Count** = 2; % **Within sex** = 9.5%). For female subjects, their preference was for Brand B (**Count** = 11; **% Within sex** = 52.4%), followed by Brand A (**Count** = 8; **% Within sex** = 38.1%) and Brand C (**Count** = 2; **% Within sex** = 9.5%).

19.4 Mann-Whitney *U* Test for Two Independent Samples

The Mann-Whitney U test is a nonparametric test for a between-subjects design using two levels of an independent variable and scores that are measured at least at the ordinal level. It is often used in place of the *t* test for independent groups when there is an extreme violation of the normality assumption or when the data are scaled at a level that is not appropriate for the *t* test.

Suppose that the following data have been collected representing monthly incomes for teachers employed in two different private schools. The researcher wishes to determine whether there is a substantial difference between these incomes.

School A		School B	
Subjects	**Weekly Income ($)**	**Subjects**	**Weekly Income ($)**
s1	870	s10	1310
s2	720	s11	940
s3	650	s12	770
s4	540	s13	880
s5	670	s14	1160
s6	760	s15	900
s7	730	s16	870
s8	820	s17	760
s9	104	s18	950
		s19	1640
		s20	1270
		s21	770

19.4.1 Assumptions

- The data must be from **independent random samples**.
- The data must be **measured at least at the ordinal level**.
- The **underlying dimension of the dependent variable is continuous in nature**, even though the actual measurements may be only ordinal in nature.

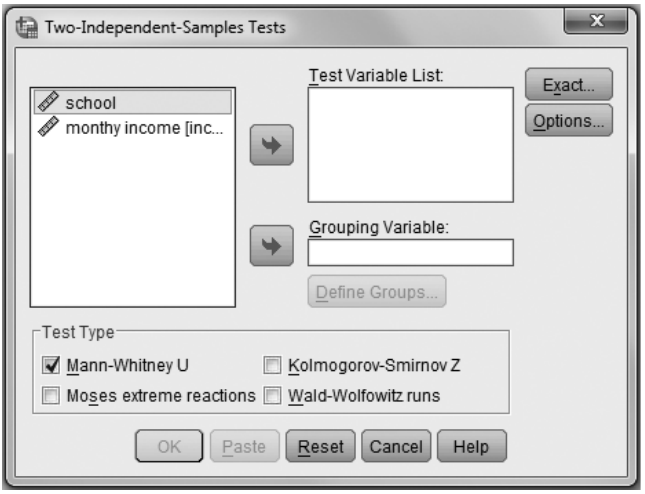
19.4.2 Data Entry Format

The data set has been saved under the name **EX19b.SAV**.

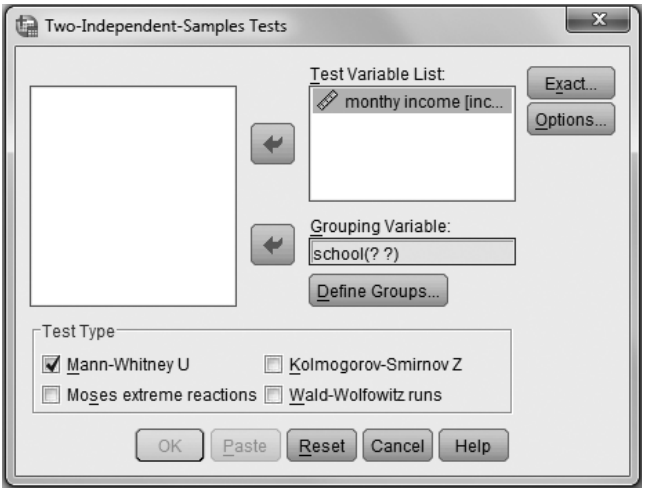
Variables	Column(s)	Code
• SCHOOL	1	1 = School A, 2 = School B
• INCOME	2	Monthly income

19.4.3 Windows Method

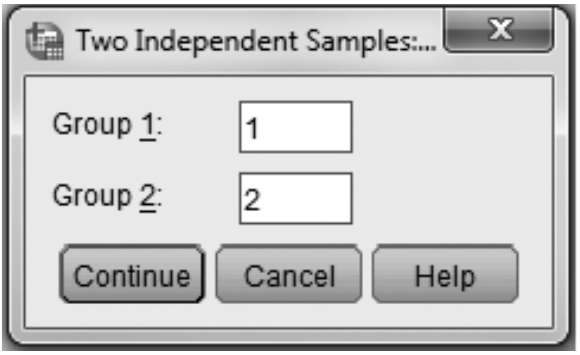
1. From the menu bar, click **Analyze**, then **Nonparametric Tests**, then **Legacy Dialogs**, and then **2 Independent Samples**. The following **Two-Independent-Samples Tests** window will open.



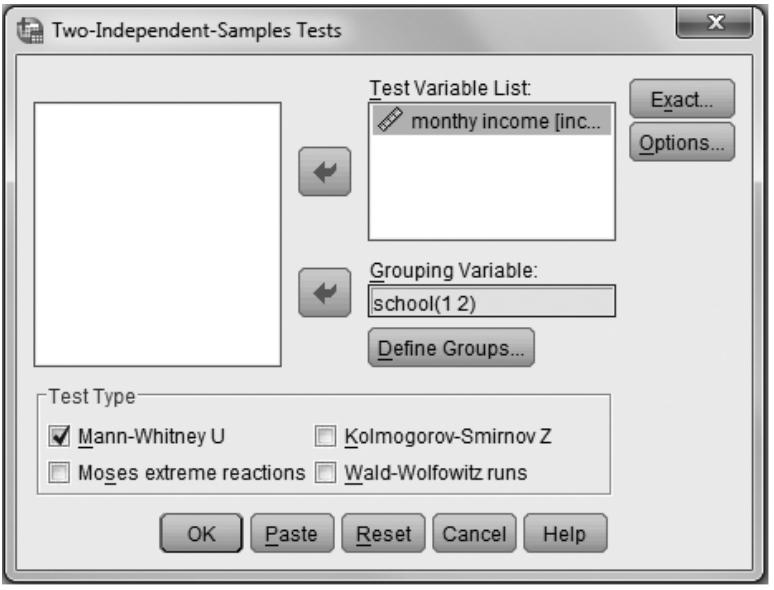
2. Since **SCHOOL** is the grouping (independent) variable, transfer it to the **Grouping Variable:** field by clicking (highlight) the variable and then clicking ▶, As **INCOME** is the test (dependent) variable, transfer it to the **Test Variable List:** field by clicking (highlight) the variable and then clicking ▶.



3. Click Define Groups... to define the ranges for the grouping variable **SCHOOL** (coded 1 = School A, 2 = School B). When the following **Two Independent Samples** window opens, type **1** in the **Group 1:** field and **2** in the **Group 2:** field, and then click Continue.



4. When the following **Two-Independent-Samples Tests** window opens, ensure that the **Mann-Whitney *U*** cell is checked. Run the analysis by clicking OK. See Table 19.4 for the results.



19.4.4 SPSS Syntax Method

```
NPAR TESTS M-W = INCOME BY SCHOOL(1,2).
```

19.4.5 SPSS Output

TABLE 19.4

Mann-Whitney Test Output

Mann-Whitney Test

Ranks

	School	*N*	Mean Rank	Sum of Ranks
Monthly income	School A	9	6.67	60.00
	School B	12	14.25	171.00
	Total	21		

Test Statistics[b]

	Monthy Income
Mann-Whitney *U*	15.000
Wilcoxon *W*	60.000
Z	−2.774
Asymp. sig. (2-tailed)	.006
Exact sig. [2*(1-tailed sig.)]	.004[a]

[a] Not corrected for ties.
[b] Grouping Variable: school.

19.4.6 Results and Interpretation

The hypothesis tested by the Mann-Whitney analysis is that the *medians* of the two groups are equal. The obtained Mann-Whitney *U* statistic is 15. This value, when corrected for tied rankings and converted to a *z*-score (critical ratio test) is significant at the 0.006 level. This means that the probability of the two medians being the same is very small. Thus, it can be concluded that there is a significant difference between the median incomes of teachers in the two private schools.

19.5 Kruskal-Wallis Test for Several Independent Samples

The Kruskal-Wallis test is a nonparametric test that is used with an independent groups design comprising more than two groups. It is a nonparametric version of the one-way ANOVA discussed in Chapter 6, and is calculated on the basis of the sums of the ranks of the combined groups. It is used when violations of assumptions underlying parametric tests (e.g., population normality, homogeneity of variance) are extreme.

Suppose that an educational psychologist is interested in determining the effectiveness of three different methods of instruction in the basic principles of

arithmetic. A total of 29 primary grade children was randomly assigned to the three "methods of instruction" conditions. The scores recorded below are the number of correct answers obtained following completion of their instruction.

Instruction Method					
A		**B**		**C**	
s1	4	s10	12	s19	1
s2	5	s11	8	s20	3
s3	4	s12	10	s21	4
s4	3	s13	5	s22	6
s5	6	s14	7	s23	8
s6	10	s15	9	s24	5
s7	1	s16	14	s25	3
s8	8	s17	9	s26	2
s9	5	s18	4	s27	2

19.5.1 Assumptions

- The data must be from **independent random samples**.
- The data must be **measured at least at the ordinal level**.
- There **must be at least five scores in each sample** to use the chi-square probabilities.

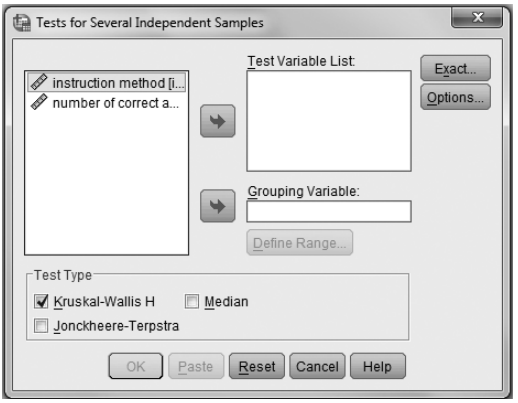
19.5.2 Data Entry Format

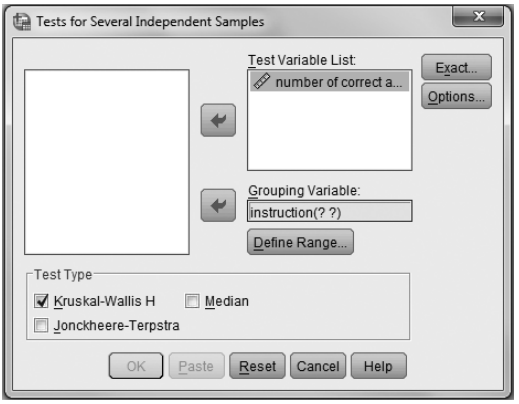
The data set has been saved under the name **EX19c.SAV**.

Variables	Column(s)	Code
• INSTRUCTION	1	1 = Method A, 2 = Method B, 3 = Method C
• SCORES	2	Number of correct responses

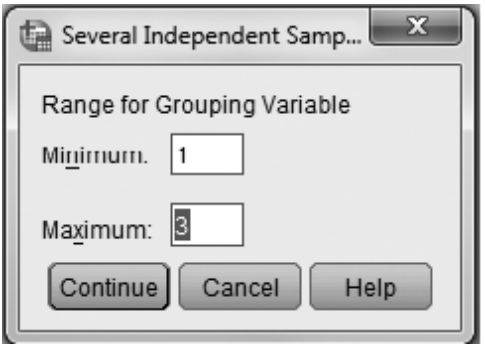
19.5.3 Windows Method

1. From the menu bar, click **Analyze**, then **Nonparametric Tests**, then **Legacy Dialogs**, and then **K Independent Samples**. The following **Tests for Several Independent Samples** window will open.
2. Since **INSTRUCTION** is the grouping (independent) variable, transfer it to the **Grouping Variable:** field by clicking (highlight) the variable and then clicking  . As **SCORES** is the test (dependent) variable, transfer it to the **Test Variable List:** field by clicking (highlight) the variable and then clicking  .

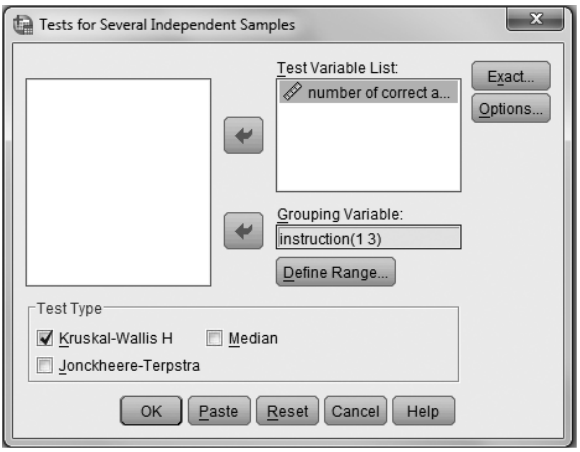




3. Click Define Range... to define the ranges for the grouping variable **INSTRUCTION** (coded 1 = Method A, 2 = Method B, 3 = Method C). When the following **Several Independent Samples** window opens, type **1** in the **Minimum:** field and **3** in the **Maximum:** field, and then click Continue.



4. When the following **Tests for Several Independent Samples** window opens, ensure that the **Kruskal-Wallis H** cell is checked. Run the analysis by clicking OK. See Table 19.5 for the results.



19.5.4 SPSS Syntax Method

```
NPAR TESTS K-W = SCORES BY INSTRUCTION(1,3).
```

19.5.5 SPSS Output

TABLE 19.5

Kruskal-Wallis Test Output

Kruskal-Wallis Test

Ranks

	Instruction Method	*N*	Mean Rank
Number of correct answers	A	9	12.72
	B	9	20.39
	C	9	8.89
	Total	27	

Test Statistics[a,b]

	Number of Correct Answers
Chi-square	9.896
d*f*	2
Asymp. sig.	.007

[a] Kruskal-Wallis Test.

[b] Grouping Variable: instruction method.

19.5.6 Results and Interpretation

The null hypothesis tested by the Kruskal-Wallis analysis is that the three instruction methods have the same effect on the number of correct responses obtained. Therefore, the samples are random samples from the same or identical population distributions. The obtained Kruskal-Wallis statistic is interpreted as a chi-square value, and is shown to be significant, $\chi^2(df = 2) = 9.89$, $p < .01$. Therefore, it can be concluded that the three instruction methods are not equally effective with respect to the number of correct answers obtained.

19.6 Wilcoxon Signed Rank Test for Two Related Samples

The Wilcoxon signed rank test is used for within-subjects design with data that are at least ordinal in scaling. When a researcher wants to analyze two sets of data obtained from the same individuals, the appropriate test to apply is the related *t* test (covered in Chapter 5). However, when there is an extreme violation of the normality assumption or when the data are not of appropriate scaling, the Wilcoxon signed rank test can be used.

Suppose a researcher is interested in whether a drug claimed by its manufacturer to improve problem-solving skills accomplished that goal. The researcher selects a sample of seven individuals and measured their problem-solving skills before and after they have taken that drug. The number of correct responses is recorded below.

	Correct Responses	
	Before Drug	**After Drug**
s1	69	71
s2	76	75
s3	80	84
s4	74	83
s5	87	90
s6	78	75
s7	82	81

19.6.1 Assumptions

- The scale of measurement within each pair must be at least ordinal in nature.
- The differences in scores must also constitute an ordinal scale.

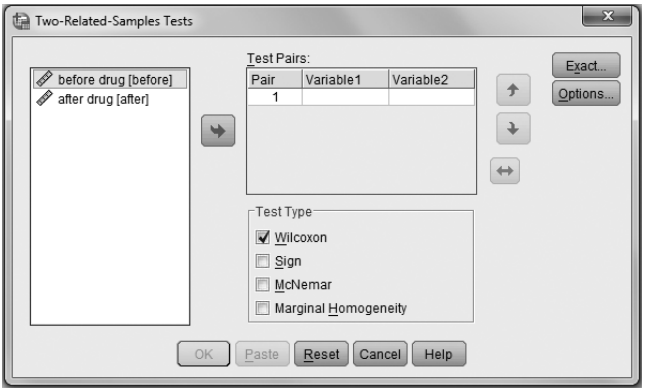
19.6.2 Data Entry Format

The data set has been saved under the name **EX19d.SAV**.

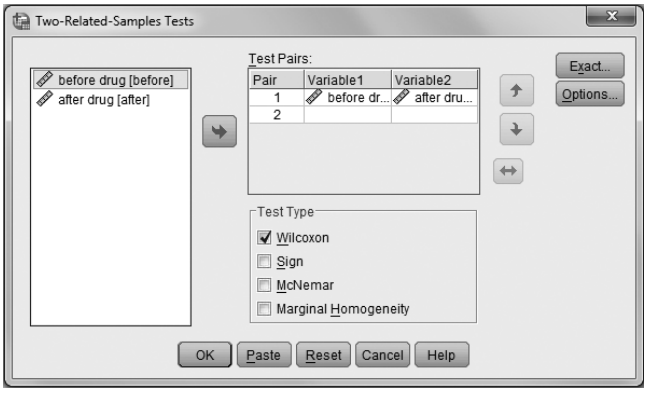
Variables	Column(s)	Code
• BEFORE	1	Number of correct responses
• AFTER	2	Number of correct responses

19.6.3 Windows Method

1. From the menu bar, click **Analyze**, then **Nonparametric Tests**, then **Legacy Dialogs**, and then **2 Related Samples**. The following **Two-Related-Samples Tests** window will open.



2. Transfer the two variables of **BEFORE** and **AFTER** to the **Variable1** and **Variable2** fields under **Test Pairs:** by clicking (highlight) them and then clicking . Ensure that the **Wilcoxon** cell is checked. Run the analysis by clicking . See Table 19.6 for the results.



19.6.4 SPSS Syntax Method

```
NPAR TESTS WILCOXON = BEFORE WITH AFTER (PAIRED).
```

19.6.5 SPSS Output

TABLE 19.6

Wilcoxon Signed Rank Test Output

Data Were Compared Using the Wilcoxon Signed Rank Test

Ranks

		N	Mean Rank	Sum of Ranks
After drug - before drug	Negative ranks	3[a]	2.50	7.50
	Positive ranks	4[b]	5.13	20.50
	Ties	0[c]		
	Total	7		

[a] after drug < before drug.
[b] after drug > before drug.
[c] after drug = before drug

Test Statistics[b]

	After drug - before drug
Z	−1.103[a].
Asymp. Sig. (2-tailed)	270

[a] Based on negative ranks.
[b] Wilcoxon Signed Rank Test.

19.6.6 Results and Interpretation

The null hypothesis tested is that the drug does not improve problem-solving skills, that is, it has no effect. In computing the test results, SPSS converts the Wicoxon statistic to a *z*-score that can be tested for significance under the normal curve. When using the 0.05 level of significance and a two-tailed test, the critical values of *z* are –1.96 and +1.96. Since the obtained value of *z* (–1.10) does not exceed these values, the researcher cannot reject the null hypothesis. That is, it is concluded that the drug has no effect.

19.7 Friedman Test for Several Related Samples

While the Wilcoxon test is used to analyze two sets of scores obtained from the same individuals, the **Friedman test** is used when there are more than two sets of scores. The Friedman test is the nonparametric alternative to a

one-way repeated measures analysis of variance. Like the Mann-Whitney and Kruskal-Wallis tests, the calculation of the Friedman test is based on ranks within each case. The scores for each variable are ranked and the mean ranks for the variables are compared.

Suppose a researcher is interested in whether problem-solving ability changes depending on the time of day. The sample consists of 11 individuals tested in the morning, afternoon, and evening. The scores recorded below are the number of correct responses achieved by each subject across the three time periods.

Problem-Solving (Number of Correct Responses)			
	Morning	**Afternoon**	**Evening**
s1	10	3	2
s2	6	5	15
s3	9	16	4
s4	11	12	5
s5	7	12	5
s6	12	6	11
s7	13	12	9
s8	12	3	2
s9	15	14	12
s10	13	12	5
s11	15	4	3

19.7.1 Assumption

- The scale of measurement within the variables must be at least ordinal in nature.
- The subjects represent a **random sample of subjects**.

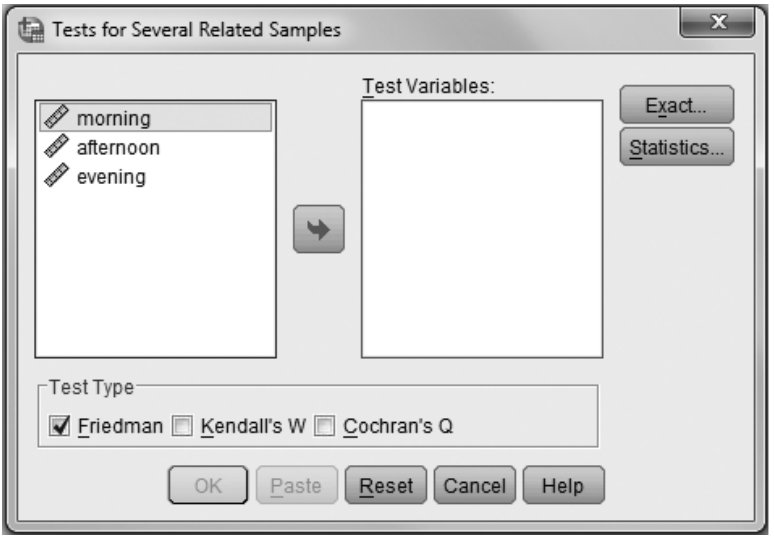
19.7.2 Data Entry Format

The data set has been saved under the name **EX19e.SAV**.

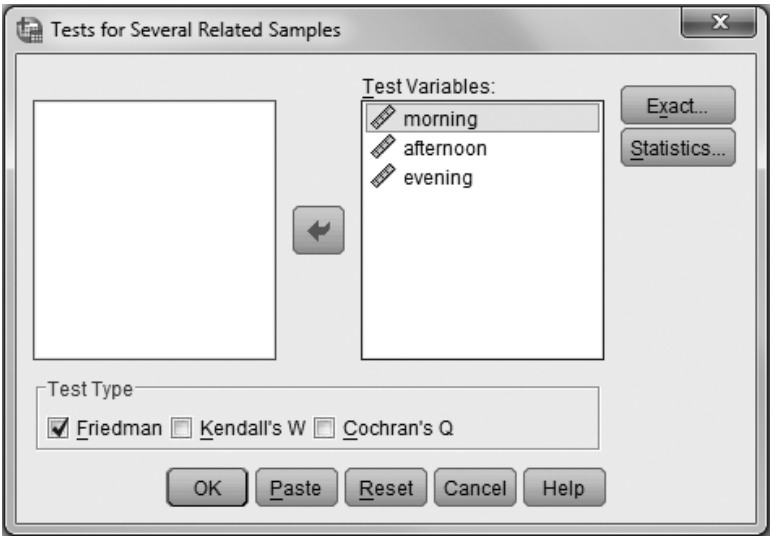
Variables	Column(s)	Code
• MORNING	1	Number of correct responses
• AFTERNOON	2	Number of correct responses
• EVENING	3	Number of correct responses

19.7.3 Windows Method

1. From the menu bar, click **Analyze**, then **Nonparametric Tests**, then **Legacy Dialogs**, and then **K Related Samples.** The following **Tests for Several Related Samples** window will open.



2. Transfer the three variables **MORNING, AFTERNOON,** and **EVENING** to the **Test Variables:** field by clicking (highlight) them and then clicking ▶. Ensure that the **Friedman** cell is checked. Run the analysis by clicking OK. See Table 19.7 for the results.



19.7.4 SPSS Syntax Method

```
NPAR TESTS FRIEDMAN = MORNING AFTERNOON EVENING.
```

19.7.5 SPSS Output

TABLE 19.7

Friedman Test Output

Friedman Test	
Ranks	
	Mean Rank
Morning	2.64
Afternoon	2.09
Evening	1.27

Test Statistics[a]	
N	11
Chi-square	10.364
d*f*	2
Asymp. sig.	.006

[a] Friedman Test.

19.7.6 Results and Interpretation

The Friedman test tests the null hypothesis that time of day makes no difference to the subjects' problem-solving ability (number of correct responses obtained). The results indicated that the Friedman χ^2 statistic is significant, $\chi^2(df = 2) = 10.36$, $p < .01$. Therefore, it can be concluded that time of day did have a significant effect on subjects' problem-solving ability.

Appendix: Summary of SPSS Syntax Files

Frequency Analysis (Chapter 2: TRIAL.SAV)

```
FREQUENCIES VARIABLES = ALL or list of variables
/STATISTICS = ALL.
```

Multiple Response Analysis (Chapter 3)

- Multiple-Dichotomy Frequency Analysis (EX3.SAV)

```
MULT RESPONSE GROUPS = REASONS 'REASONS FOR PREFERRING
  THAT PARTY' (HONEST TO LEADER(1))
/FREQUENCIES = REASONS.
```

- Multiple-Response Frequency Analysis (EX3.SAV)

```
MULT RESPONSE GROUPS = REASONS 'REASONS FOR PREFERRING
  THAT PARTY' (REASON1 TO REASON3 (1,5))
/FREQUENCIES = REASONS.
```

- Multiple-Response Cross-Tabulation Frequency Analysis (EX3.SAV)

```
MULT RESPONSE GROUPS = REASONS 'REASONS FOR PREFERRING
  THAT PARTY' (HONEST TO LEADER(1))
/VARIABLES SEX(1,2)
/TABLES = REASONS BY SEX
/CELLS = ALL.
```

t Test (Independent) (Chapter 4: EX4.SAV)

- Test of Normality Assumption

```
EXAMINE VARIABLES = WORDS
/PLOT NPPLOT
/COMPARE GROUPS
/STATISTICS DESCRIPTIVES
/CINTERVAL 95
```

```
/MISSING LISTWISE
/NOTOTAL.
```

- *t* Test for Independent Groups

```
T-TEST GROUPS = GENDER(1 2)
/MISSING = ANALYSIS
/VARIABLES = WORDS
/CRITERIA = CI(.95).
```

t Test (Related) (Chapter 5: EX5.SAV)

- Test of Normality Assumption

```
EXAMINE VARIABLES = BEFORE AFTER
/PLOT NPPLOT
/STATISTICS DESCRIPTIVES
/CINTERVAL 95
/MISSING LISTWISE
/NOTOTAL.
```

- *t* Test for Paired Samples

```
T-TEST PAIRS = BEFORE AFTER.
```

One-Way ANOVA (Chapter 6: EX6.SAV)

```
ONEWAY TIME BY SHOCK
/STATISTICS DESCRIPTIVES HOMOGENEITY
/MISSING ANALYSIS
/POSTHOC = SCHEFFE ALPHA(0.05).
```

Factorial ANOVA (Chapter 7)

- 2 × 2 ANOVA (EX7a.SAV)

```
UNIANOVA ERRORS BY STRATEGY LIST
/METHOD = SSTYPE(3)
/INTERCEPT = INCLUDE
/PLOT = PROFILE(STRATEGY*LIST)
```

```
/EMMEANS = TABLES(STRATEGY)
/EMMEANS = TABLES(LIST)
/EMMEANS = TABLES(STRATEGY*LIST)
/PRINT = HOMOGENEITY
/CRITERIA = ALPHA(.05)
/DESIGN = STRATEGY LIST STRATEGY*LIST.
```

- Data Transformation (Post Hoc Analysis)

```
IF (STRATEGY EQ 1 AND LIST EQ 1) GROUP = 1.
IF (STRATEGY EQ 1 AND LIST EQ 2) GROUP = 2.
IF (STRATEGY EQ 2 AND LIST EQ 1) GROUP = 3.
IF (STRATEGY EQ 2 AND LIST EQ 2) GROUP = 4.

VALUE LABELS GROUP 1 'STRATEGY A-EASY LIST'
                   2 'STRATEGY A-HARD LIST'
                   3 `STRATEGY B-EASY LIST'
                   4 `STRATEGY B-HARD LIST'.

ONEWAY ERRORS BY GROUP
/STATISTICS DESCRIPTIVES
/MISSING ANALYSIS
/POSTHOC = SCHEFFE ALPHA(0.05).
```

- $2 \times 2 \times 2$ ANOVA (EX7b.SAV)

```
GLM ERRORS BY STRATEGY LIST SHOCK
/PLOT = PROFILE(STRATEGY*LIST)
/PLOT = PROFILE(STRATEGY*SHOCK)
/PLOT = PROFILE(LIST*SHOCK)
/EMMEANS = TABLES(STRATEGY)
/EMMEANS = TABLES(LIST)
/EMMEANS = TABLES(SHOCK)
/EMMEANS = TABLES(STRATEGY*LIST)
/EMMEANS = TABLES(STRATEGY*SHOCK)
/EMMEANS = TABLES(LIST*SHOCK)
/EMMEANS = TABLES(STRATEGY*LIST*SHOCK).

IF (STRATEGY EQ 1 AND LIST EQ 1) GROUP = 1.
IF (STRATEGY EQ 1 AND LIST EQ 2) GROUP = 2.
IF (STRATEGY EQ 2 AND LIST EQ 1) GROUP = 3.
IF (STRATEGY EQ 2 AND LIST EQ 2) GROUP = 4.
VALUE LABELS GROUP 1 'STRATEGY A-EASY LIST'
                   2 'STRATEGY A-HARD LIST'
                   3 'STRATEGY B-EASY LIST'
                   4 'STRATEGY B-HARD LIST'.

UNIANOVA ERRORS BY SHOCK GROUP
/METHOD = SSTYPE(3)
/INTERCEPT = INCLUDE
```

```
/PLOT = PROFILE(GROUP*SHOCK)
/EMMEANS = TABLES(SHOCK)
/EMMEANS = TABLES(GROUP)
/EMMEANS = TABLES(SHOCK*GROUP)
/CRITERIA = ALPHA(.05)
/DESIGN = GROUP SHOCK SHOCK*GROUP.
```

GLM Multivariate ANOVA (Chapter 8)

- One-Sample Test (EX8a.SAV)

```
COMPUTE FIFTY = (T1-6).
COMPUTE ONEHUND = (T2-12).
COMPUTE TWOHUND = (T3-25).
COMPUTE THREEHUN = (T4-40).

GLM FIFTY ONEHUND TWOHUND THREEHUN
/PRINT = DESCRIPTIVES.
```

- Two-Sample Test (EX8b.SAV)

```
COMPUTE FIFTY = (T1-6).
COMPUTE ONEHUND = (T2-12).
COMPUTE TWOHUND = (T3-25).
COMPUTE THREEHUN = (T4-40).

GLM FIFTY ONEHUND TWOHUND THREEHUN BY SEX
/EMMEANS = TABLES(SEX)
/PRINT HOMOGENEITY.
```

- 2×2 Factorial Design (EX8c.SAV)

```
COMPUTE FIFTY = (T1-6).
COMPUTE ONEHUND = (T2-12).
COMPUTE TWOHUND = (T3-25).
COMPUTE THREEHUN = (T4-40).

GLM FIFTY ONEHUND TWOHUND THREEHUN BY SEX ETHNIC
/PRINT HOMOGENEITY
/EMMEANS = TABLES(SEX)
/EMMEANS = TABLES(ETHNIC)
/EMMEANS = TABLES(SEX*ETHNIC).

GRAPH
/LINE(MULTIPLE)MEAN(FIFTY) BY SEX BY ETHNIC.
```

```
GRAPH
/LINE(MULTIPLE)MEAN(ONEHUND) BY SEX BY ETHNIC.
GRAPH
/LINE(MULTIPLE)MEAN(TWOHUND) BY SEX BY ETHNIC.
GRAPH
/LINE(MULTIPLE)MEAN(THREEHUN) BY SEX BY ETHNIC.

IF (SEX EQ 1 AND ETHNIC EQ 1) GROUP = 1.
IF (SEX EQ 1 AND ETHNIC EQ 2) GROUP = 2.
IF (SEX EQ 2 AND ETHNIC EQ 1) GROUP = 3.
IF (SEX EQ 2 AND ETHNIC EQ 2) GROUP = 4.
VALUE LABELS GROUP 1 'MALE-WHITE' 2 'MALE-NONWHITE'
 3 'FEMALE-WHITE' 4 'FEMALE-NONWHITE'.

ONEWAY ONEHUND BY GROUP (1,4)
/RANGES = SCHEFFE(.05).
```

GLM Repeated Measures Analysis (Chapter 9)

- One-Way Repeated Measures (EX9a.SAV)

```
GLM TEMP1 TO TEMP4
/WSFACTOR = TEMP 4 REPEATED
/MEASURE = ERRORS
/EMMEANS = TABLES(TEMP) COMPARE ADJ(BONFERRONI).
```

- Doubly Multivariate (EX9b.SAV)

```
GLM LS_E TO HS_D
/WSFACTOR = SHOCK 2 REPEATED ITEMS 3 REPEATED
/MEASURE = CORRECT
/PLOT = PROFILE(ITEMS*SHOCK)
/EMMEANS = TABLES(SHOCK) COMPARE ADJ(BONFERRONI)
/EMMEANS = TABLES(ITEMS) COMPARE ADJ(BONFERRONI)
/EMMEANS = TABLES(SHOCK*ITEMS).
```

- Two-Factor Mixed Design (EX9c.SAV)

```
GLM TRIAL1 TO TRIAL3 BY GROUP
/WSFACTOR = TRIAL 3 REPEATED
/MEASURE = CORRECT
/PLOT = PROFILE(TRIAL*GROUP)
/POSTHOC = GROUP(SCHEFFE)
/EMMEANS = TABLES(GROUP)
/EMMEANS = TABLES(TRIAL) COMPARE ADJ(BONFERRONI)
/EMMEANS = TABLES(GROUP*TRIAL).
```

- Three-Factor Mixed Design (EX9d.SAV)

```
GLM TRIAL1 TO TRIAL4 BY DRUG SEX
/WSFACTOR = TRIAL 4 REPEATED
/MEASURE = ERRORS
/EMMEANS = TABLES(DRUG)
/EMMEANS = TABLES(SEX)
/EMMEANS = TABLES(DRUG*SEX)
/EMMEANS = TABLES(TRIAL) COMPARE ADJ(BONFERRONI)
/EMMEANS = TABLES(DRUG*TRIAL)
/EMMEANS = TABLES(SEX*TRIAL)
/EMMEANS = TABLES(DRUG*SEX*TRIAL)
/PLOT = PROFILE(DRUG*SEX)
/PLOT = PROFILE(TRIAL*DRUG)
/PLOT = PROFILE(TRIAL*SEX).

IF (DRUG EQ 1 AND SEX EQ 1) GROUP = 1.
IF (DRUG EQ 2 AND SEX EQ 1) GROUP = 2.
IF (DRUG EQ 1 AND SEX EQ 2) GROUP = 3.
IF (DRUG EQ 2 AND SEX EQ 2) GROUP = 4.
VALUE LABELS GROUP 1 'DRUG PRESENT-MALE'
                   2 'DRUG ABSENT-MALE'
                   3 'DRUG PRESENT-FEMALE'
                   4 'DRUG ABSENT-FEMALE'.

GLM TRIAL1 TO TRIAL4 BY GROUP
/WSFACTOR = TRIAL 4 REPEATED
/MEASURE = ERRORS
/PLOT = PROFILE(TRIAL*GROUP)
/EMMEANS = TABLES(GROUP*TRIAL).
```

Correlation Analysis (Chapter 10)

- Scatterplot (CORR.SAV)

```
GRAPH
/SCATTERPLOT(BIVAR) = READ WITH GPA
/MISSING = LISTWISE.
```

- Pearson Product Moment Correlation (CORR.SAV)

```
CORRELATIONS READ WITH GPA
/MISSING = PAIRWISE.
```

- Spearman Rank Order Correlation Coefficient (CORR.SAV)

```
NONPAR CORR READ_RANK WITH GPA_RANK.
```

Linear Regression (Chapter 11: CORR.SAV)

```
REGRESSION VARIABLES = (COLLECT)
/MISSING LISTWISE
/STATISTICS = DEFAULTS CI
/DEPENDENT = GPA
/METHOD = ENTER READ.
```

Factor Analysis (Chapter 12: EUTHAN_1.SAV)

```
FACTOR VARIABLES = E1 to E12
/FORMAT = SORT BLANK(.33)
/PRINT = INITIAL EXTRACTION ROTATION CORRELATION KMO
/PLOT = EIGEN
/EXTRACTION = PC
/ROTATION = OBLIMIN.
```

- Specifying Number of Factors (3)

```
FACTOR VARIABLES = E1 to E12
/FORMAT = SORT BLANK(.33)
/CRITERIA = FACTOR(2)
/EXTRACTION = PC
/ROTATION = OBLIMIN.
```

Reliability (Chapter 13: EUTHAN_1.SAV)

```
RELIABILITY VARIABLES = E1 TO E12
/SCALE(VOLUNTARY_EUTHANASIA) = E1 E2 E5 E6 E9 E10
/STATISTICS = SCALE
/SUMMARY = TOTAL.
```

Multiple Regression (Chapter 14)

- Forward Selection (DOMES.SAV)

```
COMPUTE PROVOKE = MEAN(PROVO,PASSION).
COMPUTE SELFDEF = MEAN(PROTECT,SAVE,DEFEND).
COMPUTE INSANITY = MEAN(MENTAL,INSANE,STABLE).
```

```
REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS CI(95) R ANOVA COLLIN TOL CHANGE
/CRITERIA = PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT RESPON
/METHOD = FORWARD PROVOKE SELFDEF INSANITY
/SCATTERPLOT = (*ZRESID,*ZPRED)
/RESIDUALS NORMPROB(ZRESID).
```

- Hierarchical Regression (DOMES.SAV)

```
REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS CI(95) R ANOVA COLLIN TOL CHANGE
/CRITERIA = PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT RESPON
/METHOD = ENTER SEX AGE EDUC INCOME/ENTER PROVOKE
INSANITY SELFDEF.
```

- Path Analysis (DOMES.SAV)

```
REGRESSION VARIABLES = (COLLECT)
    /STATISTICS = DEFAULTS CHA TOL CI COLLIN
    /DEPENDENT = RESPON
    /FORWARD PROVOKE SELFDEF INSANITY.
REGRESSION VARIABLES = (COLLECT)
    /STATISTICS = DEFAULTS CHA TOL CI COLLIN
    /DEPENDENT = SELFDEF
    /FORWARD PROVOKE.
REGRESSION VARIABLES = (COLLECT)
    /STATISTICS = DEFAULTS CHA TOL CI COLLIN
    /DEPENDENT = INSANITY
    /FORWARD PROVOKE.
```

Multiple Discriminant Analysis (Chapter 15)

- Test for Multivariate Outliers (Mahalanobis Distance) (DRIVE_1.SAV)

```
REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA
/CRITERIA = PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT GROUP
```

```
/METHOD = ENTER FAST_RIS DISRESP SEN_SEEK DANGER
/SAVE MAHAL.
```

- Test for Multicollinearity (DRIVE_1.SAV)

```
REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA TOL
/CRITERIA = PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT FAST_RIS
/METHOD = ENTER DISRESP SEN_SEEK DANGER.
REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA TOL
/CRITERIA = PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT DISRESP
/METHOD = ENTER FAST_RIS SEN_SEEK DANGER.
REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA TOL
/CRITERIA = PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT SEN_SEEK
/METHOD = ENTER DISRESP DANGER FAST_RIS.
REGRESSION
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA TOL
/CRITERIA = PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT DANGER
/METHOD = ENTER DISRESP FAST_RIS SEN_SEEK.
```

- Two-Group Discriminant Analysis (DRIVE_1.SAV)

```
DISCRIMINANT
/GROUPS = GROUP(1 2)
/VARIABLES = FAST_RIS DISRESP SEN_SEEK DANGER
/ANALYSIS ALL
/PRIORS SIZE
/STATISTICS = MEAN STDDEV UNIVF BOXM RAW CORR TABLE
/CLASSIFY = NONMISSING POOLED.
```

- Three-Group Discriminant Analysis (DRIVE_1.SAV)

```
DISCRIMINANT
/GROUPS = GROUP_1(1 3)
/VARIABLES = FAST_RIS DISRESP SEN_SEEK DANGER
```

```
/ANALYSIS ALL
/SAVE = SCORES
/PRIORS SIZE
/STATISTICS = MEAN STDDEV UNIVF BOXM COEFF RAW CORR
     TABLE
/PLOT = COMBINED
/CLASSIFY = NONMISSING POOLED.
```

Logistic Regression (Chapter 16) (DRIVE_1.SAV)

```
LOGISTIC REGRESSION VARIABLES GROUP
/METHOD = ENTER FAST_RIS DISRESP SEN_SEEK DANGER
/PRINT = GOODFIT CI(95)
/CRITERIA = PIN(0.05) POUT(0.10) ITERATE(20) CUT(0.5).
```

Canonical Correlation Analysis (Chapter 17) (HIV.SAV)

```
MANOVA WILLING_USE LIKELY_USE INTEND_USE CERTAIN_USE
WITH PERCEIVED_RISK PERCEIVED_SEVERITY SELF_EFFICACY
RESPONSE_EFFICACY MALADAPTIVE_COPING FEAR
/PRINT = SIGNIF(MULT EIGEN DIMENR)
/DISCRIM = (STAN ESTIM COR ALPHA(.999)).
```

Chi-Square (χ^2) Test for Single Variable Experiment—Equal Expected Frequencies (Chapter 19: EX19a.SAV)

```
NPAR TESTS CHISQUARE = COLA
/EXPECTED = EQUAL.
```

Chi-Square (χ^2) Test for Single Variable Experiment—Unequal Expected Frequencies (Chapter 19: EX19a.SAV)

```
NPAR TESTS CHISQUARE = COLA
/EXPECTED = 9 5 28.
```

Chi-Square (χ^2) Test of Independence between Two Variables (Chapter 19: EX19a.SAV)

```
CROSSTABS TABLES = COLA BY SEX
/CELLS = COUNT ROW COLUMN TOTAL EXPECTED
/STATISTICS = CHISQ.
```

Mann-Whitney *U* Test for Two Independent Samples (Chapter 19: EX19b.SAV)

```
NPAR TESTS M-W = INCOME BY SCHOOL(1,2).
```

Kruskal-Wallis Test for Several Independent Samples (Chapter 19: EX19c.SAV)

```
NPAR TESTS K-W = SCORES BY INSTRUCTION(1,3).
```

Wilcoxon Signed Rank Test for Two Related Samples (Chapter 19: EX19d.SAV)

```
NPAR TESTS WILCOXON = BEFORE WITH AFTER (PAIRED).
```

Friedman Test for Several Related Samples (Chapter 19: EX19e.SAV)

```
NPAR TESTS FRIEDMAN = MORNING AFTERNOON EVENING.
```

Missing Value Analysis

```
MVA VARIABLES = V1 TO Vn
/EM (OUTFILE = 'C:\emdata.sav').
```

Bibliography

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov and F. Csaki (Eds.), *Proceedings of the 2nd international symposium on information theory* (pp. 267–281). Budapest: Akademiai Kiado.

Akaike, H. (1987). Factor analysis and AIC. *Psychometrika*, *52*, 317–332.

Anderson, J. C., and Gerbing, D. W. (1988). Structural equation modeling in practice. A review and recommended two-step approach. *Psychological Bulletin*, *103*, 411–423.

Arbuckle, J. L., and Wothke, W. (1999). *AMOS 4.0 User's guide*. Chicago, IL: SmallWaters Corporation.

Bentler, P. M. (1980). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*, 238–246.

Bentler, P. M., and Bonnett, D. G. (1980). Significant tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*, 588–606.

Bentler, P. M., and Mooijaart, A. (1989). Choice of structural model via parsimony: A rationale based on precision. *Psychological Bulletin*, *106*, 315–317.

BMPD Statistical Software, Inc. (1992). *BMPD statistical software manual*, Release 7 (Vols. 1 and 2), Los Angeles, CA: BMPD Statistical Software, Inc.

Bollen, K. A., and Stine, R. (1990). Direct and indirect effects: Classical and bootstrap estimates of variability. *Sociological Methodology*, *20*, 115–140.

Browne, M. W., and Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen and J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

Byrne, B. M. (2001). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. Mahwah, NJ: Lawrence Earlbaum Associates, Inc.

Cattel, R. B. (1966, April). The scree test for the number of factors. *Multivariate Behavioral Research*, *1*, 245–276.

Chou, C.-P., and Bentler, P. M. (1993). Invariant standardized estimated parameter change for model modification in covariance structural analysis. *Multivariate Behavioral Research*, *28*, 97–110.

Cohen, J., and Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Crawford, C. B., and Koopman, P. (1979). Inter-rater reliability of scree test and mean square ratio test of number of factors. *Perceptual and Motor Skills*, *49*, 223–226.

Douglas, S., and Love, W. (1968). A general canonical correlation index. *Psychological Bulletin*, *70*, 160–163.

Edwards, A. L. (1967). *Statistical methods* (2nd ed.). New York, NY: Holt, Rinehart & Winston.

Efron, B., and Tibshirani, R. J. (1993). *An introduction to the bootstrap*. Boca Raton, FL: Chapman & Hall.

Gerbing, D. W., and Anderson, J. C. (1984). On the meaning of within-factor correlated measurement errors. *Journal of Consumer Research*, *11*, 572–580.

Hair, J. F., Anderson, R. E., Tatham, R. L., and Black, W. C. (1995). *Multivariate data analysis with readings* (4th ed.). Englewood Cliffs, NJ: Prentice-Hall.

Hair, J. F., Anderson, R. E., Tatham, R. L., and Black, W. C. (1998). *Multivariate data analysis* (5th ed.). Upper Saddle River, NJ: Prentice-Hall.

Ho, R. (1989). Why do people smoke?: Motives for the maintenance of smoking behaviour and its possible cessation. *Australian Psychologist*, *24*(3), 385–400.

Ho, R. (1998). Assessing attitudes toward euthanasia: An analysis of the subcategorical approach to right to die issues. *Personality and Individual Differences*, *25*, 719–734.

Ho, R. (1999). Factors influencing decisions to terminate life: Condition of suffering and the identity of the terminally ill. *Australian Journal of Social Issues*, *34*, 25–41.

Ho, R. (2000). Predicting intention for protective health behaviour: A test of the protection versus the ordered protection motivation model. *Australian Journal of Psychology*, *52*(2), 110–118.

Ho, R., and Venus, M. (1995). Reactions to a battered woman who kills her abusive spouse: An attributional analysis. *Australian Journal of Psychology*, *47*(3), 153–159.

Ho, R., and Yong Gee, R. (2008). Young men driving dangerously: Development of the motives for dangerous driving scale (MDDS). *Australian Journal of Psychology*, *60*(2), 91–100.

Hoyle, R. H., and Smith, G. T. (1994). Formulating clinical research hypotheses as structural equation models: A conceptual overview. *Journal of Consulting and Clinical Psychology*, *62*(3), 429–440.

Hu, L. T., and Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*, 424–453.

Hu, L. T., and Bentler, P. M. (1999). Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55.

IBM Corp. Released. (2011). *IBM SPSS Statistics for Windows, Version 20.0*. Armonk, NY: IBM Corp. (Released 2011).

Jöreskog, K. G., and Sörbom, D. (1989). *LISREL 8: A guide to program and applications*. Chicago, IL: SPSS Inc.

Jöreskog, K. G., and Sörbom, D. (1993). *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Kline, R. B. (2005). *Principles and practice of structural equation modeling*. New York, NY: The Guilford Press.

MacCullum, R. C. (1986). Specification searches in covariance structure modeling. *Psychological Bulletin*, *100*, 107–120.

MacCallum, R. C., Browne, M. W., and Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, *1*, 130–149.

Marsh, H. W., Hau, K.-T., and Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, *11*(3), 320–341.

McCall, R. B. (1990). *Fundamental statistics for the behavioral sciences* (5th ed.). New York, NY: Harcourt Brace Jovanovich.

McDonald, R. P., and Marsh, H. W. (1990). Choosing a multivariate model: Noncentrality and goodness-of-fit. *Psychological Bulletin*, *107*, 247–255.

Muthuen, B., and Kaplan, D. (1985). A comparison of methodologies for the factor analysis of nonnormal Likert variables. *British Journal of Mathematical and Statistical Psychology*, *38*, 171–189.

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers*, *32*(3), 396–402.

Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction* (3rd ed.). New York, NY: Harcourt Brace College Publishers.

Preacher, K. J., and Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, *36*(4), 717–731.

Preacher, K. J., and Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*(3), 879–891.

Rindskopf, D., and Rose, T. (1988). Some theory and applications of confirmatory second-order factor analysis. *Multivariate Behavioral Research*, *23*, 51–67.

Russell, D. W., Kahn, J. H., Spoth, R., and Altmaier, E. M. (1998). Analyzing data from experimental studies: A latent variable structural equation modelling approach. *Journal of Counseling Psychology*, *45*(1), 18–29.

Sherry, A., and Henson, R. K. (2005). Conducting and interpreting canonical correlation analysis in personality research: A user-friendly primer. *Journal of Personality Assessment*, *84*(1), 37–48.

SPSS Inc. (1998). *Statistical package for the social sciences*. Chicago, IL: SPSS Inc.

Streiner, D. L. (1998). Factors affecting reliability of interpretations of scree plots. *Psychological Reports*, *83*, 687–694.

Sweet, S. A. and Grace-Martin, K. (2002). *Data analysis with SPSS: A first course in applied statistics*. Needham Heights, MA: Allyn & Bacon.

Tabachnick, B. G., and Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Boston, MA: Allyn and Bacon.

Thompson, B. (1997). Editorial policies regarding statistical significance tests: Further comments. *Educational Researcher*, *25*(2), 26–30.

Wang, L. L., Fan, X., and Wilson, V. L. (1996). Effects of nonnormal data on parameter estimates for a model with latent and manifest variables: An empirical study. *Structural Equation Modeling*, *3*(3), 228–247.

Williams, L. J., and Holahan, P. J. (1994). Parsimony-based fit indices for multiple-indicator models. *Structural Equation Modeling*, *1*(2), 161–189.

Wood, J. M., Tataryn, D. J., and Gorsuch, R. L. (1996). Effects of under- and overextraction on principal axis factor analysis with varimax solution. *Psychological Methods*, *1*, 354–365.

Wothke, W. (1993). Nonpositive definite matrices in structural modeling. In K. A. Bollen and J. S. Long (Eds.), *Testing structural equation models* (pp. 256–293). Newbury Park, California: Sage.

Zwick, W. R., and Velicer, W. F. (1982). Factors influencing four rules for determining the number of components to retain. *Multivariate Behavioral Research*, *17*, 253–269.

Zwick, W. R., and Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, *99*, 432–442.

Statistics

Using the same accessible, hands-on approach as its best-selling predecessor, the **Handbook of Univariate and Multivariate Data Analysis with IBM SPSS, Second Edition** explains how to apply statistical tests to experimental findings, identify the assumptions underlying the tests, and interpret the findings.

New to the Second Edition

- Three new chapters on multiple discriminant analysis, logistic regression, and canonical correlation
- New section on how to deal with missing data
- Coverage of tests of assumptions, such as linearity, outliers, normality, homogeneity of variance-covariance matrices, and multicollinearity
- Discussions of the calculation of Type I error and the procedure for testing statistical significance between two correlation coefficients obtained from two samples
- Expanded coverage of factor analysis, path analysis (test of the mediation hypothesis), and structural equation modeling

The handbook guides you in selecting the right statistical test, executing a wide range of univariate and multivariate statistical tests via the Windows and syntax methods, and interpreting the output results. SPSS syntax files are located in the appendix and data sets are available on the book's CRC Press web page.

About the Author

Robert Ho is an associate professor in the Graduate School of Psychology at the Assumption University of Thailand. He received a DPhil from Waikato University. His research interests include quantitative methods, health psychology, and social psychology.

K14105

ISBN: 978-1-4398-9021-9

90000

9 781439 890219

CRC Press
Taylor & Francis Group
an informa business
www.crcpress.com

6000 Broken Sound Parkway, NW
Suite 300, Boca Raton, FL 33487
711 Third Avenue
New York, NY 10017
2 Park Square, Milton Park
Abingdon, Oxon OX14 4RN, UK