

CD ENCLOSED,
appropriate for SPSS 12
and earlier versions

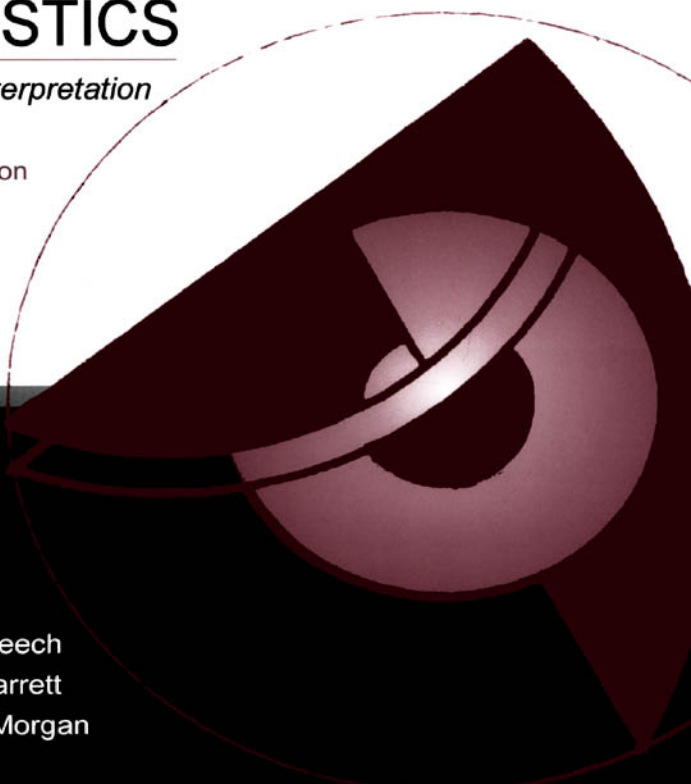
SPSS

FOR INTERMEDIATE STATISTICS

Use and Interpretation

Second Edition

Nancy L. Leech
Karen C. Barrett
George A. Morgan



SPSS for Intermediate Statistics: Use and Interpretation

Second Edition

This page intentionally left blank

SPSS for Intermediate Statistics: Use and Interpretation

Second Edition

Nancy L. Leech
University of Colorado at Denver

Karen C. Barrett
George A. Morgan
Colorado State University

In collaboration with
Joan Naden Clay
Don Quick



LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS
Mahwah, New Jersey **London**

Camera ready copy for this book was provided by the author.

Copyright © 2005 by Lawrence Erlbaum Associates, Inc.

All rights reserved. No part of this book may be reproduced in any form, by photostat, microform, retrieval system, or any other means, without prior written permission of the publisher.

Lawrence Erlbaum Associates, Inc., Publishers
10 Industrial Avenue
Mahwah, New Jersey 07430

Cover design by Kathryn Houghtaling Lacey

CIP information can be obtained by contacting the Library of Congress.

ISBN 0-8058-4790-1 (pbk. : alk. paper)

Books published by Lawrence Erlbaum Associates are printed on acid-free paper, and their bindings are chosen for strength and durability.

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

Disclaimer:

This eBook does not include the ancillary media that was packaged with the original printed version of the book.

Table of Contents

Preface	vii
1. Introduction and Review of Basic Statistics With SPSS.....	1
Variables	
Research Hypotheses and Research Questions	
A Sample Research Problem: The Modified High School and Beyond (HSB) Study	
Research Questions for the Modified HSB Study	
Frequency Distributions	
Levels of Measurement	
Descriptive Statistics	
Conclusions About Measurement and the Use of Statistics	
The Normal Curve	
Interpretation Questions	
2. Data Coding and Exploratory Analysis (EDA).....	24
Rules for Data Coding	
Exploratory Data Analysis (EDA)	
Statistical Assumptions	
Checking for Errors and Assumptions With Ordinal and Scale Variables	
Using Tables and Figures for EDA	
Transforming Variables	
Interpretation Questions	
Extra Problems	
3. Selecting and Interpreting Inferential Statistics	46
Selection of Inferential Statistics	
The General Linear Model	
Interpreting the Results of a Statistical Test	
An Example of How to Select and Interpret Inferential Statistics	
Review of Writing About Your Outputs	
Conclusion	
Interpretation Questions	
4. Several Measures of Reliability	63
Problem 4.1: Cronbach's Alpha for the Motivation Scale	
Problems 4.2 & 4.3: Cronbach's Alpha for the Competence and Pleasure Scales	
Problem 4.4: Test-Retest Reliability Using Correlation	
Problem 4.5: Cohen's Kappa With Nominal Data	
Interpretation Questions	
Extra Problems	
5. Exploratory Factor Analysis and Principal Components Analysis	76
Problem 5.1: Factor Analysis on Math Attitude Variables	
Problem 5.2: Principal Components Analysis on Achievement Variables	
Interpretation Questions	
Extra Problems	

6. Multiple Regression	90
Problem 6.1: Using the Simultaneous Method to Compute Multiple Regression	
Problem 6.2: Simultaneous Regression Correcting Multicollinearity	
Problem 6.3: Hierarchical Multiple Linear Regression	
Interpretation Questions	
7. Logistic Regression and Discriminant Analysis.....	109
Problem 7.1: Logistic Regression	
Problem 7.2: Hierarchical Logistic Regression	
Problem 7.3: Discriminant Analysis (DA)	
Interpretation Questions	
8. Factorial ANOVA and ANCOVA	129
Problem 8.1: Factorial (2-Way) ANOVA	
Problem 8.2: Post Hoc Analysis of a Significant Interaction	
Problem 8.3: Analysis of Covariance (ANCOVA)	
Interpretation Questions	
Extra Problems	
9. Repeated Measures and Mixed ANOVAs	147
The Product Data Set	
Problem 9.1: Repeated Measures ANOVA	
Problem 9.2: The Friedman Nonparametric Test for Several Related Samples	
Problem 9.3: Mixed ANOVA	
Interpretation Questions	
10. Multivariate Analysis of Variance (MANOVA) and Canonical Correlation	162
Problem 10.1: GLM Single-Factor Multivariate Analysis of Variance	
Problem 10.2: GLM Two-Factor Multivariate Analysis of Variance	
Problem 10.3: Mixed MANOVA	
Problem 10.4: Canonical Correlation	
Interpretation Questions	

Appendices

A. Quick Reference Guide to SPSS Procedures	Joan Naden Clay	188
C. Getting Started with SPSS.....		206
D. Making Figures and Tables	Don Quick	213
E. Answers to Odd Numbered Interpretation Questions.....		226
For Further Reading.....		232
Index		233

PREFACE

This book is designed to help students learn how to analyze and interpret research data with intermediate statistics. It is intended to be a supplemental text in an intermediate statistics course in the behavioral sciences or education and it can be used in conjunction with any mainstream text. We have found that the book makes SPSS for windows easy to use so that it is not necessary to have a formal, instructional computer lab; you should be able to learn how to use SPSS on your own with this book. Access to the SPSS program and some familiarity with Windows is all that is required. Although SPSS for Windows is quite easy to use, there is such a wide variety of options and statistics that knowing which ones to use and how to interpret the printouts can be difficult, so this book is intended to help with these challenges.

SPSS 12 and Earlier Versions

We use SPSS 12 for Windows in this book, but, except for enhanced tables and graphics, there are only minor differences from versions 10 and 11. In fact, as far as the procedures demonstrated, in this book there are only a few major differences between versions 7 and 12. We also expect future Windows versions to be similar. You should not have much difficulty if you have access to SPSS versions 7 through 9. Our students have used this book, or earlier editions of it, with all of these versions of SPSS; both the procedures and outputs are quite similar.

Goals of This Book

This book demonstrates how to produce a variety of statistics that are usually included in intermediate statistics courses, plus some (e.g., reliability measures) that are useful for doing research. Our goal is to describe the use and interpretation of these statistics as much as possible in nontechnical, jargon-free language.

Helping you learn how to choose the appropriate statistics, interpret the outputs, and develop skills in writing about the meaning of the results are the main goals of this book.

Thus, we have included material on:

- 1) How the appropriate **choice of a statistic** is based on the design of the research.
- 2) How to use SPSS to **answer research questions**.
- 3) How to **interpret SPSS outputs**.
- 4) How to **write about the outputs** in the Results section of a paper.

This information will help you develop skills that cover a range of steps in the research process: design, data collection, data entry, data analysis, interpretation of outputs, and writing results. The modified high school and beyond data set (HSB) used in this book is similar to one you might have for a thesis, dissertation, or research project. Therefore, we think it can serve as a model for your analysis. The compact disk (CD) packaged with the book contains the HSB data file and several other data sets used for the extra problems at the end of each chapter. However, **you will need to have access to or purchase the SPSS program**. Partially to make the text more readable, we have chosen not to cite many references in the text; however, we have provided a short bibliography of some of the books and articles that we have found useful. We assume that most students will use this book in conjunction with a class that has a textbook; it will help you to read more about each statistic before doing the assignments. Our "For Further Reading" list should also help.

Our companion book, Morgan, Leech, Gloeckner, and Barrett (2004), *SPSS for Introductory Statistics: Use and Interpretation*, also published by Lawrence Erlbaum Associates, is on the "For Further Reading" list at the end of this book. We think that you will find it useful if you need to

Preface

review how to do introductory statistics including the ones such as *t* tests, chi-square, and correlation.

Special Features

Several user friendly features of this book include:

1. The **key SPSS windows** that you see when performing the statistical analyses. This has been helpful to “visual learners.”
2. The **complete outputs** for the analyses that we have done so you can see what you will get, after some editing in SPSS to make the outputs fit better on the pages.
3. **Callout boxes** on the outputs that point out parts of the output to focus on and indicate what they mean.
4. For each output, a boxed **interpretation section** that will help you understand the output.
5. Specially developed flow charts and tables to help you **select an appropriate inferential statistic** and tell you how to **interpret statistical significance and effect sizes** (in Chapter 3). This chapter also provides an extended example of how to identify and write a research problem, several research questions, and a results paragraph for a *t* test and correlation.
6. For the statistics in chapters 4-10, an example of **how to write about the output** and make a table for a thesis, dissertation or research paper.
7. **Interpretation questions** that stimulate you to think about the information in the chapter and outputs.
8. Several **extra SPSS problems** at the end of each chapter for you to run with SPSS and discuss.
9. A **Quick Reference Guide** to SPSS (Appendix A) which provides information about many SPSS commands not discussed in the chapters.
10. Information (in Appendix B) on how to **get started with SPSS**.
11. A step by step guide to (Appendix C) **making APA tables** with MsWord.
12. **Answers** to the odd numbered interpretation questions (Appendix D).
13. **Several data sets on a CD**. These realistic data sets are packaged with the book to provide you with data to be used to solve the chapter problems and the extra problems at the end of each chapter.

Overview of the Chapters

Our approach in this book is to present how to use and interpret SPSS in the context of proceeding as if the HSB data were the actual data from your research project. However, before starting the SPSS assignments, we have three introductory chapters. The first chapter is an introduction and review of research design and how it would apply to analyzing the HSB data. In addition chapter includes a review of measurement and descriptive statistics. Chapter 2 discusses rules for coding data, exploratory data analysis (EDA), and assumptions. Much of what is done in this chapter involves preliminary analyses to get ready to answer the research questions that you might state in a report.

Chapter 3 provides a brief overview of research designs (between groups and within subjects). This chapter provides flowcharts and tables useful for selecting an appropriate statistic. Also included is an overview of how to interpret and write about the results of a basic inferential statistic. This section includes not only testing for statistical significance but also a discussion of effect size measures and guidelines for interpreting them.

Chapters 4-10 are designed to answer several research questions. Solving the problems in these chapters should give you a good idea of some of the intermediate statistics that can be computed with SPSS. Hopefully, seeing how the research questions and design lead naturally to the choice

Preface

of statistics will become apparent after using this book. In addition, it is our hope that interpreting what you get back from the computer will become more clear after doing these assignments, studying the outputs, answering the interpretation questions, and doing the extra SPSS problems.

Our Approach to Research Questions, Measurement, and Selection of Statistics

In Chapters 1 and 3, our approach is somewhat nontraditional because we have found that students have a great deal of difficulty with some aspects of research and statistics but not others. Most can learn formulas and “crunch” the numbers quite easily and accurately with a calculator or with a computer. However, many have trouble knowing what statistics to use and how to interpret the results. They do not seem to have a “big picture” or see how research design and measurement influence data analysis. Part of the problem is inconsistent terminology. For these reasons, we have tried to present a semantically consistent and coherent picture of how research design leads to three basic kinds of research questions (difference, associational, and descriptive) which, in turn, lead to three kinds or groups of statistics with the same names. We realize that these and other attempts to develop and utilize a consistent framework are both nontraditional and somewhat of an oversimplification. However, we think the framework and consistency pay off in terms of student understanding and ability to actually use statistics to answer their research questions. Instructors who are not persuaded that this framework is useful can skip Chapters 1 and 3 and still have a book that helps their students use and interpret SPSS.

Major Changes and Additions to This Edition

The following changes and additions are based on our experiences using the book with students, feedback from reviewers and other users, and the revisions in policy and best practice specified by the APA Task Force on Statistical Inference (1999) and the 5th Edition of the *APA Publication Manual* (2001).

1. ***Effect size.*** We discuss effect size in addition to statistical significance in the interpretation sections to be consistent with the requirements of the revised APA manual. Because SPSS does not provide effect sizes for all the demonstrated statistics, we often show how to estimate or compute them by hand.
2. ***Writing about outputs.*** We include examples of how to write about and make APA type tables from the information in SPSS outputs. We have found the step from interpretation to writing quite difficult for students so we now put more emphasis on writing.
3. ***Assumptions.*** When each statistic is introduced, we have a brief section about its assumptions and when it is appropriate to select that statistic for the problem or question at hand.
4. ***Testing assumptions.*** We have expanded emphasis on exploratory data analysis (EDA) and how to test assumptions.
5. ***Quick Reference Guide for SPSS procedures.*** We have condensed several of the appendixes of the first edition into the alphabetically organized Appendix A, which is somewhat like a glossary. It includes how to do basic statistics that are not included in this text, and procedures like print and save, which are tasks you will use several times and/or may already know. It also includes brief directions of how to do things like import a file from Excel or export to PowerPoint, do split files, and make 3-D figures.
6. ***Extra SPSS problems.*** We have developed additional extra problems, to give you more practice in running and interpreting SPSS.
7. ***Reliability assessment.*** We include a chapter on ways of assessing reliability including Cronbach’s alpha, Cohen’s kappa, and correlation. More emphasis on reliability and testing assumptions is consistent with our strategy of presenting SPSS procedures that students would use in an actual research project.
8. ***Principal Components Analysis and Exploratory Factor Analysis.*** We have added a section on exploratory factor analysis to increase students’ choices when using these types of analyses.

Preface

9. **Interpretation questions.** We have added more interpretation questions to each chapter because we have found them useful for student understanding. We include the answers to the odd numbered questions in Appendix C for self-study.

Bullets, Arrows, Bold and Italics

To help you do the problems with SPSS, we have developed some conventions. We use bullets to indicate actions in SPSS Windows that you will take. For example:

- Highlight *gender* and *math achievement*.
- Click on the arrow to move the variables into the right hand box.
- Click on **Options** to get Fig 2.16.
- Check **Mean, Std Deviation, Minimum, and Maximum**.
- Click on **Continue**.

Note that the words in italics are variable names and words in bold are words that you will see in the SPSS Windows and utilize to produce the desired output. In the text they are spelled and capitalized as you see them in the Windows. Bold is also used to identify key terms when they are introduced, defined, or important to understanding.

The words you will see in the pull down menus are given in bold with arrows between them. For example:

- Select **Analyze => Descriptive Statistics => Frequencies**
(This means pull down the Analyze menu, then slide your cursor down to Descriptive Statistics and over to Frequencies and click.)

Occasionally, we have used underlines to emphasize critical points or commands.

Acknowledgements

This SPSS book is consistent with and could be used as a supplement for Gliner and Morgan, (2000) *Research Methods in Applied Settings: An Integrated Approach to Design and Analysis*, also published by Erlbaum. In fact, some sections of chapters 1 and 3 have been only slightly modified from that text. For this we thank Jeff Gliner, the first author of that book. Although Orlando Griego is not an author on this revision of our SPSS book, it still shows the imprint of his student friendly writing style.

We would like to acknowledge the assistance of the many students in our education and human development classes who have used earlier versions of this book and provided helpful suggestions for improvement. We could not have completed the task or made it look so good without our technology consultant, Don Quick, our word processors, Linda White and Catherine Lamana, and several capable work study students including Rae Russell, Katie Jones, Erica Snyder, and Jennifer Musser. Jikyeong Kang, Bill Sears, LaVon Blaesi, Mei-Huei Tsay and Sheridan Green assisted with classes and the development of materials for the DOS and earlier Windows versions of the assignments. Laura Jensen, Lisa Vogel, Don Quick, James Lyall, Joan Anderson, and Yasmine Andrews helped with writing or editing parts of the manuscript or earlier editions. Jeff Gliner, Jerry Vaske, Jim zumBrunnen, Laura Goodwin, David MacPhee, Gene Gloeckner, James O. Benedict, Barry Cohen, John Ruscio, Tim Urdan, and Steve Knotek provided reviews and suggestions for improving the text. Joan Clay and Don Quick wrote helpful appendices for this edition. Bob Fetch and Ray Yang provided helpful feedback on the readability

Preface

and user friendliness of the text. We also acknowledge the financial assistance of two instructional improvement grants from the College of Applied Human Sciences at Colorado State University. Finally, the patience of our families enabled us to complete the task, without too much family strain.

N. L., K. B., and G. M.
Fort Collins, Colorado
July, 2004

This page intentionally left blank

SPSS for Intermediate Statistics: Use and Interpretation

Second Edition

This page intentionally left blank

CHAPTER 1

Introduction

This chapter will review important information about measurement and descriptive statistics and provide an overview of the expanded high school and beyond (HSB) data set, which will be used in this chapter and throughout the book to demonstrate the use and interpretation of the several statistics that are presented. First, we provide a brief review of some key terms, as we will use them in this book.

Variables

Variables are key elements in research. A **variable** is defined as a characteristic of the participants or situation for a given study that has different values in that study. A variable must be able to vary or have different values or levels.¹ For example, *gender* is a variable because it has two levels, female or male. *Age* is a variable that has a large number of values. *Type of treatment/intervention* (or *type of curriculum*) is a variable if there is more than one treatment or a treatment and a control group. *Number of days to learn something or to recover from an ailment* are common measures of the effect of a treatment and, thus, are also variables. Similarly, *amount of mathematics knowledge* is a variable because it can vary from none to a lot. If a concept has only one value in a particular study, it is not a variable; it is a constant. Thus, ethnic group is not a variable if all participants are European American. Gender is not a variable if all participants in a study are female.

In quantitative research, variables are **defined operationally** and are commonly divided into **independent variables** (active or attribute), **dependent variables**, and **extraneous variables**. Each of these topics will be dealt with briefly in the following sections.

Operational definitions of variables. An operational definition describes or defines a variable in terms of the operations or techniques used to make it happen or measure it. When quantitative researchers describe the variables in their study, they specify what they mean by demonstrating how they measured the variable. Demographic variables like age, gender, or ethnic group are usually measured simply by asking the participant to choose the appropriate category from a list. Types of treatment (or curriculum) are usually operationally defined much more extensively by describing what was done during the treatment or new curriculum. Likewise, abstract concepts like mathematics knowledge, self-concept, or mathematics anxiety need to be defined operationally by spelling out in some detail how they were measured in a particular study. To do this, the investigator may provide sample questions, append the actual instrument, or provide a reference where more information can be found.

Independent Variables

In this book, we will refer to two types of independent variables: **active** and **attribute**. It is important to distinguish between these types when we discuss the results of a study.

¹ To help you, we have identified the SPSS variable names, labels, and values using italics (e.g., *gender* and *male*). Sometimes italics are also used to emphasize a word. We have put in bold the terms used in the SPSS windows and outputs (e.g., **SPSS Data Editor**) and other key terms when they are introduced, defined, or are important to understanding. Underlines are used to emphasize critical points. Bullets precede instructions about SPSS actions (e.g., click, highlight).

Active or manipulated independent variables. An active independent variable is a variable, such as a workshop, new curriculum, or other intervention, one level of which is given to a group of participants, within a specified period of time during the study. For example, a researcher might investigate a new kind of therapy compared to the traditional treatment. A second example might be to study the effect of a new teaching method, such as cooperative learning, on student performance. In these two examples, the variable of interest was something that was *given* to the participants. Thus, active independent variables are *given* to the participants in the study but are not necessarily given or manipulated by the experimenter. They may be given by a clinic, school, or someone other than the investigator, but from the participants' point of view, the situation was manipulated. Using this definition, the treatment is usually given after the study was planned so that there could have been (or preferably was) a pretest. Other writers have similar but, perhaps, slightly different definitions of active independent variables. **Randomized experimental and quasi-experimental studies** have an active independent variable. An active independent variable is a necessary but not sufficient condition to make cause and effect conclusions; the clearest causal conclusions can be drawn when participants are assigned randomly to conditions that are manipulated by the experimenter.

Attribute or measured independent variables. A variable that cannot be manipulated, yet is a major focus of the study, can be called an attribute independent variable. In other words, the values of the independent variable are preexisting attributes of the persons or their ongoing environment that are not systematically changed during the study. For example, education, gender, age, ethnic group, IQ, and self-esteem are attribute variables that could be used as attribute independent variables. Studies with only attribute independent variables are called **nonexperimental studies**.

In keeping with SPSS, but unlike authors of some research methods books, we do not restrict the term independent variable to those variables that are manipulated or active. We define an independent variable more broadly to include any predictors, antecedents, or *presumed* causes or influences under investigation in the study. Attributes of the participants, as well as active independent variables, fit within this definition. For the social sciences and education, attribute independent variables are especially important. Type of disability or level of disability may be the major focus of a study. Disability certainly qualifies as a variable since it can take on different values even though they are not *given* during the study. For example, cerebral palsy is different from Down syndrome, which is different from spina bifida, yet all are disabilities. Also, there are different levels of the same disability. People already have defining characteristics or attributes that place them into one of two or more categories. The different disabilities are already present when we begin our study. Thus, we might also be interested in studying a class of variables that are not given or manipulated during the study, even by other persons, schools, or clinics.

Other labels for the independent variable. SPSS uses a variety of terms in addition to independent variable; for example, **factor** (chapters 8, 9, and 10), and **covariates** (chapter 7). In other cases, (chapters 4 and 5) SPSS and statisticians do not make a distinction between the independent and dependent variable; they just label them **variables**. For example, there is no independent variable for a correlation or chi-square. However, even for chi-square and correlation, we think it is sometimes educationally useful to think of one variable as the predictor (independent variable) and the other as the outcome (dependent variable), as is the case in regression.

Values of the independent variable. SPSS uses the term **values** to describe the several options or values of a variable. These values are *not* necessarily ordered, and several other terms, **categories**, **levels**, **groups**, or **samples** are sometimes used interchangeably with the term values, especially in statistics books. Suppose that an investigator is performing a study to investigate the effect of a treatment. One

Chapter 1 - Introduction

group of participants is assigned to the treatment group. A second group does not receive the treatment. The study could be conceptualized as having one independent variable (*treatment type*), with two values or levels (*treatment* and *no treatment*). The independent variable in this example would be classified as an active independent variable. Now, suppose instead, that the investigator was interested primarily in comparing two different treatments but decided to include a third no-treatment group as a control group in the study. The study still would be conceptualized as having one active independent variable (*treatment type*), but with three values or levels (the two treatment conditions and the control condition). As an additional example, consider *gender*, which is an attribute independent variable with two values, as *male* and *female*.

Note that in SPSS each variable is given a **variable label**; moreover, the values, which are often categories, have **value labels** (e.g., male and female). Each value or level is assigned a number used by SPSS to compute statistics. It is especially important to know the value labels when the variable is **nominal** (i.e., when the values of the variable are just names and, thus, are not ordered).

Dependent Variables

The **dependent variable** is assumed to measure or assess the effect of the independent variable. It is thought of as the presumed outcome or criterion. Dependent variables are often test scores, ratings on questionnaires, readings from instruments (electrocardiogram, galvanic skin response, etc.), or measures of physical performance. When we discuss measurement, we are usually referring to the dependent variable. Dependent variables, like independent variables must have at least two values; most dependent variables have many values, varying from low to high.

SPSS also uses a number of other terms in addition to dependent variable. **Dependent list** is used in cases where you can do the same statistic several times, for a list of dependent variables (e.g., in chapter 8 with one-way ANOVA). **Grouping variable** is used in chapter 7 for discriminant analysis.

Extraneous Variables

These are variables (also called nuisance variables or, in some designs, covariates) that are not of primary interest in a particular study but could influence the dependent variable. Environmental factors (e.g., temperature or distractions), time of day, and characteristics of the experimenter, teacher, or therapist are some possible extraneous variables that need to be controlled. SPSS does not use the term extraneous variable. However, sometimes such variables are controlled using statistics that are available in SPSS.

Research Hypotheses and Research Questions

Research hypotheses are predictive statements about the relationship between variables. **Research questions** are similar to hypotheses, except that they do not entail specific predictions and are phrased in question format. For example, one might have the following research question: “Is there a difference in students’ scores on a standardized test if they took two tests in one day versus taking only one test on each of two days?” A hypothesis regarding the same issue might be: “Students who take only one test per day will score *better* on standardized tests than will students who take two tests in one day.”

We divide research questions into three broad types: **difference**, **associational**, and **descriptive** as shown in the middle of Fig 1.1. The figure also shows the general and specific purposes and the general types of statistics for each of these three types of research question.

Difference research questions. For these questions, we compare scores (on the dependent variable) of two or more different groups, each of which is composed of individuals with one of the values or levels

on the independent variable. This type of question attempts to demonstrate that groups are not the same on the dependent variable.

Associational research questions are those in which two or more variables are associated or related. This approach usually involves an attempt to see how two or more variables covary (as one grows larger, the other grows larger or smaller) or how one or more variables enables one to predict another variable.

Descriptive research questions are not answered with inferential statistics. They merely describe or summarize data, without trying to generalize to a larger population of individuals.

Figure 1.1 shows that both difference and associational questions or hypotheses are similar in that they explore the relationships between variables.² Note that difference and associational questions differ in specific purpose and the kinds of statistics they use to answer the question.

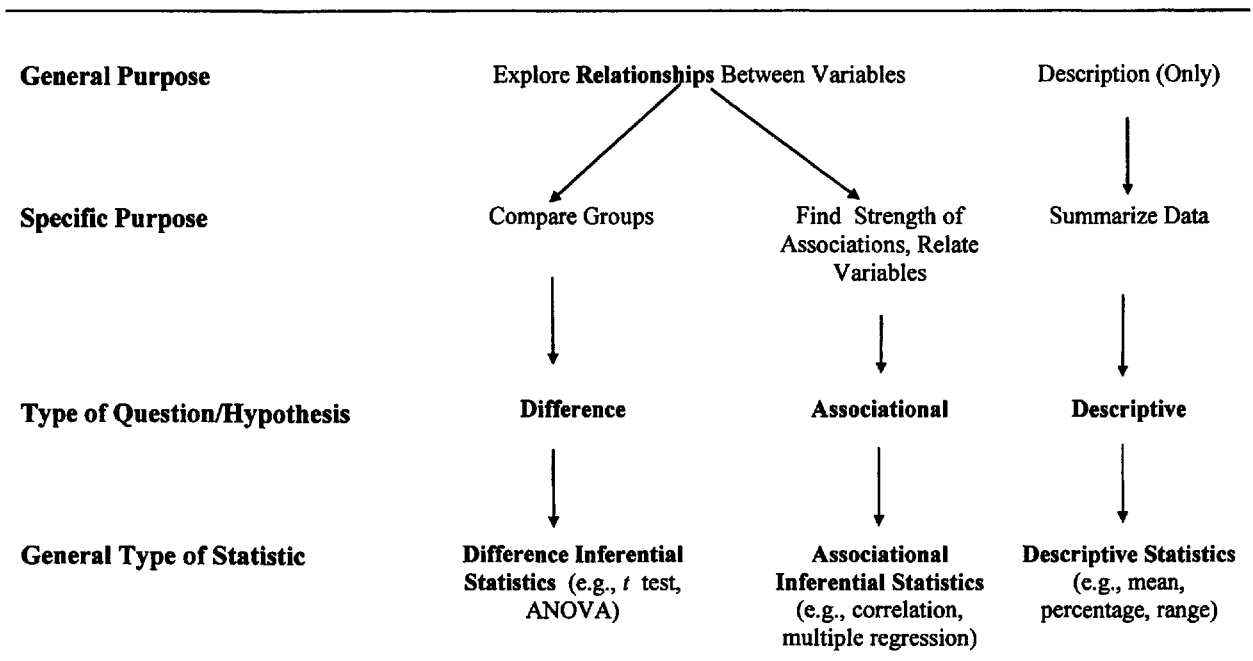


Fig. 1.1. Schematic diagram showing how the purpose and type of research question correspond to the general type of statistic used in a study.

²This similarity is in agreement with the statement by statisticians that all common parametric inferential statistics are relational. We use the term associational for the second type of research question rather than relational or correlational to distinguish it from the *general purpose* of both difference and associational questions/hypotheses, which is to study relationships. Also we wanted to distinguish between correlation, as a specific statistical technique, and the broader type of associational question and that group of statistics.

Difference versus associational inferential statistics. We think it is educationally useful to divide inferential statistics into two types, corresponding to difference and associational hypotheses or questions.³ **Difference inferential statistics** (e.g., *t* test or analysis of variance) are used for approaches that test for differences between groups. **Associational inferential statistics** test for associations or relationships between variables and use, for example, correlation or multiple regression analysis. We will utilize this contrast between difference and associational inferential statistics in chapter 3 and later in this book.

Remember that research questions are similar to hypotheses, but they are stated in question format. We think it is advisable to use the question format when one does not have a clear directional prediction and for the descriptive approach. As implied by Fig. 1.1, it is acceptable to phrase any research question that involves two variables as whether or not there is a relationship between the variables (e.g., “Is there a relationship between *gender* and *math achievement*?” or “Is there a relationship between *anxiety* and *GPA*?”). However, we think that phrasing the question as a difference or association is desirable because it helps one choose an appropriate statistic and interpret the result.

Complex Research Questions

Most research questions posed in this book involve more than two variables at a time. We call such questions and the appropriate statistics **complex**. Some of these statistics are called **multivariate** in other texts, but there is not a consistent definition of multivariate in the literature. We provide examples of how to write complex research questions in the chapter pertaining to each complex statistic.

In a factorial ANOVA, there are two (or more) independent variables and one dependent variable. We will see, in chapter 8, that although you do one factorial ANOVA, there are actually three (or more) research questions. This *set* of three questions can be considered a complex difference question because the study has two independent variables. Likewise, complex associational questions are used in studies with more than one independent variable considered together.

Table 1.1 expands our overview of research questions to include both basic and complex questions of each of the three types: **descriptive, difference, and associational**. The table also includes references to other chapters in this book and examples of the types of statistics that we include under each of the six types of questions.

A Sample Research Problem: The Modified High School and Beyond (HSB) Study

The SPSS file name of the data set used with this book is *hsbdataB.sav*; it stands for high school and beyond data. It is based on a national sample of data from more than 28,000 high school students. The current data set is a sample of 75 students drawn randomly from the larger population. The data that we have for this sample includes school outcomes such as grades and the number of mathematics courses of different types that the students took in high school. Also, there are several kinds of standardized test data and demographic data such as gender and mother’s and father’s education. To provide an example

³ We realize that all parametric inferential statistics are relational, so this dichotomy of using one type of data analysis procedure to test for differences (when there are a few values or levels of the independent variables) and another type of data analysis procedure to test for associations (when there are continuous independent variables) is somewhat artificial. Both continuous and categorical independent variables can be used in a general linear model approach to data analysis. However, we think that the distinction is useful because most researchers utilize the above dichotomy in selecting statistics for data analysis.

of questionnaire data, we have included 14 questions about mathematics attitudes. These data were developed for this book and, thus, are not really the math attitudes of the 75 students in this sample; however, they are based on real data gathered by one of the authors to study motivation. Also, we made up data for religion, ethnic group, and SAT-math, which are somewhat realistic overall. These inclusions enable us to do some additional statistical analyses.

Table 1.1. Summary of Six Types of Research Questions and Appropriate Statistics

Type of Research Question - Number of Variables	Statistics (Example)
1) Basic Descriptive Questions – One variable	Table 1.5, ch. 1 (mean, standard deviation, frequency distribution)
2) Complex Descriptive Questions – Two or more variables, but no use of inferential statistics	Ch. 2, 4, 5 (mean & SD for one variable after forming groups based on another variable, factor analysis, measures of reliability)
3) Basic/Single Factor Difference Questions – One independent and one dependent variable. Independent variable usually has a few levels (ordered or not).	Table 3.1, QRG (<i>t</i> test, one-way ANOVA)
4) Complex/Multi Factor Difference Question – Three or more variables. Usually two or a few independent variables and one (or more) dependent variables.	Table 3.3, ch. 8, 9, 10 (factorial ANOVA, MANOVA)
5) Basic Associational Questions – One independent variable and one dependent variable. Usually at least five <i>ordered</i> levels for both variables. Often they are continuous.	Table 3.2, QRG (correlation tested for significance)
6) Complex/Multivariate Associational Questions – Two or more independent variables and one dependent variable. Often five or more ordered levels for all variables but some or all can be dichotomous variables.	Table 3.4, ch. 6, 7 (multiple or logistic regression)

Note: Many studies have more than one dependent variable. It is common to treat each one separately (i.e., to do several *t* tests, ANOVAs, correlations, or multiple regressions). However, there are complex statistics (e.g., MANOVA and canonical correlation) used to treat several dependent variables together in one analysis. QRG = Quick Reference Guide, see Appendix A.

The Research Problem

Imagine that you are interested in the general problem of what factors seem to influence mathematics achievement at the end of high school. You might have some hunches or hypotheses about such factors based on your experience and your reading of the research and popular literature. Some factors that might influence mathematics achievement are commonly called demographics: for example, *gender*, *ethnic group*, and *mother's and father's education*. A probable influence would be the mathematics courses that the student has taken. We might speculate that *grades in mathematics* and in other subjects could have an

impact on *math achievement*.⁴ However, other variables, such as *students' IQs* or *parents' encouragement and assistance*, could be the actual causes of both high grades and math achievement. Such variables could influence what courses one took, the grades one received, and might be correlates of the demographic variables. We might wonder how spatial performance scores, such as *pattern* or *mosaic pattern test* scores and *visualization scores* might enable a more complete understanding of the problem, and whether these skills seem to be influenced by the same factors as *math achievement*.

The HSB Variables

Before we state the research problem and questions in more formal ways, we need to step back and discuss the types of variables and the approaches that might be used to study the above problem. We need to identify the independent/antecedent (presumed causes) variables, the dependent/outcome variable(s), and any extraneous variables.

The primary dependent variable. Given the above research problem which focuses on mathematics achievement at the end of the senior year, the primary dependent variable is *math achievement*.

Independent and extraneous variables. The *number of math courses taken* up to that point is best considered to be an antecedent or independent variable in this study. What about *father's* and *mother's education* and *gender*? How would you classify *gender* and *parents' education* in terms of the type of variable? What about *grades*? Like the *number of math courses*, these variables would usually be considered independent variables because they occurred before the math achievement test. However, some of these variables, specifically parental education, might be viewed as extraneous variables that need to be "controlled." *Visualization* and *mosaic pattern test* scores probably could be either independent or dependent variables depending upon the specific research question, because they were measured at approximately the same time as math achievement, at the end of the senior year. Note that student's class is a constant and is not a variable in this study because all the participants are high school seniors (i.e., it does not vary; it is the population of interest).

Types of independent variables. As we discussed previously, independent variables can be **active** (given to the participant during the study or manipulated by the investigator) or **attributes** of the participants or their environments. Are there any **active** independent variables in this study? No! There is no intervention, new curriculum, or similar treatment. All the independent variables, then, are attribute variables because they are attributes or characteristics of these high school students. Given that all the independent variables are attributes, the research approach cannot be experimental. This means that we will *not* be able to draw definite conclusions about cause and effect (i.e., we will find out what is related to math achievement, but we will not know for sure what causes or influences math achievement).

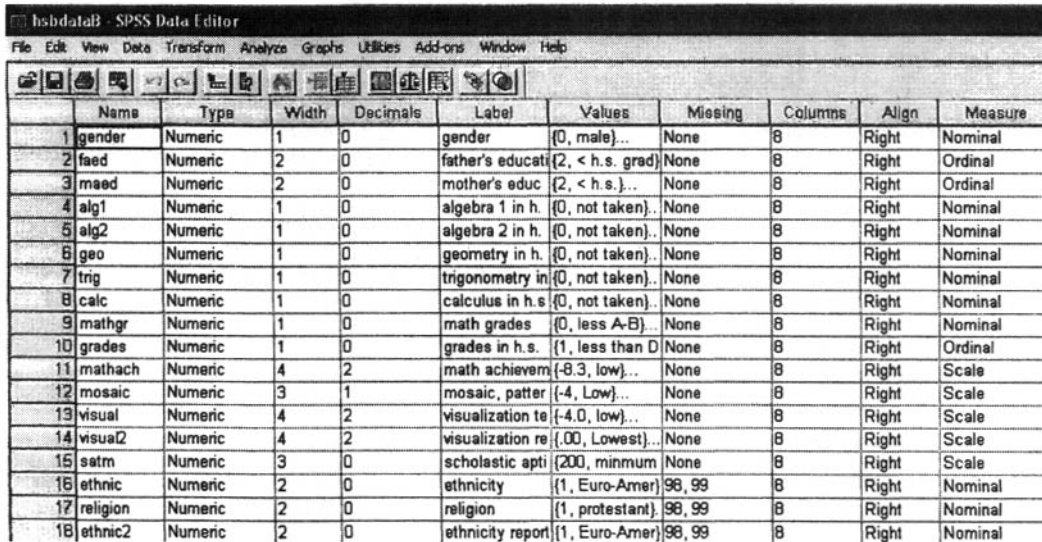
Now we will examine the *hsbdataB.sav* that you will use to study this complex research problem. We have provided a CD that contains the data for each of the 75 participants on 45 variables. The variables in the *hsbdataB.sav* file have already been labeled (see Fig 1.2) and entered (see Fig 1.3) to enable you to get started on analyses quickly. The CD in this book contains several SPSS data files for you to use, but it does not include the actual SPSS program, which you will have to have access to in order to do the assignments.

The SPSS Variable View

Figure 1.2 is a piece of what SPSS calls the **variable view** in the **SPSS Data Editor** for the *hsbdataB.sav* file. Figure 1.2 shows information about each of the first 18 variables. When you open this file and click

⁴ We have decided to use the short version of mathematics (i.e., math) throughout the book to save space and because it is used in common language.

on **Variable View** at the bottom left corner of the screen, this is what you will see. What is included in the variable view screen is described in more detail in Appendix B, Getting Started. Here, focus on the Name, Label, Values, and Missing columns. **Name** is a short name for each variable (e.g., *faed* or *alg1*).⁵ **Label** is a longer label for the variable (e.g., *father's education* or *algebra 1 in h.s.*). The **Values** column contains the **value labels**, but you can see only the label for one value at a time (e.g., 0=male). That is, you cannot see that 1=female unless you click on the row for that variable under the value column. The **Missing** column indicates whether there are any special, user-identified missing values. **None** just means that there are no special missing values, just the usual *SPSS system missing* value, which is a blank.



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	gender	Numeric	1	0	gender	(0, male)...	None	8	Right	Nominal
2	faed	Numeric	2	0	father's education	(2, < h.s. grad)	None	8	Right	Ordinal
3	maed	Numeric	2	0	mother's education	(2, < h.s.)...	None	8	Right	Ordinal
4	alg1	Numeric	1	0	algebra 1 in h.s.	(0, not taken)...	None	8	Right	Nominal
5	alg2	Numeric	1	0	algebra 2 in h.s.	(0, not taken)...	None	8	Right	Nominal
6	geo	Numeric	1	0	geometry in h.s.	(0, not taken)...	None	8	Right	Nominal
7	trig	Numeric	1	0	trigonometry in h.s.	(0, not taken)...	None	8	Right	Nominal
8	calc	Numeric	1	0	calculus in h.s.	(0, not taken)...	None	8	Right	Nominal
9	mathgr	Numeric	1	0	math grades	(0, less A-B)...	None	8	Right	Nominal
10	grades	Numeric	1	0	grades in h.s.	(1, less than D)	None	8	Right	Ordinal
11	mathach	Numeric	4	2	math achievement	(-8.3, low)...	None	8	Right	Scale
12	mosaic	Numeric	3	1	mosaic, pattern	(-4, Low)...	None	8	Right	Scale
13	visual	Numeric	4	2	visualization test	(-4.0, low)...	None	8	Right	Scale
14	visual2	Numeric	4	2	visualization retest	(.00, Lowest)...	None	8	Right	Scale
15	satm	Numeric	3	0	scholastic aptitude	(200, minimum)	None	8	Right	Scale
16	ethnic	Numeric	2	0	ethnicity	(1, Euro-Amer)	98, 99	8	Right	Nominal
17	religion	Numeric	2	0	religion	(1, protestant)	98, 99	8	Right	Nominal
18	ethnic2	Numeric	2	0	ethnicity report	(1, Euro-Amer)	98, 99	8	Right	Nominal

Fig. 1.2. Part of the hsbdataB.sav variable view in the SPSS data editor.

Variables in the Modified HSB Data Set

The 45 variables shown in Table 1.2 (with the values/levels or range of their values in parentheses) are found in the hsbdata.savB file on the CD in the back of the book. Note that variables 33-38 and 42-44 were computed from the math attitude variables (19-32).

The variables of *ethnic* and *religion* were added to the original HSB data set to provide true nominal (unordered) variables with a few (4 and 3) levels or values. In addition, for *ethnic* and *religion*, we have made two missing value codes to illustrate this possibility. All other variables use blanks, the SPSS system missing value, for missing data. For *ethnicity*, 98 indicates *multiethnic* and other. For *religion*, all the high school students who were not *protestant* or *catholic* or said they had *no religion* were coded 98 and considered to be missing because none of the other religions had enough members to make a reasonable size group. Those who left the ethnicity or religion questions blank were coded as 99, also missing.

⁵ In SPSS 7-11, the variable name had to be 8 characters or less. In SPSS 12, it can be longer, but we recommend that you keep it short. If a longer name is used with SPSS 7-11, the name will be truncated. SPSS names must start with a letter and must not contain blank spaces or certain special characters (e.g., !, ?, ', or *).

Table 1.2. HSB Variable Descriptions

Name	Label (and Values)
Demographic School and Test Variables	
1. gender	gender (0 = male, 1 = female).
2. faed	father's education (2 = less than h.s. to 10 = PhD/MD).
3. maed	mother's education (2 = less than h.s. grad to 10 = PhD/MD).
4. alg1	algebra 1 in h.s. (1 = taken, 0 = not taken)
5. alg2	algebra 2 in h.s. (1 = taken, 0 = not taken)
6. geo	geometry in h.s. (1 = taken, 0 = not taken)
7. trig	trigonometry in h.s. (1 = taken, 0 = not taken)
8. calc	calculus in h.s. (1 = taken, 0 = not taken)
9. mathgr	math grades (0 = low, 1 = high)
10. grades	grades in h.s. (1 = less than a D average to 8 = mostly an A average)
11. mathach	math achievement score (-8.33 to 25). ⁶ This is a test something like the ACT math.
12. mosaic	mosaic, pattern test score (-4 to 56). This is a test of pattern recognition ability involving the detection of relationships in patterns of tiles.
13. visual	visualization score (-4 to 16). This is a 16-item test that assesses visualization in three dimensions (i.e., how a three-dimensional object would look if its spatial position were changed).
14. visual2	visualization retest – the visualization test score students obtained when they retook the test a month or so later.
15. satm	scholastic aptitude test – math (200 = lowest, 800 = highest possible)
16. ethnic	ethnicity (1 = Euro-American, 2 = African-American, 3 = Latino-American, 4 = Asian-American, 98 = other or multiethnic, chose 2 or more, 99 = missing, left blank)
17. religion	religion (1 = protestant, 2 = catholic, 3 = no religion, 98 = chose one of several other religions, 99 = left blank)
18. ethnic2	ethnicity reported by student (same as values for ethnic)
Math Attitude Questions 1–14 (Rated from 1 = very atypical to 4 = very typical)	
19. item01	motivation - "I practice math skills until I can do them well."
20. item02	pleasure - "I feel happy after solving a hard problem."
21. item03	competence - "I solve math problems quickly."
22. item04	(low) motiv - "I give up easily instead of persisting if a math problem is difficult."
23. item05	(low)comp - "I am a little slow catching on to new topics in math."
24. item06	(low)pleas - "I do not get much pleasure out of math problems."
25. item07	motivation - "I prefer to figure out how to solve problems without asking for help."
26. item08	(low)motiv - "I do not keep at it very long when a math problem is challenging."
27. item09	competence - "I am very competent at math."
28. item10	(low)pleas - "I smile only a little (or not at all) when I solve a math problem."
29. item11	(low)comp - "I have some difficulties doing math as well as other kids my age."

⁶Negative test scores result from a penalty for guessing.

30. item12 *motivation* - "I try to complete my math problems even if it takes a long time to finish."
31. item13 *motivation* - "I explore all possible solutions of a complex problem before going on to another one."
32. item14 *pleasure* - "I really enjoy doing math problems."

Variables Computed From the Above Variables

33. item04r *item04 reversed* (4 now = high motivation)
34. item05r *item05 reversed* (4 now = high competence)
35. item08r *item08 reversed* (4 now = high motivation)
36. item11r *item11 reversed* (4 now = high competence)
37. competence *competence scale*. An average computed as follows: (item03 + item05r + item09 + item11r)/4
38. motivation *motivation scale* (item01 + item04r + item07 + item08r + item12 + item13)/6
39. mathcrs *math courses taken* (0 = none, 5 = all five)
40. faedRevis *father's educ revised* (1 = HS grad or less, 2 = some college, 3 = BS or more)
41. maedRevis *mother's educ revised* (1 = HS grad or less, 2 = some college, 3 = BS or more)
42. item06r *item06 reversed* (4 now = high pleasure)
43. item10r *item10 reversed* (4 now = high pleasure)
44. pleasure *pleasure scale* (item02 + item06r + item10r + item14)/4
45. parEduc *parents' education* (average of the unrevised mother's and father's educations)

The Raw HSB Data and Data Editor

Figure 1.3 is a piece of the hsbdata.savB file showing the first 11 student participants for variables 1 through 17 (*gender* through *religion*). Notice the short variable names (e.g. *faed*, *alg1*, etc.) at the top of the hsbdataB file. Be aware that the participants are listed down the left side of the page, and the variables are always listed across the top. You will always enter data this way. If a variable is measured more than once, such as *visual* and *visual2* (see Fig 1.3), it will be entered as two variables with slightly different names.

	gender	faed	maed	alg1	alg2	geo	trig	calc	mathgr	grades	mathach	mosaic	visual	visual2	satm	ethnic	religion
1	1	10	10	0	0	0	0	0	0	4	9.00	31.0	8.75	7.00	500	2	1
2	1	2	2	0	0	0	0	0	0	5	10.33	56.0	4.75	4.00	460	3	2
3	1	2	2	0	0	0	0	0	1	6	7.67	25.0	4.75	3.00	420	2	2
4	0	3	3	1	0	0	0	0	0	3	5.00	22.0	1.00	1.00	400	1	1
5	1		3	0	0	0	0	0	0	3	-1.67	17.5	2.25	2.00	450	2	1
6	1	3	2	0	0	0	0	0	1	5	1.00	23.5	1.00	.00	250	3	2
7	0	9	6	1	1	1	0	0	0	6	12.00	28.5	2.50	2.00	480	1	2
8	1	5	3	1	0	0	0	0	0	4	8.00	29.5	3.50	3.00	430	1	3
9	1	3	3	1	0	0	0	0	1	7	13.00	28.0	3.50	2.00	490	1	98
10	0	8	2	0	0	0	0	0	0	5	3.67	27.5	3.75	3.00	300	2	1
11	0	3	4	1	1	1	1	1	0	6	21.00	27.0	11.00	9.50	500	4	99

Fig. 1.3. Part of the hsbdataB data view in the SPSS data editor.

Note that in Fig. 1.3, most of the values are single digits, but *mathach*, *mosaic*, and *visual* include some decimals and even negative numbers. Notice also that some cells, like *father's education* for participant 5, are blank because a datum is missing. Perhaps participant 5 did not know her father's education. Blank

is the system missing value that can be used for any missing data in an SPSS data file. We suggest that you leave missing data blank; however, you may run across “user defined” missing data codes like -1, 9, 98, or 99 in other researchers’ data (see *religion*, subjects 9 and 11).

Research Questions for the Modified HSB Study⁷

In this book, we will generate a large number of research questions from the modified HSB data set. In this section, we list some of the possible questions you might ask in order to give you an idea of the range of types of questions that one might have in a typical research project like a thesis or dissertation. For review, we start with basic *descriptive questions* and some questions about assumptions. Also for review, we list some basic *difference* and *associational questions* that can be answered with basic (two-variable) inferential statistics such as a *t* test or correlation. These statistics are not discussed in this book but how to compute them can be found in the Quick Reference Guide (Appendix A). Finally, we pose a number of complex questions that can be answered with the statistics discussed in this book.

- 1) Often, we start with **basic descriptive questions** about the demographics of the sample. Thus, we could answer, with the outputs in chapter 2, the following basic descriptive questions: “What is the average educational level of the fathers of the students in this sample?” “What percentage of the students is male and what percentage is female?”
- 2) In chapter 2, we also examine whether the continuous variables are distributed normally, an **assumption** of many statistics. One question is “Are the frequency distributions of the math achievement scores markedly **skewed**; that is, different from the **normal curve distribution**?”
- 3) How to make a table crosstabulating two categorical variables (ones with a few values or categories) is described in the Quick Reference Guide. Crosstabulation and the **chi-square** statistic can answer research questions such as “Is there a relationship between *gender* and *math grades* (high or low)?
- 4) In the Quick Reference Guide, we will also describe how to answer **basic associational research questions** (using **Pearson product-moment** correlation coefficients) such as, “Is there a positive association/relationship between *grades in high school* and *math achievement*?” A **correlation matrix** of all the correlations among several key variables including *math achievement* will provide the basis for computing **factor analysis** in chapter 5 and **multiple regression** in chapter 6. In chapter 4, correlation is also used to assess **reliability**.
- 5) **Basic difference questions** such as “Do males and females differ on *math achievement*?” and “Are there differences among the three *father’s education* groups in regard to average scores on *math achievement*?” can be answered with a *t* test or **one-way ANOVA**, as described in the Quick Reference Guide.
- 6) **Complex difference questions** will be answered with **factorial ANOVA** in chapter 8. One *set* of three questions is as follows: (1) “Is there a difference between students who have fathers with no

⁷ The High School and Beyond (HSB) study was conducted by the National Opinion Research Center (1980). The example discussed here and throughout the book is based on 13 variables obtained from a random sample of 75 out of 28,240 high school seniors. These variables include achievement scores, grades, and demographics. The raw data for the 13 variables were slightly modified from data in an appendix in Hinkle, Wiersma, and Jurs (1994). That file had no missing data, which is unusual in behavioral science research, so we made some.

college, some college, or a BS or more with respect to the student's *math achievement*?" (2) "Is there a difference between students who had a B or better *math grade* average and those with less than a B average on a *math achievement* test at the end of high school?" and (3) "Is there an **interaction** between *father's education* and *math grades* with respect to *math achievement*?"

- 7) How well can you predict *math achievement* from a combination of four variables: *motivation scale*, *grades in high school*, *parents' education*, and *gender*? This *complex associational question* can be answered with **multiple regression**, as discussed in chapter 6. If the dependent variable, *math achievement*, were dichotomous (high vs. low achievement), **logistic regression** or possibly **discriminant analysis**, discussed in chapter 7, would be appropriate to answer this question.
- 8) Are there differences between the three *father's education* groups on a linear combination of *grades in high school*, *math achievement*, and *visualization test*? This complex difference question can be answered with a **single-factor multivariate analysis of variance (MANOVA)**, as discussed in chapter 10.

More complex questions can be found in chapters 4-10 under each of the several "problems" in those chapters.

This introduction to the research problem and questions raised by the HSB data set should help make the assignments meaningful, and it should provide a guide and examples for your own research.

Frequency Distributions

Frequency distributions are critical to understanding our use of measurement terms. We begin this section with a discussion of frequency distributions and two examples. Frequency tables and distributions can be used whether the variable involved has ordered or unordered levels (SPSS calls them **values**). In this section, we will only consider variables with many ordered values, ranging from low to high.

A **frequency distribution** is a tally or count of the number of times each score on a single variable occurs. For example, the frequency distribution of final grades in a class of 50 students might be 7 As, 20 Bs, 18 Cs, and 5 Ds. Note that in this frequency distribution most students have Bs or Cs (grades in the middle) and similar smaller numbers have As and Ds (high and low grades). When there are small numbers of scores for the low and high values and most scores are for the middle values, the distribution is said to be **approximately normally distributed**. We will discuss the normal curve in more detail later in this chapter.

When the variable is continuous or has many ordered levels (or values), the frequency distribution usually is based on ranges of values for the variable. For example, the frequencies (number of students), shown by the bars in Fig 1.4, are for a range of points (in this case, SPSS selected a range of 50, 250-299, 300-349, 350-399, etc.). Notice that the largest number of students (about 20) has scores in the middle two bars of the range (450-550). Similar small numbers of students have very low and very high scores. The bars in the histogram form a distribution (pattern or curve) that is quite similar to the normal, bell-shaped curve shown by the line that is superimposed on the histogram. Thus, the frequency distribution of the *SAT math* scores is said to be **approximately normal**.

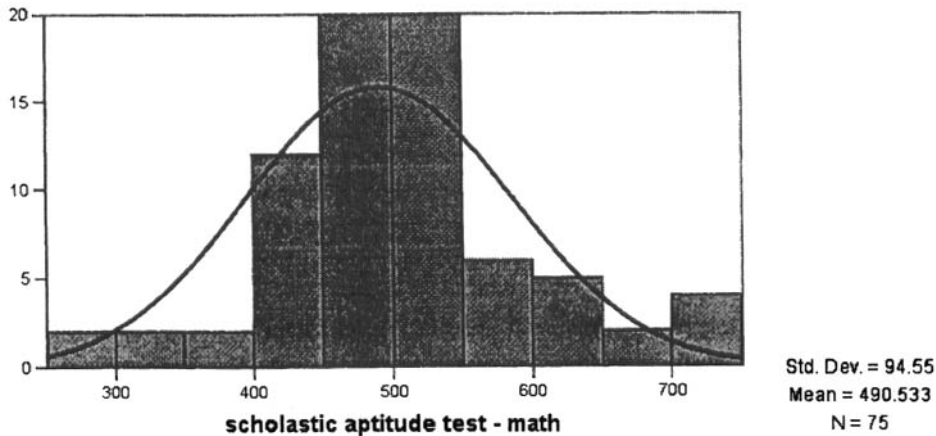


Fig. 1.4. A grouped frequency distribution for SAT math scores.

Figure 1.5 shows the frequency distribution for the *competence scale*. Notice that the bars form a pattern very different from the normal curve line. This distribution can be said to be **not normally distributed**. As we will see later in the chapter, the distribution is **negatively skewed**. That is, the tail of the curve or the extreme scores are on the low end or left side. Note how much this differs from the *SAT math* score frequency distribution. As you will see in the Measurement section (below), we call the *competence scale* variable **ordinal**.

You can create these figures yourself using the hsbdataB.sav file. Select:

- **Graphs => Histogram.**
- Then move *scholastic aptitude test – math* (or *competence scale*) into the **Variable** box.
- Then check **Display normal curve**.
- Click **OK**.

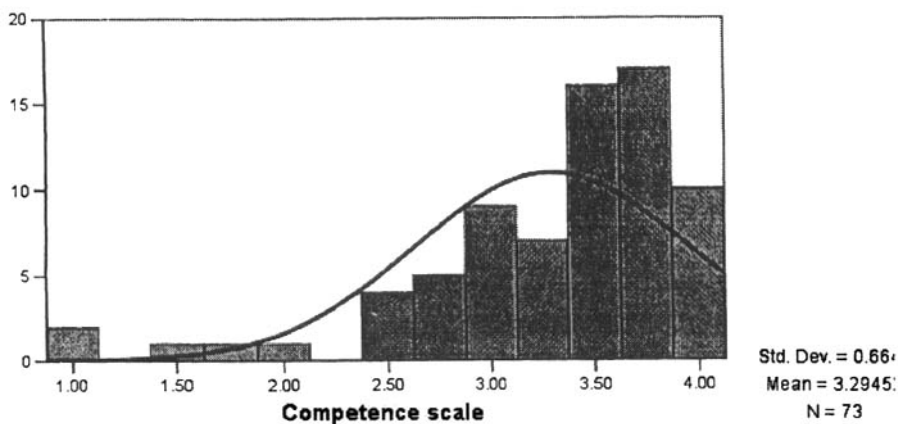


Fig. 1.5. A grouped frequency distribution for the competence scale.

Levels of Measurement

Measurement is the assignment of numbers or symbols to the different characteristics (values) of variables according to rules. In order to understand your variables, it is important to know their level of measurement. Depending on the level of measurement of a variable, the data can mean different things.

For example, the number 2 might indicate a score of two; it might indicate that the subject was a male; or it might indicate that the subject was ranked second in the class. To help understand these differences, types or levels of variables have been identified. It is common and traditional to discuss four levels or scales of measurement: **nominal**, **ordinal**, **interval**, and **ratio**, which vary from the unordered (nominal) to the highest level (ratio). These four traditional terms are not the same as those used in SPSS, and we think that they are not always the most useful for determining what statistics to use.

SPSS uses three terms (**nominal**, **ordinal**, and **scale**) for the levels of types or measurement. How these correspond to the traditional terms is shown in Table 1.3. When you name and label variables in SPSS, you have the opportunity to select one of these three types of measurement (see Fig 1.2). Although what you choose does *not* affect what SPSS does in most cases, an appropriate choice both indicates that you understand your data and may help guide your selection of statistics.

We believe that the terms **nominal**, **dichotomous**, **ordinal**, and **approximately normal** (for normally distributed) are usually more useful than the traditional or SPSS measurement terms for the selection and interpretation of statistics. In part, this is because statisticians disagree about the usefulness of the traditional levels of measurement in determining the appropriate selection of statistics. Furthermore, our experience is that the traditional terms are frequently misunderstood and applied inappropriately by students. Hopefully, our terms, as discussed below, are clear and useful.

Table 1.3 compares the three sets of terms and provides a summary description of our definitions of them. Professors differ in the terminology they prefer and on how much importance to place on levels or scales of measurement, so you will see all of these terms, and the others mentioned below, in textbooks and articles.

Table 1.3. Similar Traditional, SPSS, and Our Measurement Terms

Traditional Term	Traditional Definition	SPSS Term	Our Term	Our Definitions
Nominal	Two or more <u>unordered</u> categories	Nominal	Nominal	Three or more <u>unordered</u> categories.
			Dichotomous	Two categories, either ordered or unordered.
Ordinal	<u>Ordered</u> levels, in which the difference in magnitude between levels is not equal	Ordinal	Ordinal	Three or more <u>ordered</u> levels, but the frequency distribution of the scores is <u>not</u> normally distributed.
Interval & Ratio	Interval: <u>ordered</u> levels, in which the difference between levels is equal, but there is no true zero. Ratio: <u>ordered</u> levels; the difference between levels is equal, and there is a true zero.	Scale	Approximately Normal	Many (at least five) <u>ordered</u> levels or scores, with the frequency distribution of the scores being approximately normally distributed.

Nominal Variables

This is the most basic or lowest level of measurement, in which the numerals assigned to each category stand for the name of the category, but they have no implied order or value. For example, in the HSB study, the values for the *religion* variable are 1 = *protestant*, 2 = *catholic*, 3 = *no religion*. This does not mean that two protestants equal one catholic or any of the typical mathematical uses of the numerals. The same reasoning applies to many other true nominal variables, such as ethnic group, type of disability, or section number in a class schedule. In each of these cases, the categories are distinct and non-overlapping, but not ordered. Each category or group in the modified HSB variable *ethnicity* is different from each other but there is no necessary order to the categories. Thus, the categories could be numbered 1 for *Asian-American*, 2 for *Latino-American*, 3 for *African-American*, and 4 for *European-American* or the reverse or any combination of assigning one number to each category.

What this implies is that you must *not* treat the numbers used for identifying nominal categories as if they were numbers that could be used in a formula, added together, subtracted from one another, or used to compute an average. Average *ethnicity* makes no sense. However, if you ask SPSS to compute the mean *ethnicity*, it will do so and give you meaningless information. The important aspect of nominal measurement is to have clearly defined, non-overlapping or mutually exclusive categories that can be coded reliably by observers or by self-report.

Using nominal measurement does dramatically limit the statistics that can be used with your data, but it does not altogether eliminate the possible use of statistics to summarize your data and make inferences. Therefore, even when the data are unordered or nominal categories, your research may benefit from the use of appropriate statistics. Later we will discuss the types of statistics, both descriptive and inferential, that are appropriate for nominal data.

Other terms for nominal variables. Unfortunately, the literature is full of similar, but not identical terms to describe the measurement aspects of variables. **Categorical, qualitative, and discrete** are terms sometimes used interchangeably with nominal, but we think that nominal is better because it is possible to have ordered, discrete categories (e.g., low, medium, and high IQ, which we and other researchers would consider an ordinal variable). Qualitative is also used to discuss a different approach to doing research, with important differences in philosophy, assumptions, and approach to conducting research.

Dichotomous Variables

Dichotomous variables always have only two levels or categories. In some cases, they may have an implied order (e.g., *math grades* in high school are coded 0 for *less than an A or B* average and 1 for *mostly A or B*). Other dichotomous variables do not have any order to the categories (e.g., *male* or *female*). For many purposes, it is best to use the same statistics for dichotomous and nominal variables. However, a statistic such as the mean or average, which would be meaningless for a three or more category nominal variable (e.g., *ethnicity*), does have meaning when there are only two categories. For example, in the HSB data, the average *gender* is .55 (with *males* = 0 and *females* = 1). This means that 55% of the participants were *females*, the higher code. Furthermore, we will see with multiple regression that dichotomous variables, called *dummy variables*, can be used as independent variables along with other variables that are normally distributed.

Other terms for dichotomous variables. In the SPSS Variable View (e.g., see Fig 1.2), we label dichotomous variables **nominal**, and this is common in textbooks. However, please remember that dichotomous variables are really a special case and for some purposes they can be treated as if they were

normally distributed. Note that dichotomous data have two discrete categories, so these variables are sometimes called **discrete variables** or **categorical variables**.

Ordinal Variables

In ordinal measurement, there are not only mutually exclusive categories as in nominal scales, but the categories are ordered from low to high, such that ranks could be assigned (e.g., 1st, 2nd, 3rd). Thus, in an ordinal scale, one knows which participant is highest or most preferred on a dimension, but the intervals between the various categories are not equal. Often, whether or not the intervals between categories are equal is a matter of judgment. Our definition of **ordinal** focuses on whether the frequency counts for each category or value are distributed like the bell-shaped, normal curve with more responses in the middle categories and fewer in the lowest and highest categories. If not approximately normal, we would call the variable ordinal. Ordered variables with only a few categories (say 2-4) would also be called ordinal. As indicated in Table 1.3, however, the traditional definition of ordinal focuses on whether the differences between pairs of levels are equal.

Other terms for ordinal variables. Some authors use the term **ranks** interchangeably with ordinal. However, most analyses that are designed for use with ordinal data (nonparametric tests) rank the data as a part of the procedure, assuming that the data you are entering are not already ranked. Moreover, the process of ranking changes the distribution of data such that it can be used in many analyses ordinarily requiring normally distributed data. Ordinal data can be **categorical**, and so that term is sometimes used to include both nominal and ordinal data.

Approximately Normal (or Scale) Variables

Approximately normally distributed variables not only have levels or scores that are *ordered* from low to high, but also, as stated in Table 1.3, the frequencies of the scores are approximately normally distributed. That is, most scores are somewhere in the middle with similar smaller numbers of low and high scores. Thus, a 5-point Likert scale, such as strongly agree to strongly disagree, would be considered normal if the frequency distribution was approximately normal. We think normality, because it is an assumption of many statistics, should be the focus of this highest level of measurement. Many normal variables are continuous (i.e., they have an infinite number of possible values within some range). If not continuous, we suggest that there be at least five ordered values or levels and that they have an implicit, underlying continuous nature. For example, a 5-point Likert scale has only five response categories but, in theory, a person's rating could fall anywhere between 1 and 5.

Other terms for approximately normal variables. **Continuous**, **dimensional**, and **quantitative** are some terms that you will see in the literature for variables that vary from low to high and are assumed to be normally distributed. SPSS uses **scale**, as previously noted.

Traditional measurement terminology uses the terms **interval** and **ratio**. SPSS does not use these terms, but because they are common in the literature and overlap with the term **scale**, we will describe them briefly. **Interval** variables have ordered categories that are equally spaced (i.e., have equal intervals between them). Most physical measurements (*length*, *weight*, *temperature*, etc.) have equal intervals between them. Many physical measurements (*length* and *weight*), in fact, not only have equal intervals between the levels or scores, but also a true zero, which means in the above examples, zero length or weight. Such variables are called **ratio** variables. Our Fahrenheit temperature scale and almost all psychological scales do not have a true zero and, thus, even if they are very well constructed equal interval scales, it is not possible to say that zero degrees Fahrenheit involves the absence of something or that one has no intelligence or no extroversion or no attitude of a certain type. The differences between interval scales and ratio scales are not important for us because we can do all of the types of statistics that

we have available with interval data. SPSS terminology supports this non-distinction by using the term **scale** for both interval and ratio data. In fact, the more important thing, because it is an assumption of most parametric statistics, is that the variables be approximately normally distributed, not whether they have equal intervals.

How to Distinguish Between the Types of Measurement

Distinguishing between nominal and ordered variables. When you label variables in SPSS, the **Measure** column (see Fig. 1.2) provides only three choices: nominal, ordinal, or scale. How do you decide? We suggest that *if the variable* has only two levels, you call it nominal even though it is often hard to tell whether such a *dichotomous* variable (e.g., Yes or No, Pass or Fail) is unordered or ordered. Although some such dichotomous variables are clearly nominal (e.g., *gender*) and others are clearly ordered (e.g., *math grades*--high and low), all dichotomous variables form a special case, as discussed above.

If there are three or more categories, it is usually fairly easy to tell whether the categories are ordered or not, so students and researchers should be able to distinguish between nominal and ordinal data. That is good because this distinction makes a lot of difference in choosing and interpreting appropriate statistics.

Distinguishing between ordinal and normal variables. Is a 5-point Likert scale ordinal or approximately normal? Using our definitions, there is a way to test variables to see whether it is more reasonable to treat them as *normal* variables or *ordinal* variables. Unfortunately, you will not know for sure until after the data have been collected and preliminary analyses are done.

Table 1.4 provides a summary of the characteristics and examples of our four types of measurement. It should provide a good review of the concept of type or level of measurement of a variable.

Table 1.4. Characteristics and Examples of the Four Types of Measurement

	Nominal	Dichotomous	Ordinal	Normal
Characteristics	3+ levels Not ordered True categories Names, labels	2 levels Ordered or not	3+ levels Ordered levels Unequal intervals between levels Not normally distributed	5+ levels Ordered levels Approximately normally distributed Equal intervals between levels
Examples	Ethnicity Religion Curriculum type Hair color	Gender Math grades (high vs. low)	Competence scale Mother's education	SAT math Math achievement Height

Remember that in SPSS, there are only three measurement types or levels, and the researcher is the one who determines if the variable is called nominal, ordinal, or scale. We have called dichotomous variables nominal in our hsbdata file.

Descriptive Statistics

Measures of Central Tendency

In chapter 2, we will obtain descriptive statistics to summarize data and check assumptions; we will also produce plots that visually display distributions of variables. Here we review the meaning of descriptive statistics.

Three measures of the center of a distribution are commonly used: **mean**, **median**, and **mode**. Any of them can be used with normally distributed data; however, with ordinal data, the mean of the raw scores is usually not appropriate. Especially if one is computing certain statistics, the mean of the *ranked* scores of ordinal data provides useful information. With nominal data, the mode is the only appropriate measure.

Mean. The arithmetic average or mean takes into account all of the available information in computing the central tendency of a frequency distribution. Thus, it is usually the statistic of choice, assuming that the data are normally distributed data. The mean is computed by adding up all the raw scores and dividing by the number of scores ($M = \Sigma X/N$).

Median. The middle score or median is the appropriate measure of central tendency for ordinal level raw data. The median is a better measure of central tendency than the mean when the frequency distribution is skewed. For example, the median income of 100 mid-level workers and one millionaire reflects the central tendency of the group better (and is substantially lower) than the mean income. The average or mean would be inflated in this example by the income of the one millionaire. For normally distributed data, the median is in the center of the **box and whisker plot**.

Mode. The most common category, or mode can be used with any kind of data but generally provides the least precise information about central tendency. Moreover, if one's data are continuous, there often are multiple modes, none of which truly represents the "typical" score. In fact if there are multiple modes, SPSS provides only the lowest one. One would use the mode as the measure of central tendency if the variable is nominal or if you want a quick non-calculated measure. The mode is the tallest bar in a bar graph or histogram.

You also can compute the **Mean**, **Median**, and **Mode**, plus other descriptive statistics with SPSS by using the Frequencies command.

To get Fig 1.6, select:

- **Analyze => Descriptive Statistics => Frequencies => *scholastic aptitude test – math* => Statistics
=> Mean, Median, and Mode => Continue => OK.**

Note in Fig. 1.6 that the mean and median are very similar, which is in agreement with our conclusion from Fig. 1.4 that *SATM* is approximately normally distributed. Note that the mode is 500, as shown in Fig. 1.4 by the tallest bars.

Statistics		
scholastic aptitude test - math		
N	Valid	75
	Missing	0
Mean		490.53
Median		490.00
Mode		500

Fig. 1.6. Central tendency measures using the SPSS frequencies command.

Measures of Variability

Variability tells us about the spread or dispersion of the scores. At one extreme, if all of the scores in a distribution are the same, there is no variability. If the scores are all different and widely spaced apart, the variability will be high. The **range** (highest minus lowest score) is the crudest measure of variability but does give an indication of the spread in scores if they are ordered.

Standard Deviation. This common measure of variability, is most appropriate when one has normally distributed data, although the mean of ranked ordinal data may be useful in some cases. The standard deviation is based on the deviation of each score (x) from the mean of all the scores (M). Those deviation scores are squared and then summed ($\Sigma(x-M)^2$). This sum is divided by $N-1$, and, finally, the square root is taken ($SD = \sqrt{\Sigma(x-M)^2/N-1}$).

Interquartile range. For ordinal data, the interquartile range, seen in the **box plot** as the distance between the top and bottom of the box, is a very useful measure of variability. Note that the whiskers indicate the expected range, and scores outside that range are shown as outliers.

With nominal data, none of the above variability measures (range, standard deviation, or interquartile range) are appropriate. Instead, for nominal data, one would need to ask how many different categories there are and what the percentages or frequency counts are in each category to get some idea of variability. Minimum and maximum frequency may provide some indication of distribution as well.

Conclusions About Measurement and the Use of Statistics

Statistics based on means and standard deviations are valid for normally distributed or **normal** data. Typically, these data are used in the most powerful tests called **parametric** statistics. However, if the data are ordered but grossly non-normal (i.e., **ordinal**), means and standard deviations may not give meaningful answers. Then the median and a **nonparametric** test would be preferred. Nonparametric tests typically have somewhat less **power** than parametric tests (they are less able to demonstrate truly significant effects), but the sacrifice in power for nonparametric tests based on ranks usually is relatively minor. If the data are **nominal**, one would have to use the mode or counts. In this case, there would be a major sacrifice in power.

Table 1.5 summarizes much of the above information about the appropriate use of various kinds of descriptive statistics given nominal, dichotomous, ordinal, or normal data.

Table 1.5. Selection of Appropriate Descriptive Statistics and Plots for Levels of Measurement

Plots	Nominal	Dichotomous	Ordinal	Normal
Frequency Distribution	Yes ^a	Yes	Yes	OK ^b
Bar Chart	Yes	Yes	Yes	OK
Histogram	No ^c	No	OK	Yes
Frequency Polygon	No	No	OK	Yes
Box and Whiskers Plot	No	No	Yes	Yes
Central Tendency				
Mean	No	OK	Of ranks, OK	Yes
Median	No	OK = Mode	Yes	OK
Mode	Yes	Yes	OK	OK
Variability				
Range	No	Always 1	Yes	Yes
Standard Deviation	No	No	Of ranks, OK	Yes
Interquartile range	No	No	OK	OK
How many categories	Yes	Always 2	OK	Not if truly continuous
Shape				
Skewness	No	No	Yes	Yes

^aYes means a good choice with this level of measurement.

^bOK means OK to use, but not the best choice at this level of measurement.

^cNo means not appropriate at this level of measurement.

The Normal Curve

Figure 1.7 is an example of a normal curve. The frequency distributions of many of the variables used in the behavioral sciences are distributed approximately as a normal curve. Examples of such variables that approximately fit a normal curve are height, weight, intelligence, and many personality variables. Notice that for each of these examples, most people would fall toward the middle of the curve, with fewer people at the extremes. If the average height of men in the United States were 5'10", then this height would be in the middle of the curve. The heights of men who are taller than 5'10" would be to the right of the middle on the curve, and those of men who are shorter than 5'10" would be to the left of the middle on the curve, with only a few men 7 feet or 5 feet tall.

The normal curve can be thought of as derived from a frequency distribution. It is theoretically formed from counting an "infinite" number of occurrences of a variable. Usually when the normal curve is depicted, only the X axis (horizontal) is shown. To determine how a frequency distribution is obtained, you could take a fair coin, and flip it 10 times, and record the number of heads on this first set or trial. Then flip it another 10 times and record the number of heads. If you had nothing better to do, you could do 100 trials. After performing this task, you could plot the number of times that the coin turned up heads out of each trial of 10. What would you expect? Of course, the largest number of trials probably would show 5 heads out of 10. There would be very few, if any trials, where 0, 1, 9, or 10 heads occur. It could happen, but the probability is quite low, which brings us to a probability distribution. If we performed this experiment 100 times, or 1,000 times, or 1,000,000 times, the frequency distribution would "fill in" and look more and more like a normal curve.

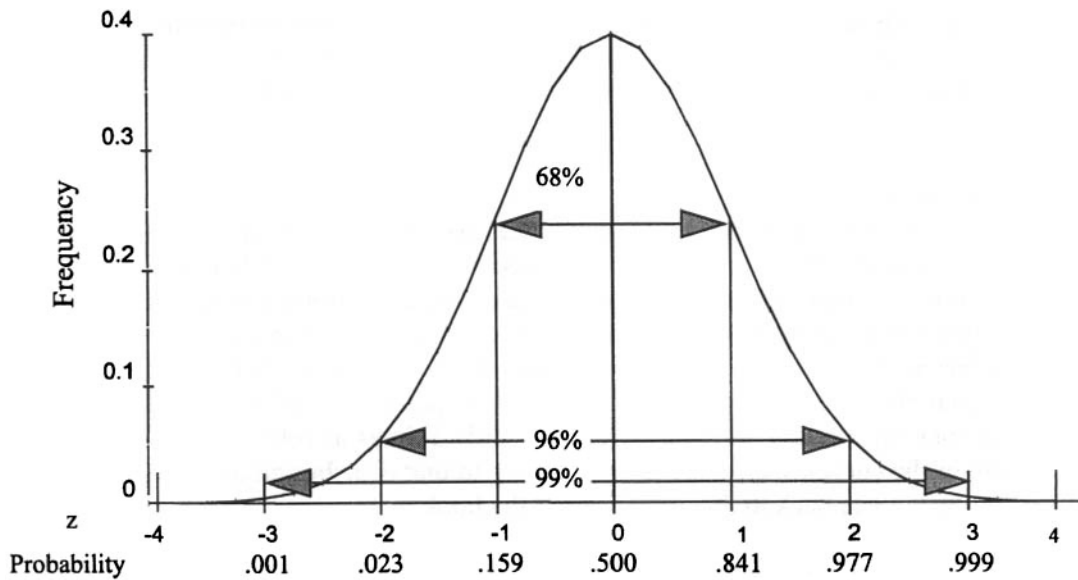


Fig. 1.7. Frequency distribution and probability distribution for the normal curve.

Properties of the Normal Curve

The normal curve has five properties that are always present.

1. The normal curve is unimodal. It has one “hump,” and this hump is in the middle of the distribution. The most frequent value is in the middle.
2. The mean, median, and mode are equal.
3. The curve is symmetric. If you fold the normal curve in half, the right side would fit perfectly with the left side; that is, it is not **skewed**.
4. The range is infinite. This means that the extremes approach but never touch the X axis.
5. The curve is neither too peaked nor too flat and its tails are neither too short nor too long; it has no **kurtosis**.

Non-Normally Shaped Distributions

Skewness. If one tail of a frequency distribution is longer than the other, and if the mean and median are different, the curve is skewed. Because most common inferential statistics (e.g., *t* test) assume that the dependent variable is normally distributed, it is important that we know if our variables are highly skewed.

Figure 1.5 showed a frequency distribution that is skewed to the left. This is called a negative skew. A perfectly normal curve has a skewness of zero (0.0). The curve in Fig. 1.5, for the *competence* scale, has a skewness statistic of -1.63, which indicates that the curve is quite different from a normal curve. We will use a somewhat arbitrary guideline that if the skewness is more than +1.0 or less than -1.0, the distribution is markedly skewed and it would be prudent to use a nonparametric (ordinal type) statistic. However, some parametric statistics, such as the two-tailed *t* test and ANOVA, are quite robust so even a skewness of more than +/-1 may not change the results much. We will provide more examples and discuss this more in chapter 2.

Kurtosis. If a frequency distribution is more peaked than the normal curve, it is said to have positive kurtosis and is called leptokurtic. Note in Fig. 1.4 that the *SAT-math* histogram is peaked (i.e., the bar for 500 extends above the normal curve line), and thus there is some positive kurtosis. If a frequency

distribution is relatively flat with heavy tails, it has negative kurtosis and is called platykurtic. Although SPSS can easily compute a kurtosis value for any variable using an option in the Frequencies command, usually we will not do so because kurtosis does not seem to affect the results of most statistical analyses very much.

Areas Under the Normal Curve

The normal curve is also a probability distribution. Visualize that the area under the normal curve is equal to 1.0. Therefore, portions of this curve could be expressed as fractions of 1.0. For example, if we assume that 5'10" is the average height of men in the United States, then the probability of a man being 5'10" or taller is .5. The probability of a man being over 6'3" or less than 5'5" is considerably smaller. It is important to be able to conceptualize the normal curve as a probability distribution because statistical convention sets acceptable probability levels for rejecting the null hypothesis at .05 or .01. As we shall see, when events or outcomes happen very infrequently, that is, only 5 times in 100 or 1 time in 100 (way out in the left or right tail of the curve), we wonder if they belong to that distribution or perhaps to a different distribution. We will come back to this point later in the book.

All normal curves can be divided into areas or units in terms of the standard deviation. Approximately 34% of the area under the normal curve is between the mean and one standard deviation above or below the mean (see Fig. 1.7 again). If we include both the area to the right *and* to the left of the mean, 68% of the area under the normal curve is within one standard deviation from the mean. Another approximately 13.5% of the area under the normal curve is accounted for by adding a second standard deviation to the first standard deviation. In other words, two standard deviations to the right of the mean account for an area of approximately 47.5%, and two standard deviations to the left *and* right of the mean make up an area of approximately 95% of the normal curve. If we were to subtract 95% from 100%, the remaining 5% relates to that ever present probability or *p* value of 0.05 needed for statistical significance. Values not falling within two standard deviations of the mean are seen as relatively rare events.

The Standard Normal Curve

All normal curves can be converted into standard normal curves by setting the mean equal to zero and the standard deviation equal to one. Since all normal curves have the same proportion of the curve within one standard deviation, two standard deviations, and so on, of the mean, this conversion allows comparisons among normal curves with different means and standard deviations. The normal distribution, has the standard normal distribution units underneath. These units are referred to as *z* scores. If you examine the normal curve table in any statistics book, you can find the areas under the curve for one standard deviation ($z = 1$), two standard deviations ($z = 2$), and so on. As described in the Quick Reference Guide (Appendix A), it is easy for SPSS to convert raw scores into *standard scores*. This is often done when one wants to aggregate or add together several scores that have quite different means and standard deviations.

Interpretation Questions

- 1.1 What is the difference between the independent variable and the dependent variable?
- 1.2 Compare the terms *active independent variable* and *attribute independent variable*. What are the similarities and differences?

Chapter 1 - Introduction

- 1.3 What kind of independent variable is necessary to infer cause? Can one always infer cause from this type of independent variable? If so, why? If not, when can one clearly infer cause and when might causal inferences be more questionable?
- 1.4 Compare and contrast associational, difference, and descriptive types of research questions.
- 1.5 Write three research questions and a corresponding hypothesis regarding variables of interest to you but not in the HSB data set (one associational, one difference, and one descriptive question).
- 1.6 Using one or more of the following HSB variables, *religion*, *mosaic score*, and *visualization score*:
 - a) Write an associational question
 - b) Write a difference question
 - c) Write a descriptive question
- 1.7 If you have categorical, ordered data (such as low income, middle income, high income) what type of measurement would you have? Why?
- 1.8
 - a) What are the differences between nominal, dichotomous, ordinal, and normal variables?
 - b) In social science research, why isn't it important to distinguish between interval and ratio variables?
- 1.9 What percent of the area under the standard normal curve is between the mean and one standard deviation above the mean?
- 1.10
 - a) How do z scores relate to the normal curve?
 - b) How would you interpret a z score of -3.0 ?

CHAPTER 2

Data Coding and Exploratory Analysis (EDA)

Before computing any inferential statistics, it is necessary to code the data, enter the data into SPSS, and then do exploratory data analysis (EDA) as outlined below. This chapter will help you understand your data, help you to see if there are any errors, and help you to know if your data meet basic assumptions for the inferential statistics that you will compute.

Rules for Data Coding

Coding is the process of assigning numbers to the values or levels of each variable. We want to present some broad suggestions or rules to keep in mind as you proceed. These suggestions are adapted from rules proposed in Newton and Rudestam's (1999) useful book entitled *Your Statistical Consultant*. It is important to note that the recommendations we make are those we find to be most useful in most contexts, but some researchers might propose alternatives, especially for "rules" 1, 2, 4, 5, and 7 below.

1. ***All data should be numeric.*** Even though it is possible to use letters or words (string variables) as data, it is not desirable to do so with SPSS. For example, we could code gender as M for male and F for female, but in order to do most statistics with SPSS, you would have to convert the letters or words to numbers. It is easier to do this conversion before entering the data into the computer. We decided to code females as 1 and males as 0. This is called **dummy coding**. In essence, the 0 means "not female." We could, of course, code males as 1 and females as 0. However, it is crucial that you be consistent in your coding and have a way to remind yourself and others of how you did the coding. In the Quick Reference Guide (Appendix A), we show how you can provide such a record, called a **codebook**.

2. ***Each variable for each case or participant must occupy the same column in the SPSS Data Editor.*** With SPSS it is important that data from each participant occupies only one line (row), and each column must contain data on the same variable for all the participants. The SPSS data editor, into which you will enter data, facilitates this by putting the variable names that you choose at the top of each column, as you saw in Fig. 1.3. If a variable is measured more than once for each participant (e.g., pretest and posttest), it needs to be entered on the same row in separate columns with somewhat different names like *mathpre* and *mathpost*.

3. ***All values (codes) for a variable must be mutually exclusive.*** That is, only one value or number can be recorded for each variable. Some items may allow participants to check more than one response. In that case, the item should be divided into a separate variable for each possible response choice, with one value of each variable (usually "1") corresponding to yes (checked) and the other (usually "0") to no (not checked).

Usually, items should be phrased so that persons would logically choose only one of the provided options, and all possible options are provided. A final category labeled "other" may be provided in cases where all possible options cannot be listed, but these "other" responses are usually quite diverse and, thus, usually not very useful for statistical purposes.

4. ***Each variable should be coded to obtain maximum information.*** Do not collapse categories or values when you set up the codes for them. If needed, let the computer do it later. In general, it is desirable to code and enter data in as detailed a form as available. Thus, enter item scores, ages, GPAs, etc. for each participant if you know them. It is good to practice asking participants to provide information that is quite specific. However, you should be careful not to ask questions that are so specific that the respondent may

not know the answer or may not feel comfortable providing it. For example, you will obtain more specific information by asking participants to state their GPA to two decimals than if you asked them to select from a few broad categories (e.g., less than 2.0, 2.0-2.49, 2.50-2.99, etc.). However, if students don't know their exact GPA or don't want to reveal it precisely, they may leave the question blank, guess, or write in a difficult to interpret answer.

These issues might lead you to provide a few categories, each with a relatively narrow range of values, for variables such as age, weight, and income. Never collapse such categories before you enter the data into SPSS. For example, if you had age categories for university undergraduates 16-18, 18-20, 21-23, and so on and you realize that there are only a few students in the below 18 group, keep the codes as they are for now. Later you can make a new category of 20 or under by using an SPSS function, **Transform => Recode**. If you collapse categories before you enter the data, the information is lost.

5. For each participant, there must be a code or value for each variable. These codes should be numbers, except for variables for which the data are missing. We recommend using blanks when data are missing or unusable, because SPSS is designed to handle blanks as missing values. However, sometimes you may have more than one type of missing data, such as items left blank *and* those that had an answer that was not appropriate or usable. In this case, you may assign numeric codes such as 98 and 99 to them, but you must tell SPSS that these codes are for missing values, or SPSS will treat them as actual data.

6. Apply any coding rules consistently for all participants. This means that if you decide to treat a certain type of response as, say, missing for one person, you must do the same for all other participants.

7. Use high numbers (value or codes) for the “agree,” “good,” or “positive” end of a variable that is ordered. However, if the variable has a negative label (e.g., aggressiveness), higher numbers should refer to more of the trait. Sometimes you will see questionnaires that use 1 for “strongly agree,” and 5 for “strongly disagree.” This is not wrong as long as it is clear and consistent. However, you are less likely to get confused when interpreting your results if high values have positive meaning, indicate that something was done (e.g., an algebra 1 course was taken), or indicate more of the characteristic.

Make a coding form and/or codebook. Now you need to make some decisions about how to code the data, especially data that are not already in numerical form. When the responses provided by participants are numbers, the variable is said to be “self-coding.” You can just enter the number that was circled or checked on the questionnaire. On the other hand, variables such as *gender* or *ethnicity* have no intrinsic values associated with them, so a number has to be assigned to each level or value.

Fix problems with the completed questionnaires. Now examine the questionnaires (or other new data) for incomplete, unclear, or double answers. The researcher needs to use rules to handle these problems and note the decision on the questionnaires or on a master “coding instructions” sheet or file so that the same rules are used for all cases. For each type of incomplete, blank, unclear, or double answer, you need to make a rule for what to do. As much as possible, you should make these rules before data collection, but there may well be some unanticipated issues. It is important that you apply the rules consistently for all similar problems so as not to bias your results.

Missing data create problems in later data analysis, especially for complex statistics. Thus, we want to use as much of the data provided as is reasonable. The important thing here is that you *must* treat all similar problems the same way. If a participant only answered some of the questions, there will be lots of missing data for that person. We could have a rule such as if half the items were blank or invalid, we would throw out that whole questionnaire as invalid. In your research report, you should state how many questionnaires were thrown out and for what reason(s). If a participant circled two responses (e.g., 3 and 4 on a 5-point Likert scale), a reasonable decision would be to enter the average or midpoint, 3.50.

Clean up completed questionnaires. Once you have made your rules and decided how to handle each problem, you need to make these rules clear to the person entering the data. A common procedure would be to write your decisions on the questionnaires, perhaps in a different color. You also need to have a master file and hard copy of all the rules that you used.

In the process of understanding your data, different types of analyses and plots will be generated depending on what level of measurement you have. Therefore, it is important to identify whether each of your variables is **nominal**, **dichotomous**, **ordinal**, or **normal** (SPSS uses the term **scale**; see chapter 1). Keep in mind that there are times when whether you call a variable ordinal or scale might change based on your EDA. For example, a variable that you considered to be ordinal may be normally distributed, and, thus, better labeled as scale. Remember that making the appropriate choice indicates that you understand your data and should help guide your selection of a statistic.

Exploratory Data Analysis (EDA)

What Is EDA?

After the data are entered into SPSS, the first step to complete (before running any inferential statistics) is EDA, which involves computing various descriptive statistics and graphs. **Exploratory Data Analysis** is used to examine and get to know your data. Chapter 1 and especially this chapter focus on ways to do exploratory data analysis with SPSS. EDA is important to do for several reasons:

1. To see if there are problems in the data such as outliers, non-normal distributions, problems with coding, missing values, and/or errors inputting the data.
2. To examine the extent to which the assumptions of the statistics that you plan to use are met.

In addition to these two reasons which are discussed in this chapter, one could also do EDA for other purposes such as:

3. To get basic information regarding the demographics of subjects to report in the Method section or Results section.
4. To examine relationships between variables to determine how to conduct the hypothesis-testing analyses. For example, correlations can be used to see if two or more variables are so highly related that they should be combined (aggregated) for further analyses and/or if only one of them should be included in the central analyses. We created *parents' education* by combining *father's* and *mother's education*, because they are quite highly correlated.

How to Do EDA

There are two general methods used for EDA: generating plots of the data and generating numbers from your data. Both are important and can be very helpful methods of investigating the data. Descriptive Statistics (including the minimum, maximum, mean, standard deviation, and skewness), frequency distribution tables, boxplots, histograms, and stem and leaf plots are a few procedures used in EDA.

After collecting data and inputting it into SPSS, many students jump immediately to doing inferential statistics (e.g., *t* tests and ANOVAs). Don't do this! Many times there are errors or problems with the data that need to be located and either fixed or at least noted before doing any inferential statistics.

Chapter 2 – Data Coding and EDA

At this point, you are probably asking “Why?” or “I’ll do that boring descriptive stuff later while I am writing the methods section.” Wait! Being patient can alleviate many problems down the road.

In the next two sections, we discuss checking for errors and checking assumptions. Some of this discussion reviews basic material, but it is so important that it is worth going over again.

Check for Errors

There are many ways to check for errors; for example:

1. As mentioned above, look over the raw data (questionnaires, interviews, or observation forms) to see if there are inconsistencies, double coding, obvious errors, etc. Do this before entering the data into the computer.
2. Check some, or preferably all, of the raw data (e.g., questionnaires) against the data in your **SPSS Data Editor** file to be sure that errors were not made in the data entry.
3. Compare the minimum and maximum values for each variable in your Descriptives output with the allowable range of values in your codebook.
4. Examine the means and standard deviations to see if they look reasonable, given what you know about the variables.
5. Examine the *N* column to see if any variables have a lot of missing data, which can be a problem when you do statistics with two or more variables. Missing data could also indicate that there was a problem in data entry.
6. Look for outliers in the data.

Check the Assumptions

As noted above, exploratory data analysis can be used to check the assumptions of a statistic. Several assumptions are common to more than one statistic, so in this chapter we will provide an introduction to how to test for them. First, we will define statistical assumptions and briefly discuss several of the most common.

Statistical Assumptions

Every statistical test has assumptions. Statistical assumptions are much like the directions for appropriate use of a product found in an owner’s manual. **Assumptions explain when it is and isn’t reasonable to perform a specific statistical test.** When the *t* test was developed, for example, the person who developed it needed to make certain assumptions about the distribution of scores, etc., in order to be able to calculate the statistic accurately. If these assumptions are not met, the value that SPSS calculates, which tells the researcher whether or not the results are statistically significant, will not be completely accurate and may even lead the researcher to draw the wrong conclusion about the results. In each chapter, the appropriate inferential statistics and their assumptions are described.

Parametric tests. These include most of the familiar ones (e.g., *t* test, analysis of variance, Pearson correlation, and almost all of the statistics discussed in chapters 4-10). They usually have more assumptions than nonparametric tests. Parametric tests were designed for data that have certain characteristics, including approximately normal distributions.

Some parametric statistics have been found to be “robust” to one or more of their assumptions. **Robust** means that the assumption can be violated without damaging the validity of the statistic. For example, one assumption of ANOVA is that the dependent variable is normally distributed for each group. Statisticians who have studied these statistics have found that even when data are not completely normally distributed (e.g., they are somewhat skewed), they still can be used.

Nonparametric tests. These tests (e.g., chi-square, Mann-Whitney U, Spearman rho) have fewer assumptions and often can be used when the assumptions of a parametric test are violated. For example, they do not require normal distributions of variables or homogeneity of variances. Unfortunately, there are few nonparametric tests similar to the intermediate statistics discussed in the book so we will have little to say about them here.

Common Assumptions

Homogeneity of variances. Both the *t* test and ANOVA may be affected quite a lot if the variances (standard deviation squared) of the groups to be compared are substantially different, especially if the number of participants in each group differs markedly. Thus, this is often a critical assumption to meet or correct for. Fortunately, SPSS provides the Levene test to check this assumption.

Normality. As mentioned above, many parametric statistics assume that certain variables are distributed approximately normally. That is, the frequency distribution would look like a symmetrical bell-shaped or normal curve, with most subjects having values in the mid range and with a smaller number of subjects with high and low scores. A distribution that is asymmetrical, with more high than low scores (or vice versa), is **skewed**. Thus, it is important to check skewness. Most statistics books do not provide advice about how to decide whether a variable is at least approximately normal. SPSS recommends that you divide the skewness by its standard error. If the result is less than 2.5 (which is approximately the $p = .01$ level), then skewness is *not* significantly different from normal. A problem with this method, aside from having to use a calculator, is that the standard error depends on the sample size, so with large samples most variables would be found to be nonnormal. A simpler guideline is that if the skewness is less than plus or minus one ($< \pm 1.0$), the variable is at least approximately normal. There are also several other ways to check for normality. In this chapter we will look at two graphical methods: boxplots and frequency polygons. However, remember that *t* (if two-tailed) and ANOVA are quite robust to violations of normality.

Independence of observations. The assumption of independence of observations is that there is no relationship between the scores for one person and those of another person. For example, if you know one subject's value on one variable (e.g., competence), then you should not be able to guess the value of that variable for any other particular participant. Sometimes, this assumption is violated because one's procedures for sampling participants create systematic bias. For example, “snowball sampling,” in which participants recommend other participants for the study, is likely to lead to nonindependence of observations because participants are likely to recommend people who are similar to themselves. Obviously, members of the same family, or the same person measured on more than one occasion, do not comprise independent observations. There are particular methods (matched samples or “repeated measures” methods) designed to deal with the nonindependence of family members or the same person measured several times, or participants who are matched on some characteristic. However, these can not easily be applied to snowball sampling because different participants recommend a different number of other participants.

Linearity. Linearity is the assumption that two variables are related in a linear fashion. If variables are linearly related, then when plotted in a scatterplot, the data will fall in straight line or in a cluster that is relatively straight. Sometimes, if the data are not linearly related (i.e., the plot looks curved), the data can be transformed to make the variables linearly related.

Checking for Errors and Assumptions With Ordinal and Scale Variables

The level of measurement of a variable you are exploring (whether it is nominal, ordinal, dichotomous, or normal/scale) influences the type of exploratory data analysis (EDA) you will want to do. Thus, we have divided this chapter by the measurement levels of the variable because, for some types of variables, certain descriptive statistics will not make sense (e.g., a mean for a nominal variable, or a boxplot for a dichotomous variable). Remember that the researcher has labeled the type of measurement as either nominal, ordinal, or scale when completing the **SPSS Data Editor Variable View**. Remember also that we decided to label dichotomous variables **nominal**, and variables that we assumed were normally distributed were labeled **scale**.

For all of the examples in this chapter, we will be using the `hsbdataB` file, which is on the CD in the back of this book. Because this chapter is meant to be an overview and review of EDA, we have formatted the chapter to be read rather than as a series of problems for you to solve, as is done in chapters 4-10. However, we have provided the syntax and parts of the SPSS outputs if you decide that you want to actually run these problems with SPSS.

See Appendix B for instructions if you need help with this or getting started with SPSS. Appendix B also shows how to set your computer to print the SPSS syntax on the output.

Using all of the HSB variables that were labeled as ordinal or scale in the SPSS Variable View, it is important to see if the means make sense (are they close to what you expected?), to examine the minimum and maximum values of the data, and to check the shape of the distribution (i.e., skewness value). One way to check these is with the SPSS **Descriptive** command. Examining your data to see if the variables are approximately normally distributed, an assumption of most of the parametric inferential statistics that we will use, is important. To understand if a variable is normally distributed, we compute the skewness index, which helps determine how much a variable's distribution deviates from the distribution of the normal curve. **Skewness** refers to the lack of symmetry in a frequency distribution. Distributions with a long "tail" to the right have a positive skew and those with a long tail on the left have a negative skew. If a frequency distribution of a variable has a large (plus or minus) skewness value, that variable is said to deviate from normality. Some of the statistics that we will use later in the book are robust or quite insensitive to violations of normality. Thus, we will assume that it is okay to use them to answer most of our research questions as long as the variables are not extremely skewed.

The **Descriptives** command will make a compact, space-efficient output. You could instead run the **Frequencies** program because you can get the same statistics with that command. (We will use the **Frequencies** command later in the chapter.) Now we will compute the mean, standard deviation, skewness, minimum, and maximum for all participants or cases on all the variables that were called ordinal or scale under measure in the **SPSS Data Editor Variable View**. We will not include the nominal variables (*ethnicity* and *religion*) or *gender*, *algebra1*, *algebra2*, *geometry*, *trigonometry*, *calculus*, and *math grades*, which are dichotomous variables.

First, we compute descriptive statistics for all of the variables that were called ordinal. These include *father's education*, *mother's education*, *grades in h.s.*, and all the item variables (*item 01* through *item 14*), *math courses taken*, *mother's education revised*, *father's education revised*, and *parents' education*.

Next, we run **Descriptives** for the scale variables: *math achievement*, *mosaic*, *visualization*, *visualization retest*, *scholastic aptitude test-math*, *competence*, *motivation*, and *pleasure scale*.

SPSS for Intermediate Statistics

Output 2.1a: Descriptives for the Ordinal Variables

DESCRIPTIVES

```
VARIABLES=faed maed grades item01 item02 item03 item04 item05 item06 item07 it
em08 item09 item10 item11 item12 item13
item14 mathcrs maedRevis faedRevis parEduc
/STATISTICS=MEAN STDDEV MIN MAX SKEWNESS .
```

Syntax or log file shows
the variables and statistics
that you requested.

Descriptive Statistics

						Skewness	
	N	Minimum	Maximum	Mean	Std. Deviation	Statistic	Std. Error
father's education	73	2	10	4.73	2.830	.684	.281
mother's education	75	2	10	4.11	2.240	1.124	.277
grades in h.s.	75	2	8	5.68	1.570	-.332	.277
item01 motivation	74	1	4	2.96	.928	-.763	.279
item02 pleasure	75	1	4	3.52	.906	-1.910	.277
item03 competence	74	1	4	2.82	.897	-.579	.279
item04 low motiv	74	1	4	2.16	.922	.422	.279
item05 low comp	75	1	4	1.61	.971	1.581	.277
item06 low pleas	75	1	4	2.43	.975	-.058	.277
item07 motivation	75	1	4	2.76	1.051	-.433	.277
item08 low motiv	75	1	4	1.95	.914	.653	.277
item09 competence	74	1	4	3.32	.760	-1.204	.279
item10 low pleas	75	1	4	1.41	.737	1.869	.277
item11 low comp	75	1	4	1.36	.747	2.497	.277
item12 motivation	75	1	4	3.00	.822	-.600	.277
item13 motivation	75	1	4	2.67	.794	-.320	.277
item14 pleasure	75	1	4	2.84	.717	-.429	.277
math courses taken	75	0	5	2.11	1.673	.325	.277
father's educ revised	73	1.00	3.00	1.7397	.85028	1.162	.277
mother's educ revised	75	1.00	3.00	1.4667	.68445	.533	.281
parents' education	75	2.00	10.00	4.3933	2.31665	.923	.277
Valid N (listwise)	69						

Output 2.1b: Descriptives for Variables Labeled as Scale

DESCRIPTIVES

```
VARIABLES=mathach mosaic visual visual2 satm competence motivation pleasure
/STATISTICS=MEAN STDDEV MIN MAX .
```

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness	
						Statistic	Std. Error
math achievement test	75	-1.67	23.67	12.5645	6.67031	.044	.277
mosaic, pattern test	75	-4.0	56.0	27.413	9.5738	.529	.277
visualization test	75	-.25	14.75	5.2433	3.91203	.536	.277
visualization retest	75	.00	9.50	4.5467	3.01816	.235	.277
scholastic aptitude test - math	75	250	730	490.53	94.553	.128	.277
competence scale	73	1.00	4.00	3.2945	.66450	-1.634	.281
motivation scale	73	1.17	4.00	2.8744	.63815	-.570	.281
pleasure scale	75	1.50	4.00	3.1300	.60454	-.682	.277
Valid N (listwise)	71						

Interpretation of Outputs 2.1a and 2.1b

These outputs provide descriptive statistics for all of the variables labeled as **ordinal** (2.1a) and **scale** (2.1b). Notice that the variables are listed down the left column of the outputs and the requested descriptive statistics are listed across the top row. The descriptive statistics included in the output are the number of subjects (N), the **Minimum** (lowest) and **Maximum** (highest) scores, the **Mean** (or average) for each variable, the **Std.** (the standard deviation), and the **Skewness** statistic and the **Std. Error** of the skewness. Note, from the bottom line of the outputs, that the **Valid N (listwise)** is 69 for Output 2.1a and 71 for 2.1b rather than 75, which is the number of participants in the data file. This is because the listwise *N* only includes the persons with *no* missing data on any variable requested in the output. Notice that several variables (e.g., *father's education* and *item01* in 2.1a, and *motivation* and *competence* in 2.1b) each have one or two participants with missing data.

Using your output to check your data for errors. For both the ordinal and scale variables, check to make sure that all **Means** seem reasonable. That is, you should check your means to see if they are within the ranges you expected (given the information in your codebook and your understanding of the variable). Next, check the output to see that the **Minimum** and **Maximum** are within the appropriate (codebook) range for each variable. If the minimum is smaller or the maximum is bigger than you expected (e.g., 100 for a variable that only has 1–50 for possible values), then you should suspect that there was an error somewhere and you need to check it out. Finally, you should check the *N* column to see if the *N*s are what you were expecting. If it happens that you have more participants missing than you expected, check the original data to see if some were entered incorrectly. Notice that *competence scale* and *motivation scale* each have a few participants missing.

Using the output to check assumptions. The main assumption that we can check from this output is normality. We won't pay much attention to the skewness for *item 01* to *item 14* and *mother's and father's education revised*. These ordinal variables have fewer than five levels, so they will not be considered to be scale even though some of them are not very skewed. We will not use the "items" as individual variables because we will be combining them to create summated variables (the *motivation* and *competence* and *pleasure* scales) before using inferential statistics.

From Output 2.1a, we can see that, of the variables with five or more levels that we called **ordinal**, four of them (*father's education*, *grades in h.s.*, *math courses taken*, and *parent's education*) are approximately normally distributed; that is, they have five or more levels and have skewness values between –1 and 1. Thus, we can assume that they are more like scale variables, and we can use inferential statistics that have the assumption of normality. To better understand these variables, it may be helpful to change the **Measure** column in the **Variable View** so that these four variables will be labeled as scale. Note that *mother's education*, with a skewness statistic of 1.12, is more skewed than is desirable.

For the variables that were labeled as scale, indicating our belief that they were normally distributed, we can now check this assumption. Look at the skewness value in Output 2.1b to see if it is between –1 and 1. From the output we see that most of these variables have skewness values between –1 and 1, but *competence* at 1.63 is quite skewed.

There are several ways to check this assumption in addition to checking the skewness value. If the mean, median, and mode, which can be obtained with the **Frequencies** command, are approximately equal, then you can assume that the distribution is approximately normally distributed. For example, the mean (490.53), median (490.00), and mode (500) for *scholastic aptitude test- math* are very similar values, and the skewness value is .128 (see Output 2.1b). Thus, we can assume that *SAT-math* is approximately normally distributed.

In addition to numerical methods for understanding your data, there are several graphical methods. SPSS can create histograms with the normal curve superimposed and also frequency polygons (line graphs) to roughly assess normality. The trouble is that visual inspection of histograms can be deceiving because some approximately normal distributions don't look very much like a normal curve.

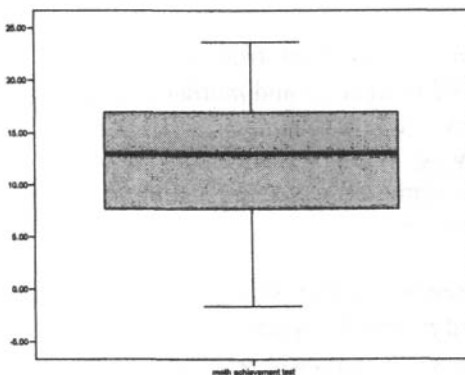
Boxplots of One or Several Variables

Boxplots and **stem-and-leaf plots** can be used to examine some HSB variables. Boxplots are a method of graphically representing ordinal and scale data. They can be made with many different combinations of variables and groups. Using boxplots for one, two, or more variables or groups in the same plot can be useful in helping you understand your data.

Output 2.2a: Boxplot of Math Achievement Test

```
EXAMINE
  VARIABLES=mathach
  /PLOT BOXPLOT
  /COMPARE GROUP
  /STATISTICS NONE
  /CINTERVAL 95
  /MISSING LISTWISE
  /NOTOTAL.
```

Explore

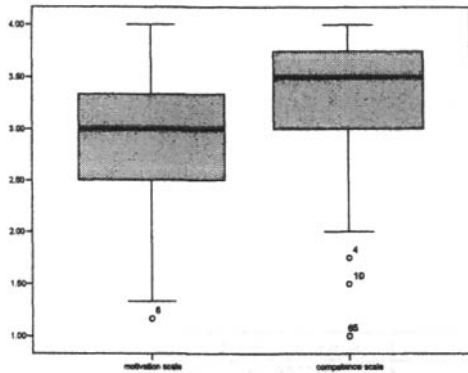


Next we will create a boxplot with more than one variable in the same plot.

Output 2.2b: Boxplots of Competence and Motivation Scales

```
EXAMINE
  VARIABLES=motivation competence
  /PLOT BOXPLOT
  /COMPARE VARIABLES
  /STATISTICS NONE
  /CINTERVAL 95
  /MISSING LISTWISE
  /NOTOTAL.
```

Explore



Notice that there are three outliers for *competence* and one for *motivation* in these boxplots.

Interpretation of Outputs 2.2a and 2.2b

Outputs 2.2a and 2.2b include boxplots generated from the **Explore** command. We did not include the **Case Processing Summary** tables that show the **Valid N**, **Missing** cases, and **Total** cases. In Output 2.2a, for *math achievement*, the valid *N* is 75, and there are no missing cases. The plot in Output 2.2a includes only one boxplot for our requested variable of *math achievement*. Each “box” represents the middle 50% of the cases, and the “whiskers” at the top and bottom of the box indicate the “expected” top and bottom 25%. If there were **outliers** there would be “O”s and if there were really extreme scores they would be shown with asterisks, above or below the end of the whiskers. Notice that there are not any Os or *s in the boxplot in Output 2.2a.

For Output 2.2b, notice that there are two separate boxplots, one for *competence* and one for *motivation*. As indicated by the Os at or below the lower whiskers, the boxplot for *competence* shows there are three outliers, and the boxplot for *motivation* indicates there is one outlier.

Using your output to check your data for errors. If there are Os or asterisks, then you need to check the raw data or score sheet to be sure there was not an error. The numbers next to the Os indicate which participants these scores belong to. This can be helpful when you want to check to see if these are errors or if they are the actual scores of the subject.

Using the output to check your data for assumptions. Boxplots can be useful for identifying variables with extreme scores, which can make the distribution skewed (i.e., nonnormal). Also if there are few outliers, if the whiskers are approximately the same length, and if the line in the box is approximately in the middle of the box, then you can assume that the variable is approximately normally distributed. Thus, *math achievement* (Output 2.2a) is near normal, *motivation* (2.2b) is approximately normal, but *competence* (2.2b) is quite skewed and not normal.

Boxplots Split by a Dichotomous Variable

If you use the syntax below, you will get an output file complete with separate statistics, stem-and-leaf plots, and boxplots for males’ and females’ math achievement scores. This is useful if you want to see if the distributions of scores are very different for the two groups, which would suggest heterogeneity of variances.

Output 2.3: Boxplots Split by Gender With Statistics and Stem-and-Leaf Plots

```

EXAMINE
  VARIABLES=mathach BY gender
  /PLOT BOXPLOT STEMLEAF
  /COMPARE GROUP
  /STATISTICS DESCRIPTIVES
  /CINTERVAL 95
  /MISSING LISTWISE
  /NOTOTAL.

```

Explore

Descriptives

gender			Statistic	Std. Error
math achievement test	male	Mean	14.7550	1.03440
		95% Confidence Interval for Mean	12.6505	
		Lower Bound	16.8595	
		Upper Bound		
		5% Trimmed Mean	14.8454	
		Median	14.3330	
		Variance	36.379	
		Std. Deviation	6.03154	
		Minimum	3.67	
		Maximum	23.7	
		Range	20.0	
		Interquartile Range	10.0005	
		Skewness	-.156	.403
		Kurtosis	-.963	.788
	female	Mean	10.7479	1.04576
		95% Confidence Interval for Mean	8.6344	
		Lower Bound	12.8615	
		Upper Bound		
		5% Trimmed Mean	10.6454	
		Median	10.3330	
		Variance	44.838	
		Std. Deviation	6.69612	
		Minimum	-1.7	
		Maximum	23.7	
		Range	25.3	
		Interquartile Range	10.5000	
		Skewness	.331	.369
		Kurtosis	-.698	.724

Note that we have circled, for males and for females, three key statistics: mean, variance, and skewness.

**math achievement test
Stem-and-Leaf Plots**

math achievement test Stem-and-Leaf Plot for
GEND= male

Frequency	Stem	&	Leaf
1.00	0	.	3
7.00	0	.	5557799
11.00	1	.	01123444444
7.00	1	.	5578899
8.00	2	.	11123333

11 persons (Frequency) had stems of 1 (scores between 10 and 14). One had 10, 2 had 11, etc.

Stem width: 10.0
Each leaf: 1 case(s)

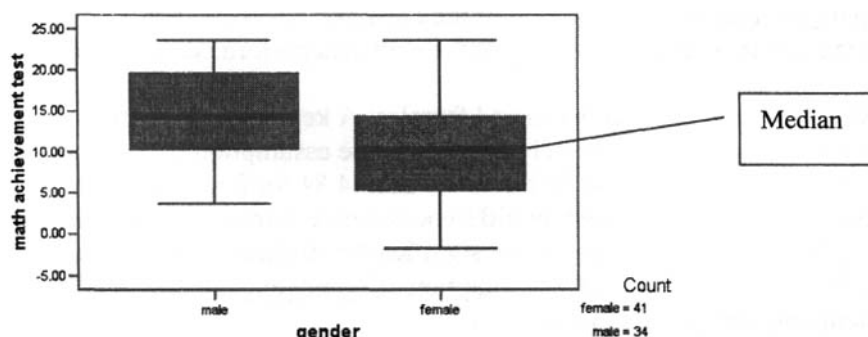
math achievement test Stem-and-Leaf Plot for
GEND= female

Chapter 2 – Data Coding and EDA

Frequency	Stem & Leaf
1.00	-0 . 1
7.00	0 . 1123344
12.00	0 . 555666778999
11.00	1 . 00002334444
5.00	1 . 77779
5.00	2 . 02233

1 person had a negative score (stem = 0) of -1.

Stem width: 10.0
Each leaf: 1 case(s)



Interpretation of Output 2.3

The first table under **Explore** provides descriptive statistics about the number of males and females with **Valid** and **Missing** data. Note that we have 34 males and 41 females with valid *math achievement test* scores.

The **Descriptives** table contains many different statistics for males and females separately, some of which (e.g., kurtosis) are not covered in this book. Note that the average *math achievement test* score is 14.76 for the males and 10.75 for females. We will discuss the variances and skewness below under assumptions.

The **Stem-and-Leaf Plots**, for each gender separately, are next. These plots are like a histogram or frequency distribution turned on the side. They give a visual impression of the distribution, and they usually show *each* person's score on the dependent variable (*math achievement*). The stem is the first digit of the score and the leaf is the second digit. Note that the legend indicates that **Stem width** equals 10. This means that entries that have 0 for the stem are less than 10, with the leaf indicating the actual number (1-9), those with 1 as the stem range from 10 to 19, etc. Note also that the legend indicates that each **leaf** equals one case. Each number in the **leaf** column represents the last digit of one person's *math achievement* score. The numbers in the **Frequency** column indicate how many participants had scores in the range represented by that stem. Thus, in the male plot, one student had a **stem** of 0 and a **leaf** of 3; that is, a score of 3. The frequency of male students with **leaves** between 5 and 9 is 7, and there were three scores of 5, two of 7, and two of 9. Eleven participants had a **stem** of 1 and a **leaf** of 0 to 4; one had a **leaf** of 0 (a score of 10); two had scores of 11, one had a score of 13, and six had a score of 14.

Boxplots are the last part of the output. There are two boxplots (one for males and one for females). By inspecting the plots, we can see that the median score for males is quite a bit higher than that for females, although there is some overlap of the boxes and substantial overlap of the full distributions. We need to be careful in concluding that males score higher than females, especially based on a small sample of students. In chapter 8, we will show how an inferential statistic (analysis of covariance) can tell us whether this apparent gender difference is actually due to gender differences in another variable (number of *math courses taken*).

Using the output to check your data for errors. Checking the boxplots and stem-and-leaf plots can help identify outliers that might be data entry errors. In this case, there aren't any.

Using the output to check your data for assumptions. As noted in the interpretation of Outputs 2.2a and 2.2b, you can tell if a variable is grossly nonnormal by looking at the boxplots. The stem-and-leaf plots provide similar information. You can also examine the skewness values for each gender separately in the table of **Descriptives** (see the circled skewness values). Note that for both males and females, the skewness values are less than one, which indicates that math achievement is approximately normal for both genders. This is an assumption of the *t* test and ANOVA, and multivariate versions of this assumption are required for many of the statistics performed in this book.

The **Descriptives** table also provides the variances for males and females. A key assumption of ANOVA and the *t* test is that the variances are approximately equal (i.e., the assumption of homogeneity of variances). Note that the variance is 36.38 for males and 44.84 for females. These do not seem grossly different, and if we did a Levene test on differences between males and females in variances on this variable, we'd find that the difference is not significantly different (see chapter 10, Morgan, Leech, Gloeckner, & Barrett, 2004). Thus, the assumption of homogeneous variances is not violated. The boxplots and stem-and-leaf plots help you see this.

Using Tables and Figures for EDA with Discrete Variables

Descriptives for Dichotomous Variables

We will now use the **Descriptives** command for each of the dichotomous variables. Once again, we could have done **Frequencies**, with or without frequency tables, but we chose Descriptives. This time we selected fewer statistics because the standard deviation, variance, and skewness values are not very meaningful with dichotomous variables.

Output 2.4: Descriptives for Dichotomous Variables

```
DESCRIPTIVES
  VARIABLES=gender alg1 alg2 geo trig calc mathgr
  /STATISTICS= MEAN MIN MAX .
```

Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean
gender	75	0	1	.55
algebra 1 in h.s.	75	0	1	.79
algebra 2 in h.s.	75	0	1	.47
geometry in h.s.	75	0	1	.48
trigonometry in h.s.	75	0	1	.27
calculus in h.s.	75	0	1	.11
math grades	75	0	1	.41
Valid N (listwise)	75			

Interpretation of Output 2.4

Output 2.4 includes only one table of **Descriptive Statistics**. Across the top row are the requested statistics of *N*, **Minimum**, **Maximum**, and **Mean**. We could have requested other statistics, but they would not be very meaningful for dichotomous variables. Down the left column are the variable labels. The *N* column indicates that all the variables have complete data. The **Valid *N* (listwise)** is 75, which also indicates that all the participants had data for each of our requested variables.

The most helpful column is the **Mean** column. You can use the mean to understand what percentage of participants fall into each of the two groups. For example, the mean of *gender* is .55, which indicates that 55% of the participants were coded as 1 (female); thus, 45% were coded 0 (male). Because the mean is greater than .50, there are more females than males. If the mean is close to 1 or 0 (e.g., algebra 1 and calculus), then splitting the data on that dichotomous variable might not be useful because there will be many participants in one group and very few participants in the other.

Using your output to check your data for errors. The **Minimum** column shows that all the dichotomous variables had “0” for a minimum, and the **Maximum** column indicates that all the variables have “1” for a maximum. This is good because it agrees with the codebook.

Frequency Tables

Displaying Frequency tables for variables can help you understand how many participants are in each level of a variable and how much missing data of various types you have. For nominal variables, most descriptive statistics are meaningless. Thus, having a frequency table is usually the best way to understand your nominal variables. We created a frequency table for *ethnicity* (a nominal variable) and for *father's education* (an ordered variable).

Output 2.5 Frequency Tables for a Nominal Variable and an Ordinal Variable

```
FREQUENCIES
  VARIABLES=ethnic faed
  /ORDER= ANALYSIS .
```

Frequencies**ethnicity**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Euro-Amer	41	54.7	56.2	56.2
	African-Amer	15	20.0	20.5	76.7
	Latino-Amer	10	13.3	13.7	90.4
	Asian-Amer	7	9.3	9.6	100.0
	Total	73	97.3	100.0	
Missing	multiethnic	1	1.3		
	blank	1	1.3		
	Total	2	2.7		
Total		75	100.0		

father's education

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	< h.s. grad	22	29.3	30.1	30.1
	h.s. grad	16	21.3	21.9	52.1
	< 2 yrs voc	3	4.0	4.1	56.2
	2 yrs voc	8	10.7	11.0	67.1
	< 2 yrs coll	4	5.3	5.5	72.6
	> 2 yrs coll	1	1.3	1.4	74.0
	coll grad	7	9.3	9.6	83.6
	master's	6	8.0	8.2	91.8
	MD/PhD	6	8.0	8.2	100.0
	Total	73	97.3	100.0	
Missing	System	2	2.7		
Total		75	100.0		

74% of fathers
have 2 years or
less of college.

Interpretation of Output 2.5

There is one **Frequency** table for *ethnicity* and one for *father's education*. The left-hand column shows the **Valid** categories (or levels or values), **Missing** values, and **Total** number of participants. The **Frequency** column gives the number of participants who had each value. The **Percent** column is the percent who had each value, including missing values. For example, in the ethnicity table, 54.7% of all participants were *Euro-American*, 20.0% were *African-American*, 13.3% were *Latino-American*, and 9.3% were *Asian-American*. There also were a total of 2.7% missing: 1.3% were *multi ethnic*, and 1.3% were left *blank*. The **valid percent** shows the percentage of those with *nonmissing* data at each value; for example, 56.2% of the 73 students with a single valid ethnic group were *Euro-Americans*. Finally, **Cumulative Percent** is the percentage of subjects in a category *plus* the categories listed above it. This last column is not very useful with nominal data, but can be quite informative for frequency distributions with several ordered categories. For example, in the distribution of father's education, 74% of the fathers had less than a bachelor's degree (i.e., they had not graduated from college).

Using your output to check your data for errors. Errors can be found by checking to see if the number missing is the number you expected. Also, if you have more than one type of missing data, and you assigned different numbers to these (e.g., 98 and 99), you will see the types listed in the first column.

Using the output to check your data for assumptions. Frequency tables are helpful for checking the levels of the variable to see if you have subjects in each one. If one of the levels does not have many subjects in it, it can create problems with difference statistics (see chapter 3 for an explanation of types of difference statistics).

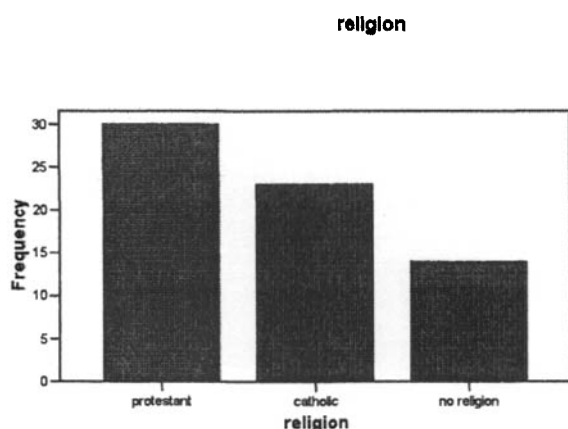
Bar Charts

With **nominal** data, you should not use a graphic that connects adjacent categories because with nominal data there is no necessary ordering of the categories or levels. Thus, it is better to make a bar graph or chart of the frequency distribution of variables like *religion*, *ethnic group*, or other nominal variables; the points that happen to be adjacent in your frequency distribution are not by necessity adjacent.

Output 2.6 Frequency Distribution Bar Chart for the Nominal Variable of Religion

```
FREQUENCIES
  VARIABLES=religion  /FORMAT=NOTABLE
  /BARCHART  FREQ
  /ORDER=  ANALYSIS
```

Frequencies



Interpretation of Output 2.6

There are two parts to the output: a statistics table (not shown) and the bar chart. The **Statistics** table indicates the valid *N* and the missing *N*. The **bar chart** presents a bar for each level of the nominal variable. Keep in mind that the order of the bars is arbitrary. The order of the bars can be changed by clicking on the plot, then clicking on the bars. The Chart Editor Properties box will appear and the level can be changed within this window.

Using your output to check your data for errors. Errors can be seen if the levels are not what you were expecting. For example, if you inadvertently entered the wrong number (a number that does not have a category label), you will see a bar for that number, with the number used as a label for the bar. Also, if there are no participants for a level, there will not be a bar for that level.

Using the output to check your data for assumptions. Bar charts are helpful if you are unsure how many categories a variable has; you can count the number of bars. You can also see the number of participants in each level. It is best if there are approximately the same number of subjects in each level if one wants to use the variable in procedures like ANOVA and MANOVA.

Some Additional Figures for Scale Variables

Histograms

Histograms (shown in chapter 1) look much like bar charts except in histograms there is no space between the boxes, indicating that there is a continuous variable theoretically underlying the scores (i.e., scores could theoretically be any point on a continuum from the lowest to highest score). Histograms can be used even if data, as measured, are not continuous, if the underlying variable is conceptualized as continuous. For example, the *competence scale* items were rated on a 4-point scale, but one could, theoretically, have any amount of competence.

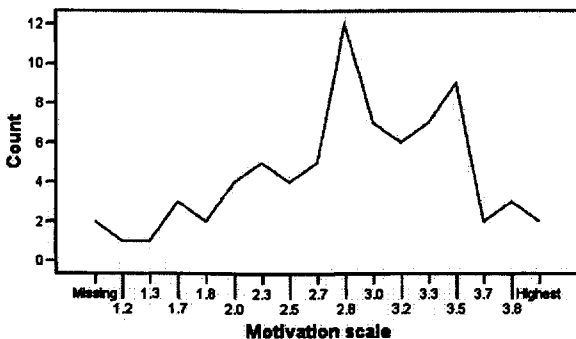
Frequency Polygons

Output 2.7 is a frequency polygon; it connects the points between the categories, and is best used with **approximately normal data**, but it can be used with ordinal data.

Output 2.7. Frequency Polygon Showing Approximately Normal Data

GRAPH
/LINE(SIMPLE)=COUNT BY motivation .

Graph



Interpretation of Output 2.7

The frequency polygon presents frequency **counts** for one variable (listed along the bottom; *motivation scale* in this example). Each level of the variable is listed along the bottom and the counts are listed along the side. We can see from this frequency polygon that the largest number of participants had a motivation level of 2.8 (12), and only one person had scores of 1.2 and 1.3.

Using your output to check your data for errors. Errors can be seen with the frequency polygon if there are values you were not expecting or if there are larger or smaller counts than you were expecting.

Using the output to check your data for assumptions. We can see that the data are somewhat normally distributed because the highest counts are in the middle, with smaller counts on either end. We would still want to check skewness values as well.

Matrix Scatterplots

To check linearity and get a visual idea about whether there is likely to be a problem with multicollinearity (see chapter 6), we can do matrix scatterplots. In this example, we used *math achievement*, *math courses taken*, *pleasure scale*, and *mosaic pattern test*.

To develop a scatterplot of *math achievement*, *math courses taken*, *pleasure scale*, and *mosaic pattern test*, follow these commands:

- **Graphs => Scatter.**
- Click on **Matrix.**
- Click on **Define.**
- Now, move *math achievement*, *math courses taken*, *pleasure scale*, and *mosaic pattern test* to the **Matrix Variables** box.
- Click on **OK.** You will get Output 2.8, the matrix scatterplot with circles for the data points.

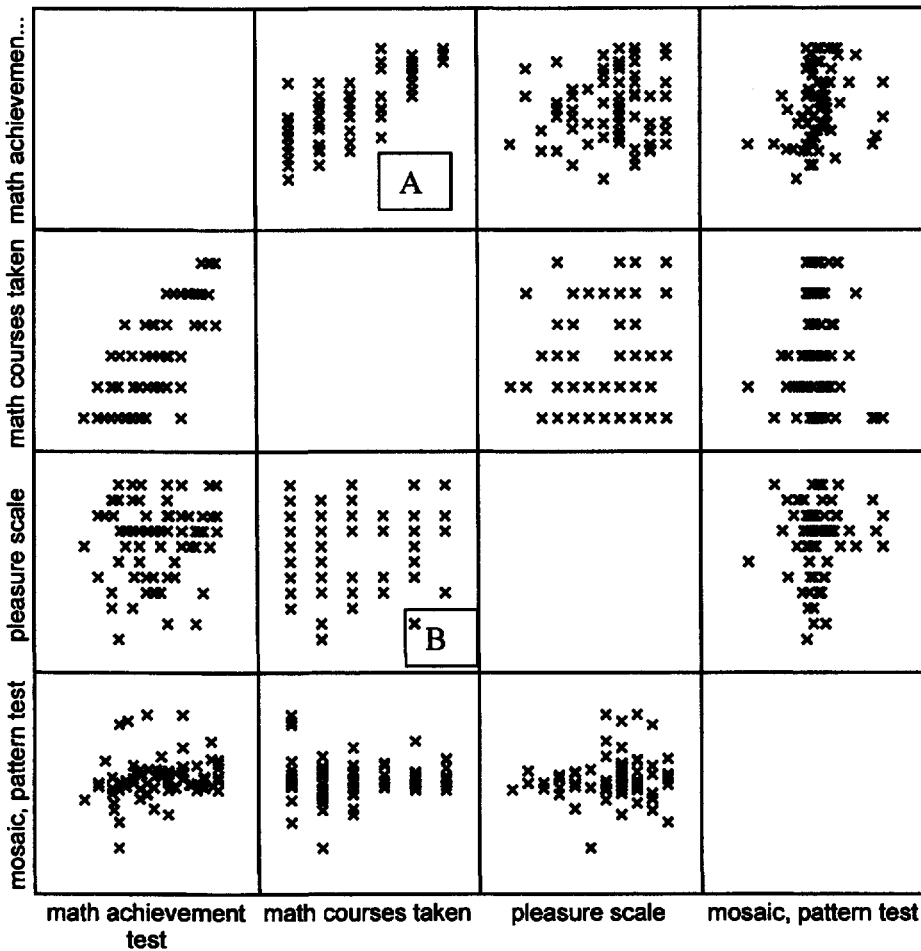
Chapter 2 – Data Coding and EDA

- To change the circles to the Xs (as we have below) double click on a data point (one of the circles). This will open the **Chart Editor**. Double click again on one of the circles in the Chart Editor. This will open the **Properties** window.
- In the **Properties** window, click on the arrow under **Type**. Select the X.
- Click on **Apply** and then **Close**.
- Close the window for the **Chart Editor** to get back to the **Output** window.

Output 2.8: Matrix Scatterplot

GRAPH

```
/SCATTERPLOT(MATRIX)=mathach mathcrs pleasure mosaic  
/MISSING=LISTWISE .
```



Interpretation of Output 2.8

The matrix scatterplot command in Output 2.8 creates multiple bivariate scatterplots. It is most helpful to create a matrix scatterplot when you need to understand the relationships among several variables. In a scatterplot, each X (or O) represents a data point. There are six bivariate scatterplots in the Output 2.8. Keep in mind that there are duplicate relationships shown (i.e., there are two bivariate scatterplots for each pair of variables); you should look at one side or the other of the empty boxes.

Using your output to check your data for errors. At times, there might be a data point that is extreme; there might be one that is not at all close to the other data points. If this is the case, it might be an outlier or it might be an error.

Using the output to check your data for assumptions. The clusters or clouds of Xs can be examined for the assumption of linearity. If a straight line can be drawn so that most of the Xs lie relatively close to it, we can assume that the two variables are related in a linear fashion (e.g., see *math achievement* and *math courses taken* indicated with an A in Output 2.8). If the Xs are not in any order or if they appear to be spread out in a cloud (e.g., see *pleasure* and *math courses taken*, indicated with an B in Output 2.8), we can assume that there is no relationship between the variables. If a scatterplot shows little relationship between two predictors, this means there is little chance of collinearity involving these variables. If it relates a predictor and a dependent variable, it means that the predictor is unlikely to contribute much to predicting the dependent variable.

However, when the clusters appear to be creating a curve, rather than a straight line, we can assume the variables might be related in a curvilinear fashion. We would not want to use them in computing a statistic that has the assumption of linearity without fixing this problem first by transforming one or both of the variables, as shown below.

Transforming Variables

If the variables do not meet the assumptions, we might consider transformation. Transformations usually increase the difficulty of interpretation of the data and may change the results of the study, so you need to be sure you really need to transform the data and also that you will be able to interpret the transformed variable in your writing.

Finding the best transformation for a variable may be guided by certain guidelines, but often trial and error is needed to determine if the transformation was successful. The most common transformations, what the transformation can fix, and the SPSS syntax commands are listed below. If you do not know which transformation is best, start where the arrow is on the figure and go up one row. After you transform a variable, you need to rerun the assumptions to see if the newly transformed variable meets the assumptions. If not, go down one row and check assumption. If assumptions are not met, go up two rows, etc.

<i>When to use</i>	<u>Transformation</u>	<u>SPSS Syntax</u>
To reduce negative skew	X^3	VAR=(X)**3
	X^2	VAR=(X)**2
Non transformed variable	X	
To reduce positive skew	log X	VAR=LG10(X)
	\sqrt{X}	VAR=SQRT(X)
	$1/X$	VAR=1/(X)
	$1/X^2$	VAR=1/(X)**2
To stretch both tails of the distribution (proportion data)	Arcsine X	VAR=ARSIN(X)
Note: Nontransformed variable=X		

Start
here.
→

Fig. 2.1. Transformation ladder.

You can do transformations either through the point and click method or with syntax. If you want to compute X^3 , X^2 , or $1/X^2$, you have to either type in the expression as written or use syntax because these functions are not available as specific choices in the **Compute** command.

To use the point and click method, follow these commands.

- **Transform => Compute.** The **Compute Variable** window will open. Type a name for the new variable in the **Target Variable** box => Click on **Type & Label** => The **Compute Variable: Type and Label** window will open => Scroll down the **Functions** box => Find the function you want to try (see Fig. 2.1 for assistance in choosing) and click on it => Click on the arrow above the box to move the function into the **Numeric Expression** box => Click on the variable you are interested in transforming => Click on the arrow to move it to the **Numeric Expression** box => **OK**. If the transformation is not listed under **Functions**, then you can just type it into the **Numeric Expression** box, or you can type it into the syntax listing. The new variable you computed will appear in the far right column in the **Data View** window. To see if the transformation was successful, retest the assumptions.

To use syntax, follow the guidelines below.

Replace "SQRT" with the chosen command from Fig. 2.1.

Put the old variable name here.

```

COMPUTE VAR = SQRT(oldvariablename) .
VARIABLE LABELS VAR 'Label for New Variable' .
EXECUTE .
    
```

This is where you should add the new label for your variable.

Output 2.9 Transformation of Competence

```

COMPUTE VAR = (competence)**2 .
VARIABLE LABELS VAR 'SquaredComp' .
EXECUTE .
DESCRIPTIVES
  VARIABLES=VAR
  /STATISTICS=MEAN STDDEV MIN MAX SKEWNESS .
    
```

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std.	Skewness	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error
SquaredComp	73	1.00	16.00	11.2894	3.73342	-.911	.281
Valid N (listwise)	73						

```

COMPUTE VAR = (competence)**3 .
VARIABLE LABELS VAR 'CubedComp' .
EXECUTE .
DESCRIPTIVES
  VARIABLES=VAR
  /STATISTICS=MEAN STDDEV MIN MAX SKEWNESS .
    
```

SPSS for Intermediate Statistics

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std.	Skewness	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error
CubedComp	73	1.00	64.00	39.6027	17.16401	-.445	.281
Valid N (listwise)	73						

Interpretation of Output 2.9

As shown in Output 2.1b, *competence* is negatively skewed (skewness = -1.634). In order to use this variable with statistics that require a normally distributed variable, we needed to transform it. Since *competence* is negatively skewed, we chose to square it to see if it would be less skewed. Output 2.9 shows the syntax for transforming *competence* by squaring it. After squaring, the skewness = -.911. We might have stopped there, but -.911 is very close to -1.0, and we want it not to be skewed even this much, so we transformed it by cubing (the original values). After transforming, the skewness value is -.445, which is still slightly skewed, but now falls well within the recommended guidelines of $< |1|$ (greater than -1 and less than +1).

Interpretation Questions

- 2.1 Using Outputs 2.1a and 2.1b: a) What is the mean *visualization test* score? b) What is the range for *grades in h.s.*? c) What is the minimum score for *mosaic pattern test*? How does this compare to the values for that variable as indicated in chapter 1? Why would this score be the minimum?
- 2.2 Using Outputs 2.1b: a) For which of the variables that we called scale is the skewness statistic more than ± 1.00 ? b) Why is the answer important? c) How many participants have some missing data? (Hint: Check chapter 1 if you don't remember the sample size.) e) What percent of students have a valid (non-missing) *motivation* or *competence* score? f) Can you tell from Output 2.1b how many are missing both *motivation* and *competence* scores?
- 2.3 Using Output 2.4: a) Can you interpret the means? Explain. b) How many participants are there all together? c) How many have complete data (nothing missing)? d) What percent are *male* (if *male*=0)? e) What percent took *algebra I*?
- 2.4 Using Output 2.5: a) 9.6% of what set of participants are *Asian-American*? b) What percent of students have fathers who had a high school education or less? c) What percent had fathers with a master's degree or higher?
- 2.5 In Output 2.8: a) Why are matrix scatterplots useful? What assumption(s) are tested by them?

Extra SPSS Problems

Using the College Student data file, do the following problems. Print your outputs and circle the key parts of the output that you discuss.

- 2.1 For the variables with five or more ordered levels, compute the skewness. Describe the results. Which variables in the data set are approximately normally distributed/scale? Which ones are ordered but not normal?

Chapter 2 – Data Coding and EDA

- 2.2 Do a stem-and-leaf plot for same-sex parent's height split by gender. Discuss the plots.
- 2.3 Which variables are nominal? Run Frequencies for the nominal variables and other variables with fewer than five levels. Comment on the results.
- 2.4 Do boxplots for student height and for hours of study. Compare the two plots.

CHAPTER 3

Selecting and Interpreting Inferential Statistics

To understand the information in this chapter, it may be necessary to remember or to review the sections in chapter 1 about **variables** and levels of **measurement** (nominal, dichotomous, ordinal, and approximately normal/scale). It is also necessary to remember the distinction we made between difference and associational research questions and between **descriptive and inferential statistics**. This chapter focuses on inferential statistics, which as the name implies, refers to statistics that make inferences about population values based on the sample data that you have collected and analyzed. What we call **difference inferential statistics** lead to inferences about the differences (usually mean differences) between groups in the populations from which the samples were drawn. **Associational inferential statistics** lead to inferences about the association or relationship between variables in the population. Thus, the purpose of inferential statistics is to enable the researcher to make generalizations beyond the specific sample data. Before we describe how to select and interpret inferential statistics, we will introduce design classifications.

General Design Classifications for Difference Questions

Many research questions focus on whether there is a significant difference between two or more groups or conditions. When a group comparison or difference question is asked, the independent variable and design can be classified as between groups or within subjects. Understanding this distinction is one essential aspect of determining the proper statistical analysis for this type of question.

Labeling difference question designs. Brief descriptive labels identify the design for other researchers and also guide us toward the proper statistics to use. We do not have design classifications for the descriptive or associational research questions, so this section only applies to difference questions. Designs are usually labeled in terms of (a) the overall type of design (between groups or within subjects), (b) the number of independent variables, and (c) the number of levels within each independent variable.

Between-groups designs. These are designs where each participant in the research is in one and only one condition or group. For example, there may be three groups (or levels or values) of the independent variable, *treatment type*. If the investigator wished to have 20 participants in each group, then 60 participants would be needed to carry out the research.

Within subjects or repeated measures designs. These designs are conceptually the opposite of between groups designs. In within subjects designs, each participant in the research receives or experiences all of the conditions or levels of the independent variable. These designs also include examples where the participants are matched by the experimenter or in some natural way (e.g., twins, husband and wife, or mother and child). When each participant is assessed more than once, these designs are also referred to as **repeated measures designs**. Repeated measures designs are common in longitudinal research and intervention research. Comparing performance on the same dependent variable assessed before and after intervention (pretest and posttest) is a common example of a repeated measures design. We might call the independent variable in such a study “time of measurement.” The HSB study did not really have a within-subjects aspect to the design. However, one of the variables is repeated (*visualization* with two levels: *visualization* test and *visualization* retest) and one is within (*education*, each student has both a *mother’s education* and *father’s education*). To demonstrate a **within subjects design** and the use of **repeated measured ANOVA**, we will use another data set, called **Product Data**, that is found on the CD in this book. This small data set has **within subjects data** ratings by each participant, one for each

of four different products (e.g., DVDs, but they could be any four stimuli). The same types of analysis could be done if, instead of each participant rating four different products in the same session, the ratings were done for satisfaction with the same product at four times. In that case, the data would be **repeated measures data**. In addition, to demonstrate a **doubly multivariate design**, in which there are repeated assessments of several measures, we will use the data set called **mixedMANOVAdata**.

Single factor designs. If the design has only one independent variable (in either a between groups design or a within subjects design), then it should be described as a basic or single factor or one-way design. (“Factor” and “way” are other names for difference independent variables.) For example, a between groups design with one independent variable that has four levels is a single factor or one-way between groups design with four levels. If the design is a within subjects design with four levels, then it would be described as a single factor, repeated measures design with four levels (e.g., the same test being given four times).

Between groups factorial designs. When there is more than one group difference independent variable, and each level of each factor (independent variable) is possible in combination with each level of each of the other factors, the design is called **factorial**. For example, a factorial design could have two independent variables (i.e., factors) *gender* and *ethnicity*, allowing for male and female members of each ethnic group. In these cases, the number of levels of *each* factor (independent variable) becomes important in the description of the design. If *gender* had two levels (i.e., males and females) and *ethnicity* had three levels (i.e., Caucasian, Hispanic, and African-American), then this design is a 2 x 3 between groups factorial design.

Mixed factorial designs. If the design has a between groups variable and a within subjects independent variable, it is called a mixed design. For example, say the independent variables are *gender* (a between-groups variable) and *time of measurement* (with pretest and posttest as within-subjects levels); this is a 2 x 2 mixed factorial design with repeated measures on the second factor. The mixed design is common in experimental studies with a pretest and posttest.

Remember, when describing a design, that each independent variable is described using one number, which is the number of levels for that variable. Thus a design description with two numbers (e.g., 3 x 4) has two independent variables or factors, which have three and four levels, respectively. The dependent variable is not part of the design description, so it was not considered in this section.

Selection of Inferential Statistics

It is time to think about how to decide which of the many possible inferential statistics to use. Because many statistical tests are introduced, don't be concerned if you don't know about all of the tests mentioned. You should come back to this chapter later, from time to time, when you have to make a decision about which statistic to use, and by then, the tests will be more familiar.

In Fig 3.1, we present eight steps to guide you in the selection of a proper statistical test for data analysis. Remember that **difference questions compare groups** and utilize the statistics, which we call difference inferential statistics. These statistics (e.g., *t* test and analysis of variance) are shown in Tables 3.1 and 3.3.

Associational questions utilize what we call associational inferential statistics. The statistics in this group examine the association or relationship between two or more variables and are shown in Tables 3.2

and 3.4. This distinction is somewhat of a simplification. There is often more than one possible statistic that can be used.

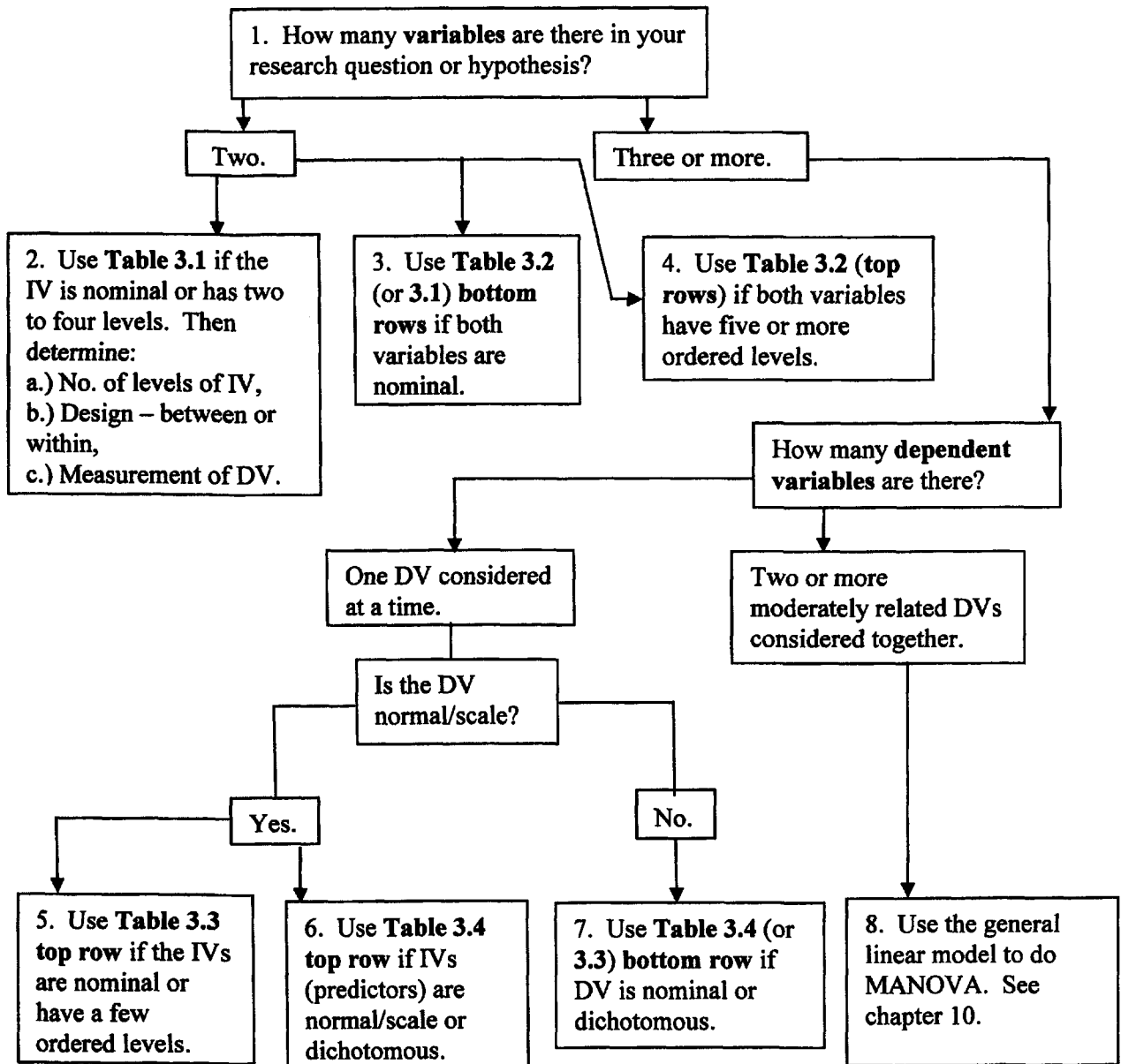


Fig. 3.1. A decision tree to help select an appropriate inferential statistic from Tables 3.1 to 3.4.

Using Tables 3.1 to 3.4 to Select Inferential Statistics

As with research questions and hypotheses discussed in chapter 1, we divide inferential statistics into basic and complex. For *basic* (or *bivariate*) statistics, there is *one* independent and *one* dependent variable and you will use Table 3.1 or 3.2. These basic statistics are discussed in more detail in our companion book, Morgan, Leech, Gloeckner, and Barrett (2004). For *complex statistics*, there are three or more variables. We decided to call them **complex** rather than **multivariate**, which is more common in the literature, because there is not unanimity about the definition of multivariate, and several complex statistics (e.g., factorial ANOVA) are not usually classified as multivariate. For complex statistics, you will use Tables 3.3 or 3.4. The complex statistics shown in Tables 3.3 and 3.4 are discussed in the

remaining chapters in this book, and assignments and outputs are given demonstrating how to compute them using SPSS 12. There are other statistics, but these four tables include most of the inferential statistics that you will encounter in reading research articles. Note that the boxes in the decision tree are numbered to correspond to the numbers in the text below, which expands some on the decision tree or flowchart.

1. Decide how many variables there are in your research question or hypothesis. If there are only two variables, use Tables 3.1 or 3.2. If there is *more* than one independent and/or more than one dependent variable (i.e., three or more variables) to be used in this analysis, use Tables 3.3 and 3.4.

Basic (Two variable) Statistics

2. If the independent variable is nominal (i.e., has unordered levels) or has a few (2-4) ordered levels, use Table 3.1. Then, your question is a **basic difference question** to compare groups. You must then determine: (a) whether there are two *or* more than two levels (also called categories or groups or samples) of your *independent* variable, (b) whether the design is between groups or within subjects, and (c) whether the measurement of the *dependent* variable is (i) normal/scale and parametric assumptions are not markedly violated, *or* (ii) ordinal, or (iii) nominal *or* dichotomous. The answers to these questions lead to a specific box in Table 3.1 and statistics such as **independent or paired sample *t* tests, one-way ANOVA, chi-square** and several other nonparametric tests.

3. If both variables are nominal or dichotomous, you could ask either a **difference question** (use the bottom row of Table 3.1, e.g., **chi-square**) or an **associational question** and use the bottom row of Table 3.2 (**phi or Cramer's V**). Note, in the second to bottom row of Table 3.2, we have included **eta**; an associational statistic used with one nominal and one normal or scale variable. We will later see it used as an effect size measure with ANOVAs. There are many other nonparametric associational measures, some of which are in Table 3.2.

4. If both variables have many (we suggest five or more) ordered levels, use Table 3.2 (top two rows). Your research question would be a **basic two variable (bivariate) associational question**. Which row you use depends on *both* variables. If both are normal/scale, then you would probably select the **Pearson product moment correlation or bivariate regression** (top row). Regression should be used if one has a clearly directional hypothesis, with an independent and dependent variable. Correlation is chosen if one is simply interested in how the two variables are related. If one or both variables are ordinal (ranks or grossly skewed), or if other assumptions are markedly violated, the second row (**Kendall's tau or Spearman rho**) is a better choice.

Complex (3 or more variable) Questions and Statistics

It is possible to break down a complex research problem or question into a series of basic (bivariate) questions and analyses. However, there are advantages to combining them into one complex analysis: additional information is provided and a more accurate overall picture of the relationships is obtained.

5. If you have one normally distributed (scale) dependent variable and two (or perhaps three or four) independent variables, each of which is nominal or has a few (2-4) ordered levels, you will use the top row of Table 3.3 and one of three types of **factorial ANOVA**. These analysis of variance (ANOVA) statistics answer **complex difference questions**.

Note, in Table 3.3, that there are no complex difference statistics in SPSS designed to be used when the dependent variable is ordinal. **Loglinear analysis** is a nonparametric statistic similar to the between group factorial ANOVA when all the variables are nominal or dichotomous (see Table 3.3).

Table 3.1. Selection of an Appropriate Inferential Statistic for Basic, Two Variable, Difference Questions or Hypotheses

Scale of Measurement of Dependent Variable ↓	COMPARE ↓	One Factor or Independent Variable with 2 Levels or Categories/Groups/Samples		One Independent Variable 3 or more Levels or Groups	
		Independent Samples or Groups (Between)	Repeated Measures or Related Samples (Within)	Independent Samples or Groups (Between)	Repeated Measures or Related Samples (Within)
Dependent Variable Approximates Normal /Scale Data and Assumptions Not Markedly Violated	MEANS	INDEPENDENT SAMPLES <i>t</i> TEST ch. 10 or ONE-WAY ANOVA QRG	PAIRED SAMPLES <i>t</i> TEST QRG	ONE-WAY ANOVA QRG	GLM REPEATED MEASURES ANOVA ch. 9
Dependent Variables Clearly Ordinal or Parametric Assumptions Markedly Violated	MEAN RANKS	MANN-WHITNEY QRG	WILCOXON QRG	KRUSKAL-WALLIS QRG	FRIEDMAN ch. 9
Dependent Variable Nominal or Dichotomous	COUNTS	CHI-SQUARE QRG	MCNEMAR QRG	CHI-SQUARE QRG	COCHRAN Q TEST QRG

QRG = Quick Reference Guide (Appendix A).
Ch. 9 refers to that chapter in this book.

Table 3.2. Selection of an Appropriate Inferential Statistic for Basic, Two Variable, Associational Questions or Hypotheses ^a

Level (Scale) of Measurement of Both Variables ↓	RELATE ↓	Two Variables or Scores for the Same or Related Subjects
Variables Are Both Normal /Scale and Assumptions Not Markedly Violated	SCORES	PEARSON (<i>r</i>) or BIVARIATE REGRESSION QRG
Both Variables at Least Ordinal Data or Assumptions Markedly Violated	RANKS	KENDALL TAU or SPEARMAN (Rho) QRG
One Variable Is Normal /Scale and One Is Nominal		ETA QRG
Both Variables Are Nominal or Dichotomous	COUNTS	PHI or CRAMER'S V QRG

a. It is acceptable to use statistics that are in the box(es) below the appropriate statistic, but there is usually some loss of power. It is not acceptable to use statistics in boxes above the appropriate statistic or ones in another column.

6. Use a **MANOVA** (second row of Table 3.3) if you have two or more normal (scale) dependent variables treated simultaneously. **MANOVA** is a better choice than several ANOVAs if the dependent variables are related statistically and conceptually.

7. The statistics in Table 3.4 are used to answer **complex associational questions**. If you have two or more independent or predictor variables and one normal (scale) dependent variable, the top row of Table 3.4 and **multiple regression** are appropriate.

8. If the dependent variable is dichotomous or nominal, consult the bottom row of Table 3.4. In general, **logistic regression** is used if some or all the independent variables are dichotomous. **Discriminant analysis** can best be used if the independent variables are all ordered/scale.

Table 3.3. Selection of the Appropriate Complex (Two or More Independent Variables) Statistic to Answer Difference Questions or Hypotheses

Dependent Variable(s) ↓	Two or More Independent Variables		
	All Between Groups	All Within Subjects	Mixed (Between & Within)
One Normal/ Scale Dependent Variable	GLM, Factorial ANOVA or ANCOVA ch. 8	GLM With Repeated Measures on All Factors ch. 9 ^a	GLM With Repeated Measures on Some Factors ch. 9 ^a
More Than One Normal/Scale Dependent Variable	GLM, Multivariate MANOVA or MANCOVA ch. 10	GLM Doubly multivariate MANOVA With Repeated Measures on All Factors ch.10 ^b	GLM Doubly multivariate MANOVA With Repeated Measures on Some Factors ch. 10
Ordinal Dependent Variable	None Common	None Common	None Common
Dichotomous Dependent Variable	LOG LINEAR QRG	None Common	None Common

^a In chapter 9, both a multivariate analysis (MANOVA) and a univariate analysis are performed.

^b In chapter 10, the doubly multivariate example also has a between groups factor, so it is actually a mixed design.

Table 3.4. Selection of the Appropriate Complex Associational Statistic for Predicting a Single Dependent/Outcome Variable From Several Independent Variables

One Dependent or Outcome Variable ↓	Several Independent or Predictor Variables		
	Normal or Scale	Some Normal Some Dichotomous (2 category)	All Dichotomous
Normal/Scale (Continuous)	MULTIPLE REGRESSION ch. 6	MULTIPLE REGRESSION ch. 6	MULTIPLE REGRESSION ch. 6
Dichotomous	DISCRIMINANT ANALYSIS ch. 7	LOGISTIC REGRESSION ch. 7	LOGISTIC REGRESSION ch. 7

Occasionally you will see a research article in which a dichotomous *dependent variable* was used with a *t* test, ANOVA, or Pearson correlation. Because of the special nature of dichotomous variables, this is not necessarily wrong, as would be the use of a nominal (three or more unordered levels) dependent variable with these parametric statistics. However, we think that it is usually a better practice to use the same statistics with dichotomous variables that you would use with nominal variables, except that it is appropriate to use dichotomous independent variables in multiple regression (see Table 3.4).

Other Multivariate (Complex) Statistics

Not shown, in part because they did not fit the format of the tables, are six complex associational statistics for analyzing a number of variables at a time, which you may see in the literature. **Cronbach's alpha**, a technique used to assess the internal consistency reliability of multiple item scales, is discussed, along with some other reliability measures, in chapter 4. In **exploratory factor analysis**, one postulates that there is a smaller set of latent variables or constructs. Factor analysis and **principal components analysis**, which is used to reduce a relatively large number of variables to a smaller number of groups of variables, are discussed in chapter 5. **Canonical correlation** involves correlation of linear combinations of one set of variables with linear combinations of another set of variables. Thus, it is useful when you have two sets of variables and want to see the patterns of correlations between the two sets. How to compute it with syntax is shown in chapter 10.

Because they cannot be computed using SPSS (without the extra program called AMOS), two other forms of complex statistics are not discussed in this book. **Structural equation models (SEM)** are models that describe relationships among latent (unobserved) variables and manifest (observed) variables. **Hierarchical Linear Models (HLM)** enable one to model nested data (data in which certain variables are only present in a subset of one's data) over time. Both SEM and HLM provide tests of the accuracy of proposed models, and both are very useful for drawing better inferences from large sets of data. However, it is important to realize that, despite the language sometimes used in discussing SEM and HLM, even they do not enable one to determine causal relationships (e.g., see the APA Task Force on Statistical Inference report, Wilkinson, et al., 1999, p. 600).

The General Linear Model

Whether or not there is a relationship between variables can be answered in two ways. For example, if each of two variables provides approximately normally distributed data with five or more levels, then Fig. 3.1 and Table 3.2 indicate that the statistic to use is either the Pearson correlation or bivariate regression, and that would be our recommendation. However, some researchers choose to divide the independent variable into a few categories such as low, medium, and high and then do a one-way ANOVA. In another example, some researchers who start with an independent variable that has only a few (say, two to four) *ordered* categories may choose to do a correlation instead of a one-way ANOVA. Although these choices are not necessarily wrong, we do not think they are the best practice. In the first example information is lost by dividing a continuous independent variable into a few categories. In the second example, there would be a restricted range, which tends to decrease the size of the correlation coefficient.

In the above examples, we recommended one of the choices, but the fact that there *are* two choices raises a bigger and more complex issue. Statisticians point out, and can prove mathematically, that the distinction between difference and associational statistics is an artificial one, in that ANOVA and multiple regression using dummy variables are often mathematically the same. In fact, SPSS calculates ANOVA and MANOVA using this regression approach. The bottom of Fig. 3.2 shows these parallels and that, although we have made a distinction between difference and associational inferential statistics, they both serve the purpose of exploring (top box) relationships and both are subsumed by the general linear model (middle box). Statisticians state that all common parametric statistics are relational. Thus, the full range of methods used to analyze one continuous dependent variable and one or more independent variables, either continuous or categorical, are mathematically similar. The model on which this is based is called the **General Linear Model**. The idea is that the relationship between the

independent and dependent variables can be expressed by an equation with weights for each of the independent/predictor variables plus an error term.

What this means is that if you have a continuous, normally distributed dependent/outcome variable and five or so levels of an ordered, normally distributed independent variable, it would be appropriate to analyze it with either bivariate regression or a one-way ANOVA. You will get a similar answer with regard to the significance level. Note in Fig. 3.1 and Table 3.3 that SPSS uses the GLM to perform a variety of statistics including factorial ANOVA and MANOVA. Although we recognize that our distinction between difference and associational parametric statistics is a simplification, we still think it is useful heuristically. We hope that this glimpse of an advanced topic is clear and helpful.

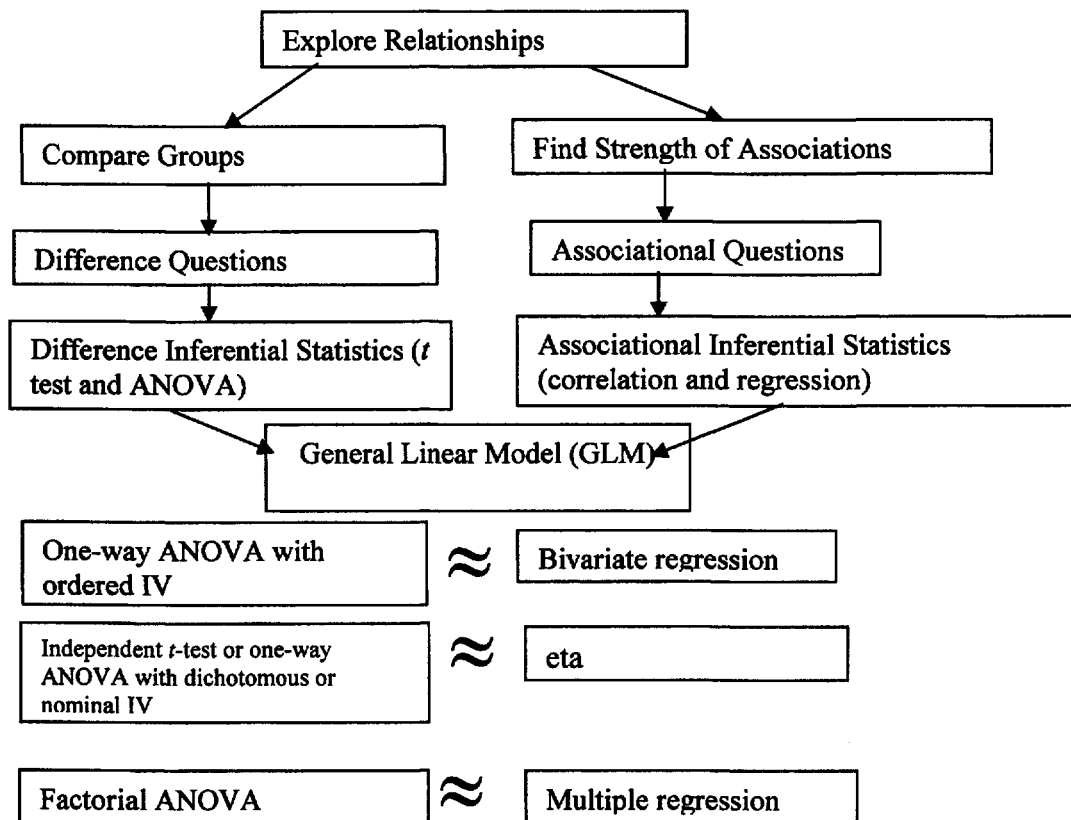


Fig. 3.2. A general linear model diagram of the selection of inferential statistics.

Interpreting the Results of a Statistical Test

In the following chapters, we present information about how to check assumptions, do SPSS analysis, and interpret complex statistics, and write about them. For each statistic, the SPSS computations produce a **calculated value** based on the specific data in your study. SPSS labels them *t*, *F*, etc. or just value.

Statistical Significance

The calculated value is compared to a **critical value** (found in a statistics table or stored in the computer's memory) that takes into account the degrees of freedom, which are usually based on the number of participants. Figure 3.3 shows how to interpret any inferential test once you know the probability level (*p* or sig.) from the computer or statistics table. In general, if the calculated value of the statistic (*t*, *F*, etc.) is relatively large, the probability or *p* is small (e.g., .05, .01, .001). If the probability

is *less than* the preset alpha level (usually .05), we can say that the results are **statistically significant** or that they are significant at the .05 level or that $p < .05$. We can also reject the null hypothesis of no difference or no relationship. Note that, using SPSS computer printouts, it is quite easy to determine statistical significance because the actual significance or probability level (p) is printed so you do not have to look up a critical value in a table. SPSS labels this p value **Sig.** so all of the common inferential statistics have a common metric. This level is also the probability of a Type I error or the probability of rejecting the null hypothesis when it is actually true. Thus, regardless of what specific statistic you use, if the Sig. or p is small (usually less than .05) the finding is *statistically significant*, and you can reject the null hypothesis of no difference and assume there is a difference (or no association).

Practical Significance Versus Statistical Significance

Students, and sometimes researchers, misinterpret statistically significant results as being practically or clinically important. But **statistical significance is not the same as practical significance or importance**. With large samples, you can find statistical significance even when the differences or associations are very small/weak. Thus, in addition to statistical significance, we will examine **effect size**. It is quite possible, with a large sample, to have a statistically significant result that is weak (i.e., has a small effect size). Remember that the null hypothesis is that there is *no* difference or *no* association. A significant result with a small effect size means that we can be very confident that there is *some* difference or association, but it is probably small and may not be practically important.

Sig. ¹	Meaning	Null Hypothesis	Interpretation
1.00	$p = 1.00$	Don't Reject	Not Statistically Significant (could be due to chance)
.50	$p = .50$	↓	↓
.06	$p = .06$		
.05	$p \leq .05$	Reject ²	Statistically Significant ³ (not likely due to chance)
.01	$p = .01$	↓	↓
.000	$p < .001$		

1. SPSS uses Sig. to indicate the significance or probability level (p) of all inferential statistics. This is just a sample of Sig. values, which would be any value from 0 to 1.

2. $p \leq .05$ is the typical alpha level that researchers use to assess whether the null hypothesis should be rejected or not. However, sometimes researchers use more liberal levels (e.g., .10 in exploratory studies) or more conservative levels (e.g., .01).

3. Statistically significant does *not* mean that the results have practical significance or importance.

Fig. 3.3. Interpreting inferential statistics using the SPSS Sig.

Confidence Intervals

One alternative to null hypothesis significance testing (NHST) is to create **confidence intervals**. These intervals provide more information than NHST and *may* provide more practical information. Suppose one knew that an increase in reading scores of 5 points would lead to a functional increase in reading performance. Two methods of instruction were compared. The result showed that students who used the new method scored statistically significantly higher than those who used the other method. According to NHST, we would reject the null hypothesis of no difference between methods and conclude that our new method is better. If we apply confidence intervals to this same study, we can determine an interval that contains the *population mean difference* 95% of the time. If the lower bound of that interval is greater

than 5 points, we conclude that using this method of instruction would lead to a practical or functional increase in reading levels. If however, the confidence interval ranged from say 1 to 11, the result would be statistically significant, but the mean difference in the population could be as little as 1 point, or as big as 11 points. Given these results, we could not be confident that there would be a practical increase in reading using the new method.

Effect Size

A statistically significant outcome does not give information about the strength or size of the outcome. Therefore, it is important to know, in addition to information on statistical significance, the size of the effect. Effect size is defined as the strength of the relationship between the independent variable and the dependent variable, and/or the magnitude of the difference between levels of the independent variable with respect to the dependent variable. Statisticians have proposed many effect size measures that fall mainly into two types or families: the *r* family and the *d* family.

The *r* family of effect size measures. One method of expressing effect sizes is in terms of strength of association. The most well-known variant of this approach is the Pearson correlation coefficient, *r*. Using Pearson *r*, effect sizes are always less than |1.0|, varying between -1.0 and +1.0 with 0 representing no effect and +1 or -1 the maximum effect. This *family* of effect sizes includes many other associational statistics as well, such as rho (*rs*), phi (*φ*), eta (*η*), and the multiple correlation (*R*).

The *d* family of effect size measures. The *d* family focuses on magnitude of difference rather than strength of association. If one compares two groups, the effect size (*d*) can be computed by subtracting the mean of the second group from the mean of the first group and dividing by the pooled standard deviation of both groups. The general formula is on the left. If the two groups have equal *ns*, the pooled *SD* is the average of the *SDs* for the two groups. When *ns* are unequal, the formula on the right is the appropriate one.

$$d = \frac{M_1 - M_2}{SD_{pooled}}$$

$$d = \frac{M_A - M_B}{\sqrt{\frac{(n_A - 1)SD_A^2 + (n_B - 1)SD_B^2}{n_A + n_B - 2}}}$$

There are many other formulas for *d* family effect sizes, but they all express effect size in standard deviation units. Thus, a *d* of .5 means that the groups differ by one half of a pooled standard deviation. Using *d*, effect sizes usually vary from 0 to + or - 1 but *d* can be more than 1.

Issues about effect size measures. Unfortunately, as just indicated, there are many different effect size measures and little agreement about which to use. Although *d* is the most commonly discussed effect size measure for differences, it is not available on SPSS outputs. However, *d* can be calculated by hand based on information in the SPSS printout. The correlation coefficient, *r*, and other measures of the strength of association such as phi (*φ*), eta² (*η*²), and *R*² are available in SPSS 12.

There is disagreement among researchers about whether it is best to express effect size as the unsquared or squared *r* family statistic (e.g., *r* or *r*²). It has been common to use the squared versions because they indicate the percentage of variance in the dependent variable that can be predicted from the independent variable(s). However, some statisticians argue that these usually small percentages give you an underestimated impression of the strength or importance of the effect. Thus, we (like Cohen, 1988) decided to use the unsquared statistics (*r*, *φ*, *η*, and *R*) as our *r* family indexes.

Although the 4th edition of the *Publication Manual of the American Psychological Association* (APA, 1994) recommended that researchers report effect sizes, relatively few researchers did so before 1999 when the APA Task Force on Statistical Inference stated that effect sizes should *always* be reported for your primary results (Wilkinson & The APA Task Force, 1999). The 5th edition (APA, 2001) adopted this recommendation of the Task Force, so more and more journal articles discuss the size of the effect as well as whether or not the result was statistically significant.

Interpreting Effect Sizes

Assuming that you have computed an effect size measure, how should it be interpreted? Table 3.5 provides guidelines for interpreting the size of the “effect” for five common effect size measures based on Cohen (1988).

Note that these guidelines are based on the effect sizes *usually found* in studies in the behavioral sciences. Thus, they do not have absolute meaning and are only relative to typical findings in these areas. For that reason, we might suggest using larger than typical instead of large, typical instead of medium, and smaller than typical instead of small. The guidelines will not apply to all subfields in the behavioral sciences, and they definitely will not apply to fields where the usually expected effects are either larger or smaller. It is advisable to examine the research literature to see if there is information about typical effect sizes on the topic.

Table 3.5. *Interpretation of the Strength of a Relationship (Effect Sizes)*

General Interpretation of the Strength of a Relationship	The <i>d</i> Family ^a	The <i>r</i> Family ^b		
	<i>d</i>	<i>r</i> and ϕ	<i>R</i>	η (eta) ^c
Very Large	$\geq 1.00^d$	$\geq .70$.70+	.45+
Large	.80	.50	.51	.37
Medium	.50	.30	.36	.24
Small	.20	.10	.14	.10

^a *d* values can vary from 0.0 to much greater than + or -1.0, but the latter is relatively uncommon.

^b *r* family values can vary from 0.0 to + or - 1.0, but except for reliability (i.e., same concept measured twice), *r* is rarely above .70. In fact, some of these statistics (e.g., ϕ) have a restricted range in certain cases; that is, the maximum ϕ may be less than 1.0.

^c Partial etas from SPSS multivariate tests are equivalent to *R*. Use *R* column.

^d The numbers in this table should be interpreted as a range of values. For example a *d* greater than .90 (or less than -.90) would be described as very large, a *d* between say .70 and .90 would be called large, and *d* between say .60 and .70 would be medium to large.

Cohen (1988) provided research examples of what we labeled small, medium, and large effects to support the suggested *d* and *r* family values. Many researchers would not consider a correlation (*r*) of .5 to be very strong because only 25% of the variance in the dependent variable is predicted. However, Cohen argued that a *d* of .8 (and an *r* of .5, which he shows are mathematically similar) are “grossly perceptible and therefore large differences, as (for example is) the mean difference in height between 13- and 18-year-old girls” (p. 27). Cohen stated that a small effect may be difficult to detect, perhaps because it is in a less well-controlled area of research. Cohen's medium size effect is “...visible to the naked eye. That is, in the course of normal experiences, one would become aware of an average difference in IQ between clerical and semi-skilled workers...” (p. 26).

Even effect size is not the same as practical significance. Although effect size measures indicate the strength of the relationship and, thus, are more relevant for practical significance or importance than statistical significance, they are not direct measures of the importance of a finding. As implied above, what constitutes a large or important effect depends on the specific area studied, the context, and the methods used. Furthermore, practical importance always involves a judgment by the researcher and the consumers (e.g., clinicians, clients, teachers, school boards) of research that takes into account such factors as cost and political considerations. A common example is that the effect size of taking daily aspirin and its effect on heart attacks is quite small, but the practical importance is high because preventing heart attacks is a life or death matter, the cost of aspirin is low, and side effects are uncommon. On the other hand, a curriculum change could have a large effect size but be judged to not be practical because of high costs and/or extensive opposition to its implementation.

Steps in Interpreting Inferential Statistics

First, decide whether to reject the null hypothesis. However, that is not enough for a full interpretation. If you find that the outcome is statistically significant, you need to answer at least two *more* questions. Figure 3.4 summarizes the steps described below about how to more fully interpret the results of an inferential statistic.

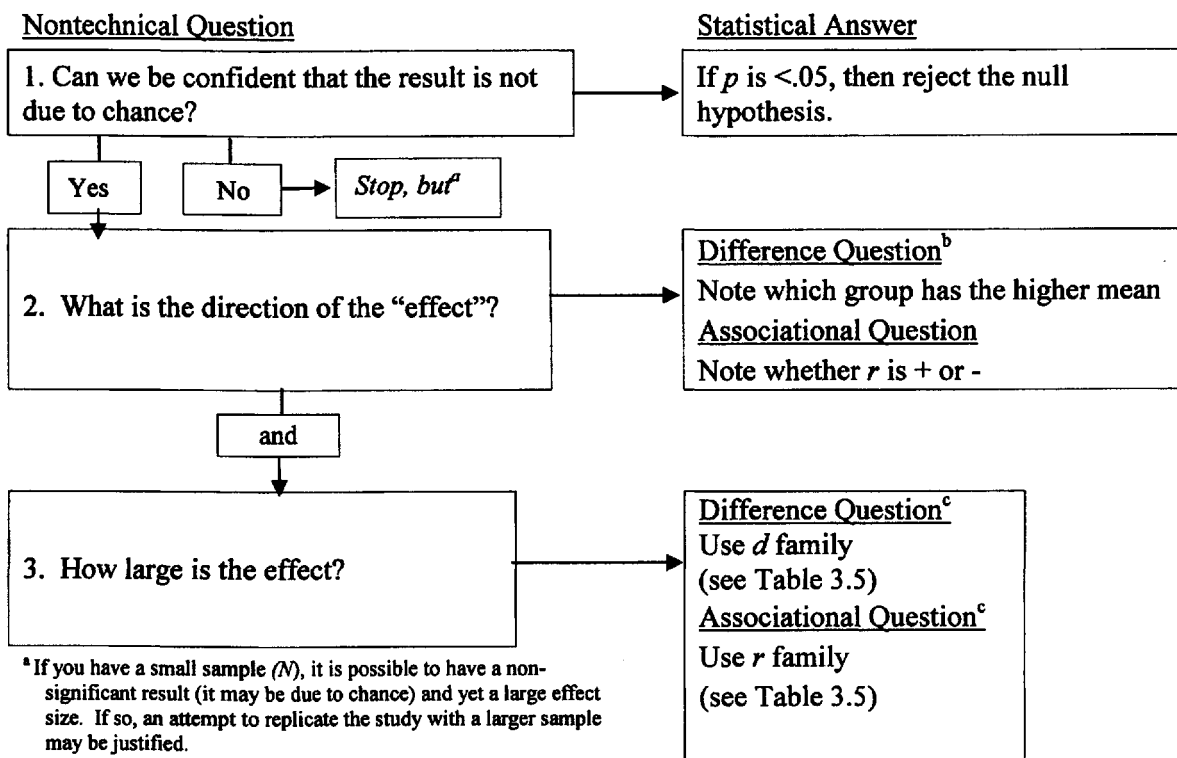


Fig. 3.4. Steps in the interpretation of an inferential statistic.

Second, what is the direction of the effect? Difference inferential statistics compare groups so it is necessary to state which group performed better. We discuss how to do this in chapters 8, 9 and 10. For associational inferential statistics (e.g., correlation), the sign is very important, so you must indicate

whether the association or relationship is positive or negative. We discuss how to interpret correlations in chapters 4, 5, and 6, as part of the chapters on reliability, factor analysis, and multiple regression.

Third, what is the size of the effect? You should include effect size, confidence intervals, or both in the description of your results. Unfortunately, SPSS does not always provide effect sizes and confidence intervals, so for some statistics we will have to compute or estimate the effect size by hand.

Fourth, but not shown in Fig. 3.4, the researcher or the consumer of the research should make a judgment about whether the result has practical or clinical significance or importance. To do so, they need to take into account the effect size, the costs of implementing change, and the probability and severity of any side effect or unintended consequences.

An Example of How to Select and Interpret Inferential Statistics

As a review, we now provide an extended example based on the HSB data. We will walk you through the process of identifying the variables, research questions, and approach, and then show how we selected appropriate statistics and interpreted the results.

Research problem. Suppose your research problem was to investigate *gender* and *math courses taken* and their relationship to *math achievement*.

Identification of the variables and their measurement. The research problem specifies three variables: *gender*, *math courses taken*, and *math achievement*. The latter appears to be the outcome or dependent variable. *Gender* and *math courses taken* are the independent or predictor variables. Because they occurred before the math exam, they are *presumed* to have an effect on *math achievement* scores. What is the level of measurement for these three variables? *Gender* is clearly dichotomous (male or female). *Math courses taken* has six ordered values, from 0 to 5 courses. These are scale data because there should be an approximately normal distribution of scores: Most students took some but not all of the math courses. Likewise, the *math achievement* test has many levels, with more scores somewhere in the middle than high or low. It is necessary to confirm that *math courses taken* and *math achievement* are at least approximately normally distributed by requesting that SPSS compute the skewness of each.

Research questions. There are a number of possible research questions that could be asked and statistics that could be used with these three variables. However, we will focus on three research questions and three inferential statistics because they answer this research problem and fit our earlier recommendations for good choices. First, we will discuss two basic research questions, given the above specification of the variables and their measurement. Then, we will discuss a complex research question that could be asked instead of research questions 1 and 2.

1. Is there a difference between male and female genders in their average *math achievement* scores?

Type of research question. Using the text, Fig. 3.1, and Table 3.1, you should see that the first question is phrased as a *basic difference question* because there are only two variables and the focus is a group difference (the difference between the male group and the female group).

Selection of an appropriate statistic. If you examine Table 3.1, you will see that the first question should be answered with an *independent samples t test* because (a) the independent variable has only two values (male and female), (b) the design is between groups (males and females form two independent

groups), and (c) the dependent variable (math achievement) is normal or scale data. We also would check other assumptions of the t test to be sure that they are not markedly violated.

Interpretation of the results for research question 1. Let's assume that about 50 students participated in the study and that $t = 2.05$. SPSS will give you the exact **Sig.**, but is $p < .05$ and, thus, t is statistically significant. (However, if you had 25 participants, this t would not have been significant because the t value necessary for statistical significance is influenced strongly by sample size. Small samples require a larger t to be significant.) Deciding whether the statistic is significant only means the result is unlikely to be due to chance. You still have to state the direction of the result and the effect size and/or the confidence interval. To **determine the direction**, we need to know the mean (average) math achievement scores for males and females. Let's assume, as is true for the HSB data, that males have the higher mean. If so, you can be quite confident that males in the population are at least a little better at math achievement, on average, than females, and you should state that males scored higher than females. If the difference was not statistically significant, it is best *not* to make any comment about which mean was higher because the difference could be due to chance. Likewise, if the difference was not significant, we recommend that you do not discuss or *interpret* the effect size. However, you should provide the means and standard deviations so that effect sizes could be computed if a researcher wanted to use this study in a meta-analysis.

Because the t was statistically significant, we would calculate d and discuss the **effect size** as shown earlier. First, figure the pooled (weighted average) standard deviation for male and female math achievement scores. Let's say that the difference between the means was 2.0 and the pooled standard deviation was 6.0; then d would be .33, a somewhat less than typical size effect. This means that the difference is less than typical of the statistically significant findings in the behavioral sciences. A d of .33 may or may not be a large enough difference to use for recommending programmatic changes (i.e., be practically significant).

Confidence intervals might help you decide if the difference in math achievement scores was large enough to have practical significance. For example, say you found (from the lower bound of the confidence interval) that you could only be confident that there was a .5 difference between males and females. Then you decide whether that is a big enough difference to justify, for example, a programmatic change.

2. Is there an association between *math courses taken* and *math achievement*?

Type of research question. This second question is phrased as a **basic associational question** because there are only two variables and both have many ordered levels. Thus, use Table 3.2 for this question.

Selection of an appropriate statistic. As you can see from Table 3.2, research question 2 should be answered with a **Pearson correlation** because both *math courses taken* and *math achievement* are normally distributed data.

Interpretation of the results for research question 2. The interpretation of r is based on similar decisions. If $r = .30$ (with 50 subjects), it would be statistically significant at the $p < .05$ level. If the r is statistically significant, you still need to discuss the direction of the correlation and effect size. Because the correlation is positive, we would say that students with a relative high number of *math courses taken* tend to perform at the high end on *math achievement* and those with few *math courses taken* tend to perform poorly on *math achievement*. The **effect size** of $r = .30$; that is, medium or typical. Note that if N were 25, $r = .30$ would not be significant. On the other hand, if N were 500 and $r = .30$, p would be

$< .0001$. With $N = 500$, even $r = .10$ would be statistically significant, indicating that you could be quite sure the association was not zero, but the effect size would be small or less than typical.

Complex research question and statistics. As stated above, there are advantages to considering the two independent variables (*gender* and *math courses taken*) together rather than separately as in research questions 1 and 2. There are at least two statistics that you will compute that can be used to consider *gender* and *math courses taken* together. The first is **multiple regression**, which is discussed in chapter 6. If you examine Table 3.4, you will see that with two (or more) independent variables that are scale and/or dichotomous and one dependent variable that is approximately normal (scale) data, an appropriate associational statistic would be multiple regression. A research question, which subsumes both questions 1 and 2 above, could be:

3. Is there a combination of *gender* and *math courses* that predicts *math achievement*?

Selection of an appropriate statistic. As just stated, multiple regression could be used to answer this question. As you can see in Table 3.4, multiple regression is appropriate because we are trying to predict a normally distributed variable (*math achievement*) from two independent variables. The independent or predictor variables are *math courses taken* (normal or scale) and *gender* (a dichotomous variable).

Based on our discussion of the general linear model (GLM), a **two-way factorial ANOVA** would be another statistic that could be used to consider both *gender* and *math courses* simultaneously. However, to use ANOVA, the many levels of *math courses taken* would have to be recoded into two or three levels (perhaps high, medium, and low). Because information is lost when you do such a recode, we would not recommend factorial ANOVA for this example. Another possible statistic to use for this example is **analysis of covariance (ANCOVA)** using *gender* as the independent variable and *math courses taken* as the covariate. We will demonstrate in chapter 8 how we can control for *gender* differences in the number of *math courses taken* by using *math courses* as a covariate.

Interpretation of the results for research question 3. We will discuss the interpretation of multiple regression results in chapter 6 and factorial ANOVA and ANCOVA in chapter 8. We will see that we will obtain more information about the relationships among these three variables by doing these complex statistics than by doing only the *t* test and correlation.

Review of Writing About Your Outputs

One of the goals of this book is to help you write a research report or thesis/dissertation using the SPSS outputs. In each of the following chapters, we will provide an **Interpretation** of each output, as well as an example of how you might write about and make a table from the results provided by the output. As a review of how to write about a *t* test and correlation, we have provided this section, which could be from a thesis based on the expanded HSB data used in the assignments in this book.

Before demonstrating how you might write about the results of research questions 1 and 2 above, we want to make several important points. Several books that will help you write a research paper and make appropriate tables are listed in For Further Reading at the back of this book. Note especially the APA manual (2001), Nicol and Pexman (1999), and Morgan, Reichart, and Harrison (2002). The examples below and in the following chapters are only one way to write about SPSS outputs. There are other good ways.

Based on your SPSS outputs, you should update the **Methods** chapter that you wrote as a proposal to include descriptive statistics about the demographics (e.g., gender, age, ethnicity) of the participants. You should add to your proposal evidence related to the reliability and validity of your measures/instruments with your sample. You also should include in your report whether statistical assumptions of the inferential statistics were met or how adjustments were made.

The **Results** chapter or section includes a description (but not a discussion) of the findings in words and tables. Your Results section should include the following numbers about each statistically significant finding (in a table or the text):

1. The value of the statistic (e.g., $t = 2.05$ or $r = .30$)
2. The degrees of freedom (often in parentheses) and for chi-square the N (e.g., $df = 2$, $N = 49$).
3. The p or Sig. Value (e.g., $p = .048$).
4. The direction of the finding (e.g., by stating which mean is larger or the sign of the correlation, if the statistic is significant).
5. An index of effect size from either the d family or the r family.

When not shown in a table, the above information should be provided in the text as shown below. In addition to the numerical information, describe your significant results in words, including the variables related, the direction of the finding, and an interpretive statement about the size/strength of the effect, perhaps based on Cohen (1988) or Table 3.5. Realize that these terms are only rough estimates of the magnitude of the “effect” based on what is typical in the behavioral sciences, but not necessarily your topic.

If your paper includes a table, it is usually not necessary or advisable to include all the details about the value of the statistic, degrees of freedom, and p in the text, because they are in the table. If you have a table, you must refer to it by number (e.g., Table 1) in the text and describe the main points, but do not repeat all of it or the table is not necessary. You should not describe, in the text, the direction of the finding or the effect size of nonsignificant findings, because the results could well be due to chance. The **Discussion** chapter or section puts the findings in context in regard to the research literature, theory, and the purposes of the study. You may also attempt to explain why the results turned out the way they did.

An Example of How to Write About the Results of a t Test and a Correlation

Results

For research question 1, there was a statistically significant difference between male and female students on math achievement, $t(48) = 2.05$, $p = .04$, $d = .33$. Males ($M = 14.70$, $SD = 5.50$) scored higher than females ($M = 12.70$, $SD = 6.50$), and the effect size was smaller than typical for this topic. The confidence interval for the difference between the means was .50 to 3.50 indicating that the difference could be as little as half a point, which is probably not a practically important difference, but it could also be as large as three and a half points.

For research question 2, there was a statistically significant positive correlation between math courses taken and math achievement $r(48) = .30$, $p = .03$. The positive correlation means that, in general, students who took more math courses tended to score higher on the math achievement test and students who did not take many math courses scored lower on math achievement. The effect size of $r = .30$ is considered medium or typical for this area of research.

We will present examples of how to write about the results of the complex statistics discussed in this book in the appropriate chapters. Note that measures of reliability (e.g., Cronbach alpha, discussed in

chapter 4) and principal components analysis (chapter 5) are usually discussed in the Methods section. Chapters 6-10 present complex statistics that might be used to answer your complex research questions.

Conclusion

After the above review, you should be ready to study each of the complex statistics in Tables 3.3 and 3.4 and learn more about their computation and interpretation. Hopefully this review has brought you up to speed. It would be wise for you to review this chapter, especially the tables and figures from time to time. If you do, you will have a good grasp of how the various statistics fit together, when to use them, and how to interpret the results. You will need this information to understand the chapters that follow.

Interpretation Questions

- 3.1 a) Is there only one appropriate statistic to use for each research design? Explain your answer.
- 3.2 When $p < .05$, what does this signify?
- 3.3 Interpret the following related to effect size:
a) $d = .25$ c) $R = .53$ e) $d = 1.15$
b) $r = .35$ d) $r = .13$ f) $\eta = .38$
- 3.4 The confidence interval of the difference between means was $-.30$ to 4.0 . Explain what this indicates.
- 3.5 What statistic would you use if you had two independent variables, income group ($< \$10,000$, $\$10,000$ - $\$30,000$, $> \$30,000$) and ethnic group (Hispanic, Caucasian, African-American), and one normally distributed dependent variable (self-efficacy at work). Explain.
- 3.6 What statistic would you use if you had one independent variable, geographic location (North, South, East, West), and one dependent variable (satisfaction with living environment, Yes or No?) Explain.
- 3.7 What statistic would you use if you had three normally distributed (scale) independent variables and one dichotomous independent variable (weight of participants, age of participants, height of participants and gender) and one dependent variable (positive self-image), which is normally distributed. Explain.
- 3.8 What statistic would you use if you had one between groups independent variable, one repeated measures independent variable, each with two levels, and one normally distributed dependent variable?
- 3.9 What statistic would you use if you had one, repeated measures, independent variable with two levels and one nominal dependent variable?
- 3.10 What statistic would you use if you had one, between groups, independent variable with four levels and three normally distributed dependent variables?
- 3.11 What statistic would you use if you had three normally distributed independent variables, one dichotomous independent variable, and one dichotomous dependent variable?

CHAPTER 4

Several Measures of Reliability

This assignment illustrates several of the methods for computing instrument reliability.

Internal consistency reliability, for multiple item scales. In this assignment, we will compute the most commonly used type of internal consistency reliability, Cronbach's coefficient **alpha**. This measure indicates the consistency of a multiple item scale. Alpha is typically used when you have several Likert-type items that are summed to make a composite score or **summated scale**. Alpha is based on the mean or average correlation of each item in the scale with every other item. In the social science literature, alpha is widely used, because it provides a measure of reliability that can be obtained from one testing session or one administration of a questionnaire. In problems 4.1, 4.2, and 4.3, you compute alphas for the three math attitude scales (*motivation, competence, and pleasure*) that items 1 to 14 were designed to assess.

Reliability for one score/measure. In problem 4.4, you compute a correlation coefficient to check the reliability of *visualization scores*. Several types of reliability can be illustrated by this correlation. If the *visualization retest* was each participant's score from retaking the test a month or so after they initially took it, then the correlation would be a measure of **test-retest reliability**. On the other hand, imagine that *visualization retest* was a score on an alternative/parallel or equivalent version of the visualization test; then this would be a measure of **equivalent forms reliability**. If two different raters scored the *visualization test* and *retest*, the correlation could be used as an index of **interrater reliability**. This latter type of reliability is needed when behaviors or answers to questions involve some degree of judgment (e.g., when there are open-ended questions or ratings based on observations).

Reliability for nominal variables. There are several other methods of computing interrater or interobserver reliability. Above, we mentioned using a correlation coefficient. In problem 4.5, **Cohen's kappa** is used to assess interobserver agreement when the data are nominal.

Assumptions for Measures of Reliability

When two or more measures, items, or assessments are viewed as measuring the same variable, reliability can be assessed. Reliability is used to indicate the extent to which the different items, measures, or assessments are consistent with one another (hopefully, in measuring that variable) and the extent to which each measure is free from measurement error. It is assumed that each item or score is composed of a true score measuring the underlying construct, plus error because there is error in the measurement of anything. Therefore, one assumption is that the measures or items will be related systematically to one another in a linear manner because they are believed to be measures of the same construct. In addition, because true error should not be correlated systematically with anything, a second assumption is that the errors (residual) for the different measures or assessments are uncorrelated. If errors are correlated, this means that the residual is not simply error; rather, the different measures not only have the proposed underlying variable in common, but they also have something else systematic in common and reliability estimates may be inflated. An example of a situation in which the assumption of uncorrelated errors might be violated would be when items all are parts of a cognitive test that is timed. The performance features that are affected by timing the test, in addition to the cognitive skills involved, might systematically affect responses to the items. The best way to determine whether part of the reliability score is due to these extraneous variables is by doing multiple types of reliability assessments (e.g., alternate forms and test-retest).

Conditions for Measures of Reliability

A condition that is necessary for measures of reliability is that all parts being used need to be comparable. If you use split-half, then both halves of the test need to be equivalent. If you use alpha (which we demonstrate in this chapter), then it is assumed that every item is measuring the same underlying construct. It is assumed that respondents should answer similarly on the parts being compared, with any differences being due to measurement error.

- Retrieve your data file: **hsbdataB**.

Problem 4.1: Cronbach's Alpha for the Motivation Scale

The motivation score is composed of six items that were rated on 4-point Likert scales, from very atypical (1) to very typical (4). Do these items go together (interrelate) well enough to add them together for future use as a composite variable labeled *motivation*?

4.1. What is the internal consistency reliability of the math attitude scale that we labeled *motivation*?

Note that you do not actually use the computed motivation scale score. The **Reliability** program uses the individual items to create the scale temporarily.

- Let's do reliability analysis for the motivation scale. Click on **Analyze => Scale => Reliability Analysis**. You should get a dialog box like Fig. 4.1.
- Now move the variables *item01*, *item04 reversed*, *item07*, *item08 reversed*, *item12*, and *item13* (the motivation questions) to the **Items** box. Be sure to use *item04 reversed* and *item08 reversed* (not *item04* and *item08*).
- Click on the **List item labels** box. Be sure the **Model** is **Alpha** (refer to Fig. 4.1).

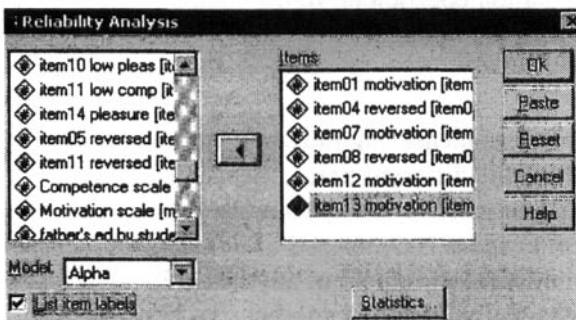


Fig. 4.1. Reliability analysis.

- Click on **Statistics** in the **Reliability Analysis** dialog box and you will see something similar to Fig. 4.2.
- Check the following items: **Item**, **Scale**, **Scale if item deleted** (under Descriptives), **Correlations** (under Inter-Item), **Means**, and **Correlations** (under Summaries).
- Click on **Continue** then **OK**. Compare your syntax and output to Output 4.1.

Chapter 4 – Several Measures of Reliability

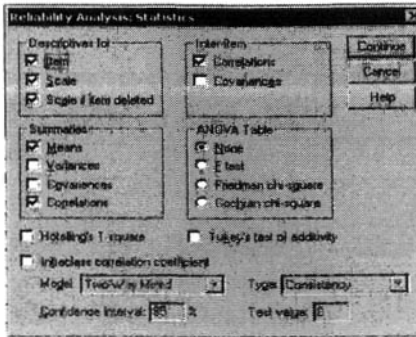


Fig. 4.2. Reliability analysis: Statistics.

Output 4.1: Cronbach's Alpha for the Math Attitude Motivation Scale

```
RELIABILITY
/VARIABLES=item01 item04r item07 item08r item12 item13
/FORMAT=LABELS
/SCALE(ALPHA)=ALL/MODEL=ALPHA
/STATISTICS=DESCRIPTIVE SCALE CORR
/SUMMARY=TOTAL MEANS CORR .
```

Reliability

Warnings

The covariance matrix is calculated and used in the analysis.

Case Processing Summary

		N	%
Cases	Valid	73	97.3
	Excluded(a)	2	2.7
	Total	75	100.0

a Listwise deletion based on all variables in the procedure.

Inter-Item Correlation Matrix

	item01 motivation	item04 reversed	item07 motivation	item08 reversed	item12 motivation	item13 motivation
item01 motivation	1.000	.250	.464	.296	.179	.167
item04 reversed	.250	1.000	.551	.576	.382	.316
item07 motivation	.464	.551	1.000	.587	.344	.360
item08 reversed	.296	.576	.587	1.000	.389	.311
item12 motivation	.179	.382	.344	.389	1.000	.603
item13 motivation	.167	.316	.360	.311	.603	1.000

The covariance matrix is calculated and used in the analysis.

SPSS for Intermediate Statistics

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.791	.790	6

This is the Cronbach's alpha coefficient.

Item Statistics

	Mean	Std. Deviation	N
item01 motivation	2.96	.934	73
item04 reversed	2.82	.918	73
item07 motivation	2.75	1.064	73
item08 reversed	3.05	.911	73
item12 motivation	2.99	.825	73
item13 motivation	2.67	.800	73

Descriptive statistics for each item.

Summary Item Statistics

	Mean	Minimum	Maximum	Range	Maximum / Minimum	Variance	N of Items
Item Means	2.874	2.671	3.055	.384	1.144	.022	73
Inter-Item Covariances	.385	.167	.603	.436	3.604	.020	73

The covariance matrix is calculated and used in the analysis.

A key table.

If correlation is moderately high to high (e.g., .40+), then the item will make a good component of a summated rating scale. You may want to modify or delete items with low correlations. Item 1 is a bit low if one wants to consider all items to measure the same thing. Note that the alpha increases a little if Item 1 is deleted.

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
item01 motivation	14.29	11.402	.378	.226	.798
item04 reversed	14.42	10.303	.596	.536	.746
item07 motivation	14.49	9.170	.676	.519	.723
item08 reversed	14.19	10.185	.627	.588	.738
item12 motivation	14.26	11.140	.516	.866	.765
item13 motivation	14.58	11.442	.476	.853	.774

Scale Statistics

Mean	Variance	Std. Deviation	N of Items
17.25	14.661	3.829	6

The average for the 6-item summated scale score for the 73 participants.

Interpretation of Output 4.1

The first section lists the items (item01, etc.) that you requested to be included in this scale and their labels. Selecting **List item labels** in Fig. 4.1 produced it. Next is a table of descriptive statistics for each item, produced by checking **Item** in Fig. 4.2. The third table is a matrix showing the **inter-item correlations** of every item in the scale with every other item. The next table shows the number of participants with no missing data on these variables (73), and three one-line tables providing summary descriptive statistics for a) the **Scale** (sum of the six motivation items), b) the **Item Means**, produced under summaries, and c) **correlations**, under summaries. The latter two tell you, for example, the average, minimum, and maximum of the item means and of the inter-item correlations.

The final table, which we think is the most important, is produced if you check the **Scale if item deleted** under **Descriptives for** in the dialog box displayed in Fig. 4.2. This table, labeled **Item-total Statistics**, provides five pieces of information for each item in the scale. The two we find most useful are the **Corrected Item-Total Correlation** and the **Alpha if Item Deleted**. The former is the correlation of each specific item with the sum/total of the other items in the scale. If this correlation is moderately high or high, say .40 or above, the item is probably at least moderately correlated with most of the other items and will make a good component of this summated rating scale. Items with lower item-total correlations do not fit into this scale as well, psychometrically. If the item-total correlation is negative or too low (less than .30), it is wise to examine the item for wording problems and conceptual fit. You may want to modify or delete such items. You can tell from the last column what the alpha would be if you deleted an item. Compare this to the alpha for the scale with all six items included, which is given at the bottom of the output. Deleting a poor item will usually make the alpha go up, but it will usually make only a small difference in the alpha, unless the scale has only a few items (e.g., fewer than five) because alpha is based on the number of items as well as their average intercorrelations.

In general, you will use the unstandardized **alpha**, unless the items in the scale have quite different means and standard deviations. For example, if we were to make a summated scale from *math achievement*, *grades*, and *visualization*, we would use the standardized alpha. As with other reliability coefficients, alpha should be above .70; however, it is common to see journal articles where one or more scales have somewhat lower alphas (e.g., in the .60-.69 range), especially if there are only a handful of items in the scale. A very high alpha (e.g., greater than .90) probably means that the items are repetitious or that you have more items in the scale than are really necessary for a reliable measure of the concept.

Example of How to Write About Problems 4.1

Method

To assess whether the six items that were summed to create the motivation score formed a reliable scale, Cronbach's alpha was computed. The alpha for the six items was .79, which indicates that the items form a scale that has reasonable internal consistency reliability. Similarly, the alpha for the competence scale (.80) indicated good internal consistency, but the .69 alpha for the pleasure scale indicated minimally adequate reliability.

Problems 4.2 & 4.3: Cronbach's Alpha for the Competence and Pleasure Scales

Again, is it reasonable to add these items together to form summated measures of the concepts of *competence* and *pleasure*?

4.2. What is the internal consistency reliability of the *competence scale*?

4.3. What is the internal consistency reliability of the *pleasure scale*?

Let's repeat the same steps as before except check the reliability of the following scales and then compare your output to 4.2 and 4.3.

- For the *competence scale*, use *item03*, *item05 reversed*, *item09*, *item11 reversed*
- For the *pleasure scale*, use *item02*, *item06 reversed*, *item10 reversed*, *item14*

Output 4.2: Cronbach's Alpha for the Math Attitude Competence Scale

```
RELIABILITY
/VARIABLES=item03 item05r item09 item11r
/FORMAT=LABELS
/SCALE (ALPHA)=ALL/MODEL=ALPHA
/STATISTICS=DESCRIPTIVE SCALE CORR
/SUMMARY=TOTAL MEANS CORR .
```

Reliability

Warnings

The covariance matrix is calculated and used in the analysis.

Case Processing Summary

		N	%
Cases	Valid	73	97.3
	Excluded (a)	2	2.7
	Total	75	100.0

a Listwise deletion based on all variables in the procedure.

Inter-Item Correlation Matrix

	item03 competence	item05 reversed	item09 competence	item11 reversed
item03 competence	1.000	.742	.325	.513
item05 reversed	.742	1.000	.340	.609
item09 competence	.325	.340	1.000	.399
item11 reversed	.513	.609	.399	1.000

The covariance matrix is calculated and used in the analysis.

Chapter 4 – Several Measures of Reliability

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.796	.792	4

Item Statistics

	Mean	Std. Deviation	N
item03 competence	2.82	.903	73
item05 reversed	3.41	.940	73
item09 competence	3.32	.762	73
item11 reversed	3.63	.755	73

Summary Item Statistics

	Mean	Minimum	Maximum	Range	Maximum / Minimum	Variance	N of Items
Item Means	3.295	2.822	3.630	.808	1.286	.117	73
Inter-Item Covariances	.488	.325	.742	.417	2.281	.025	73

The covariance matrix is calculated and used in the analysis.

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
item03 competence	10.36	3.844	.680	.769	.706
item05 reversed	9.77	3.570	.735	.872	.675
item09 competence	9.86	5.092	.405	.548	.832
item11 reversed	9.55	4.473	.633	.842	.736

Scale Statistics

Mean	Variance	Std. Deviation	N of Items
13.18	7.065	2.658	4

Output 4.3: Cronbach's Alpha for the Math Attitude Pleasure Scale

```

RELIABILITY
/VARIABLES=item02 item06r item10r item14
/FORMAT=LABELS
/SCALE(ALPHA)=ALL/MODEL=ALPHA
/STATISTICS=DESCRIPTIVE SCALE CORR
/SUMMARY=TOTAL MEANS CORR .

```

Reliability**Warnings**

The covariance matrix is calculated and used in the analysis.

Case Processing Summary

		N	%
Cases	Valid	75	100.0
	Excluded (a)	0	.0
	Total	75	100.0

a Listwise deletion based on all variables in the procedure.

Inter-Item Correlation Matrix

	item02 pleasure	item06 reversed	item10 reversed	item14 pleasure
item02 pleasure	1.000	.285	.347	.504
item06 reversed	.285	1.000	.203	.461
item10 reversed	.347	.203	1.000	.436
item14 pleasure	.504	.461	.436	1.000

The covariance matrix is calculated and used in the analysis.

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.688	.704	4

Item Statistics

	Mean	Std. Deviation	N
item02 pleasure	3.5200	.90584	75
item06 reversed	2.5733	.97500	75
item10 reversed	3.5867	.73693	75
item14 pleasure	2.8400	.71735	75

Chapter 4 – Several Measures of Reliability

Summary Item Statistics

	Mean	Minimum	Maximum	Range	Maximum / Minimum	Variance	N of Items
Item Means	3.130	2.573	3.587	1.013	1.394	.252	75
Inter-Item Covariances	.373	.203	.504	.301	2.488	.012	75

The covariance matrix is calculated and used in the analysis.

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
item02 pleasure	9.0000	3.405	.485	.880	.615
item06 reversed	9.9467	3.457	.397	.881	.685
item10 reversed	8.9333	4.090	.407	.929	.662
item14 pleasure	9.6800	3.572	.649	.979	.528

Scale Statistics

Mean	Variance	Std. Deviation	N of Items
12.5200	5.848	2.41817	4

Problem 4.4: Test-Retest Reliability Using Correlation

4.4. Is there support for the test-retest reliability of the *visualization test scores*?

Let's do a Pearson r for *visualization test* and *visualization retest* scores.

- Click on **Analyze** => **Correlate** => **Bivariate**.
- Move variables *visualization* and *visualization retest* into the variable box.
- Do not select **flag significant correlations** because reliability coefficients should be positive and greater than .70; statistical significance is not enough.
- Click on **Options**.
- Click on **Means and Standard deviations**.
- Click on **Continue** and **OK**. Do your syntax and output look like Output 4.4?

Output 4.4: Pearson r for the Visualization Score

```

CORRELATIONS
/VARIABLES=visual visual2
/PRINT=TWOTAIL SIG
/STATISTICS DESCRIPTIVES
/MISSING=PAIRWISE .

```

Correlations

Descriptive Statistics

	Mean	Std. Deviation	N
visualization test	5.2433	3.91203	75
visualization retest	4.5467	3.01816	75

Correlations

		visualization test	visualization retest
visualization test	Pearson Correlation	1	.885
	Sig. (2-tailed)	.	.000
	N	75	75
visualization retest	Pearson Correlation	.885	1
	Sig. (2-tailed)	.000	
	N	75	75

Correlation between the first and second visualization test. It should be high, not just significant.

75 participants have both visualization scores.

Interpretation of Output 4.4

The first table provides the descriptive statistics for the two variables, *visualization test* and *visualization retest*. The second table indicates that the correlation of the two visualization scores is very high ($r = .89$) so there is strong support for the test-retest reliability of the *visualization* score. This correlation is significant, $p < .001$, but we are not as concerned about the significance as we are having the correlation be high.

Example of How to Write About Problem 4.4

Method

A Pearson's correlation was computed to assess test-retest reliability of the visualization test scores, $r(75) = .89$. This indicates that there is good test-retest reliability.

Problem 4.5: Cohen's Kappa With Nominal Data

When we have two categorical variables with the *same* values (usually two raters' observations or scores using the same codes), we can compute Cohen's kappa to check the reliability or agreement between the measures. Cohen's kappa is preferable over simple percentage agreement because it corrects for the probability that raters will agree due to chance alone. In the *hsbdataB*, the variable *ethnicity* is the ethnicity of the student as reported in the school records. The variable *ethnicity reported by student* is the ethnicity of the student reported by the student. Thus, we can compute Cohen's kappa to check the agreement between these two ratings. The question is, How reliable are the school's records?

4.5. What is the reliability coefficient for the *ethnicity* codes (based on school records) and *ethnicity reported by the student*?

Chapter 4 – Several Measures of Reliability

To compute the kappa:

- Click on **Analyze => Descriptive Statistics => Crosstabs**.
- Move *ethnicity* to the **Rows** box and *ethnic reported by student* to the **Columns** box.
- Click on **Kappa** in the **Statistics** dialog box.
- Click on **Continue** to go back to the Crosstabs dialog window.
- Then click on **Cells** and request the **Observed** cell counts and **Total** under **Percentages**.
- Click on **Continue** and then **OK**. Compare your syntax and output to Output 4.5.

Output 4.5: Cohen's Kappa With Nominal Data

```
CROSSTABS
  /TABLES=ethnic BY ethnic2
  /FORMAT=AVALUE TABLES
  /STATISTIC=KAPPA
  /CELLS= COUNT TOTAL .
```

Crosstabs

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
ethnicity * ethnicity reported by student	71	94.7%	4	5.3%	75	100.0%

ethnicity * ethnicity reported by student Crosstabulation

			ethnicity reported by student				Total
			Euro-Amer	African-Amer	Latino-Amer	Asian-Amer	
ethnicity	Euro-Amer	Count	40	1	0	0	41
		% of Total	56.3%	1.4%	.0%	.0%	57.7%
	African-Amer	Count	2	11	1	0	14
		% of Total	2.8%	15.5%	1.4%	.0%	19.7%
	Latino-Amer	Count	0	1	8	0	9
		% of Total	.0%	1.4%	11.3%	.0%	12.7%
	Asian-Amer	Count	0	1	0	6	7
		% of Total	.0%	1.4%	.0%	8.5%	9.9%
Total	Count	42	14	9	6	71	
	% of Total	59.2%	19.7%	12.7%	8.5%	100.0%	

One of six disagreements. They are in squares off the diagonal.

Agreements between school records and students' answers are shown in circles on the diagonal.

As a measure of reliability, kappa should be high (usually $\geq .70$) not just statistically significant.

Symmetric Measures

	Value	Asymp. Std. Error ^a	Approx. T ^b	Approx. Sig.
Measure of Agreement Kappa	.858	.054	11.163	.000
N of Valid Cases	71			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Interpretation of Output 4.5

The **Case Processing Summary** table shows that 71 students have data on both variables, and 4 students have missing data. The **Crosstabulation** table of *ethnicity* and *ethnicity reported by student* is next. The cases where the school records and the student agree are on the diagonal and circled. There are 65 ($40 + 11 + 8 + 6$) students with such agreement or consistency. The **Symmetric Measures** table shows that $\text{kappa} = .86$, which is very good. Because **kappa** is a measure of reliability, it usually should be .70 or greater. It is not enough to be significantly greater than zero. However, because it corrects for chance, it tends to be somewhat lower than some other measures of interobserver reliability, such as percentage agreement.

Example of How to Write About Problem 4.5

Method

Cohen's kappa was computed to check the reliability of reports of student ethnicity by the student and by school records. The resulting kappa of .86 indicates that both school records and students' reports provide similar information about students' ethnicity.

Interpretation Questions

- 4.1. Using Outputs 4.1, 4.2, and 4.3, make a table indicating the mean interitem correlation and the alpha coefficient for each of the scales. Discuss the relationship between mean interitem correlation and alpha, and how this is affected by the number of items.
- 4.2. For the *competence scale*, what item has the *lowest* corrected item-total correlation? What would be the alpha if that item were deleted from the scale?
- 4.3. For the *pleasure scale* (Output 4.3), what item has the highest item-total correlation? Comment on how alpha would change if that item were deleted.
- 4.4. Using Output 4.4: a) What is the reliability of the *visualization* score? b) Is it acceptable? c) As indicated above, this assessment of reliability could have been used to indicate test-retest reliability, alternate forms reliability, or interrater reliability, depending on exactly how the *visualization* and *visualization retest* scores were obtained. Select two of these types of reliability, and indicate what information about the measure(s) of *visualization* each would

Chapter 4 – Several Measures of Reliability

provide. d) If this were a measure of interrater reliability, what would be the procedure for measuring *visualization* and *visualization retest* score?

- 4.5 Using Output 4.5: What is the interrater reliability of the ethnicity codes? What does this mean?

Extra SPSS Problems

The extra SPSS problems at the end of this, and each of the following chapters, use data sets provided by SPSS. The name of the SPSS data set is provided in each problem.

- 4.1 Using the *satisf.sav* data, determine the internal consistency reliability (Cronbach's coefficient alpha) of a proposed six-item satisfaction scale. Use the price, variety, organization, service, item quality, and overall satisfaction items, which are 5-point Likert-type ratings. In other words, do the six satisfaction items interrelate well enough to be used to create a composite score for satisfaction? Explain.
- 4.2 A judge from each of seven different countries (Russia, U.S., etc.) rated 300 participants in an international competition on a 10-point scale. In addition, an "armchair enthusiast" rated the participants. What is the interrater reliability of the armchair enthusiast (judge 8) with each of the other seven judges? Use the *judges.sav* data. Comment.
- 4.3 Two consultants rated 20 sites for a future research project. What is the level of agreement or interrater reliability (using Cohen's Kappa) between the raters? Use the *site.sav* data. Comment.

CHAPTER 5

Exploratory Factor Analysis and Principal Components Analysis

Exploratory factor analysis (EFA) and principal components analysis (PCA) both are methods that are used to help investigators represent a large number of relationships among interval-level variables in a simpler (more parsimonious) way. Both of these approaches allow the computer to determine which, of a fairly large set of items, “hang together” as a group, or are answered most similarly by the participants. A related approach, **confirmatory factor analysis**, in which one tests very specific models of how variables are related to underlying constructs (conceptual variables), is beyond the scope of this book and will not be discussed.

The primary difference, conceptually, between **exploratory factor analysis** and **principal components analysis** is that in EFA, one postulates that there is a smaller set of unobserved (latent) variables or constructs that underlie the variables that actually were observed or measured; whereas, in PCA, one is simply trying to mathematically derive a relatively small number of variables to use to convey as much of the information in the observed/measured variables as possible. In other words, EFA is directed at *understanding* the relations among variables by understanding the constructs that underlie them, whereas PCA is simply directed toward enabling one to use fewer variables to provide the same information that one would obtain from a larger set of variables. There are actually a number of different ways of computing factors for factor analysis; in this chapter, we will only use one of these methods, **principal axis factor analysis** (PA). We selected this approach because it is highly similar mathematically to PCA. The primary difference, computationally, between PCA and PA is that in the former, the analysis typically is performed on an ordinary correlation matrix, complete with the correlations of each item or variable with itself, whereas in PA factor analysis, the correlation matrix is modified such that the correlations of each item with itself are replaced with a “communality”—a measure of that item’s relation to all other items (usually a squared multiple correlation). Thus, PCA is trying to reproduce all information (variance and covariance) associated with the set of variables, whereas PA factor analysis is directed at understanding only the covariation among variables.

Conditions for Exploratory Factor Analysis and Principal Components Analysis

There are two main conditions necessary for factor analysis and principal components analysis. The first is that there need to be relationships between the variables. Further, the larger the sample size, especially in relation to the number of variables, the more reliable the resulting factors usually are. Sample size is less crucial for factor analysis to the extent that the communalities of items with the other items are high, or at least relatively high and variable. Ordinary principal axis factor analysis should never be done if the number of items/variables is greater than the number of participants.

Assumptions for Exploratory Factor Analysis and Principal Components Analysis

The methods of extracting factors and components that are used in this book do not make strong distributional assumptions; normality is important only to the extent that skewness or outliers affect the observed correlations or if significance tests are performed (which is rare for EFA and PCA). The normality of the distributions can be checked by computing the skewness statistic. Maximum likelihood estimation, which we will not cover, does require multivariate normality—the variables need to be normally distributed and the joint distribution of all the variables should be normal. Because both principal axis factor analysis and principal components analysis are based on correlations, independent sampling is required and the variables should be related to each other (in pairs) in a linear fashion. The assumption of linearity can be assessed with matrix scatterplots, as shown in chapter 2. Finally, at least

many of the variables should be correlated at a moderate level. Factor analysis and principal components analysis are seeking to explain or reproduce the correlation matrix, which would not be a sensible thing to do if the correlations all hover around zero. Bartlett's test of sphericity addresses this assumption. Moreover, if correlations are too high, this may cause problems with obtaining a mathematical solution to the factor analysis problem.

- Retrieve your data file: **hsbdataB.sav**

Problem 5.1: Factor Analysis on Math Attitude Variables

In Problem 5.1, we perform a principal axis factor analysis on the math attitude variables. Factor analysis is more appropriate than PCA when one has the belief that there are latent variables underlying the variables or items one measured. In this example, we have beliefs about the constructs underlying the math attitude questions; we believe that there are three constructs: *motivation*, *competence*, and *pleasure*. Now, we want to see if the items that were written to index each of these constructs actually do “hang together.” That is, we wish to determine empirically whether participants’ responses to the motivation questions are more similar to each other than to their responses to the competence items, and so on. This is considered exploratory factor analysis even though we have some ideas about the structure of the data because our hypotheses regarding the model are not very specific; we do not have specific predictions about the size of the relation of each observed variable to each latent variable, etc.

5.1 Run a factor analysis using the principal axis factoring method and specify the number of factors to be three (because our conceptualization is that there are three math attitude scales or factors: *motivation*, *competence*, and *pleasure*).

- **Analyze => Data Reduction => Factor** to get Fig. 5.1.
- Next, select the variables *item01* through *item14*. Do not include *item04r* or any of the other reversed items because we are including the unreversed versions of those same items.

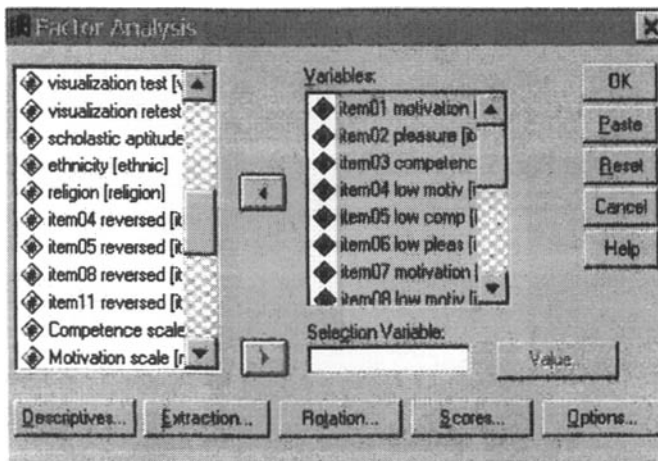


Fig. 5.1. Factor analysis.

- Now click on **Descriptives** to produce Fig. 5.2.
- Then click on the following: **Initial solution** (under **Statistics**), **Coefficients**, **Determinant**, **KMO** and **Bartlett's test of sphericity** (under **Correlation Matrix**).

- Click on **Continue**.

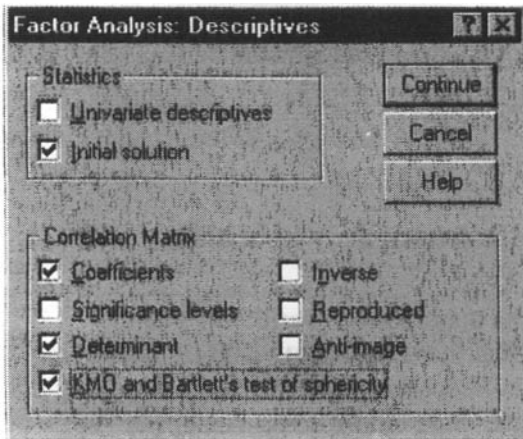


Fig. 5.2. Factor analysis: Descriptives.

- Next, click on **Extraction** at the bottom of Fig. 5.1. This will give you Fig. 5.3.
- Select **Principal axis factoring** from the **Methods** pull-down.
- **Unclick Unrotated factor solution** (under **Display**). Some researchers also would check the **Scree plot** box, but we decided to do that only in Problem 5.2.
- Check **Number of factors** under **Extract** and type **3** in the box. This setting instructs the computer to extract three math attitude factors.
- Click on **Continue**.

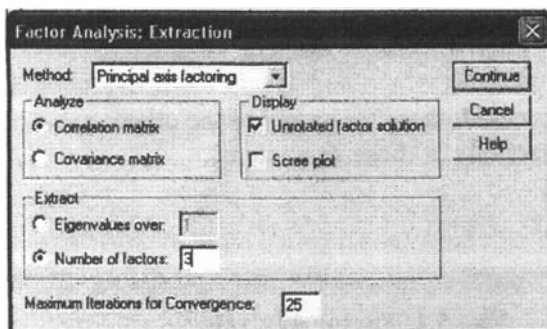


Fig. 5.3. Extraction method to produce principal axis factoring.

- Now click on **Rotation** in Fig. 5.1, which will give you Fig. 5.4.
- Click on **Varimax**, then make sure **Rotated solution** is also checked.
- Click on **Continue**.

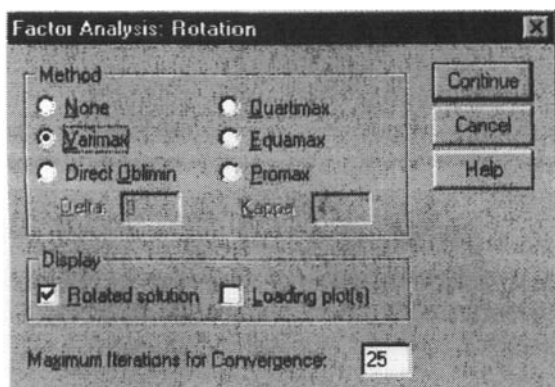


Fig. 5.4. Factor analysis: Rotation.

Chapter 5 – Exploratory Factor Analysis and Principal Components Analysis

- Next, click on **Options**, which will give you Fig. 5.5.
- Click on **Sorted by size**.
- Click on **Suppress absolute values less than** and type **.3** (point 3) in the box (see Fig. 5.5). Suppressing small factor loadings makes the output easier to read.
- Click on **Continue** then **OK**. Compare Output 5.1 to your output and syntax.

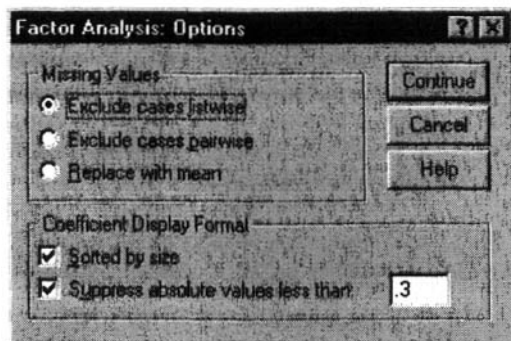
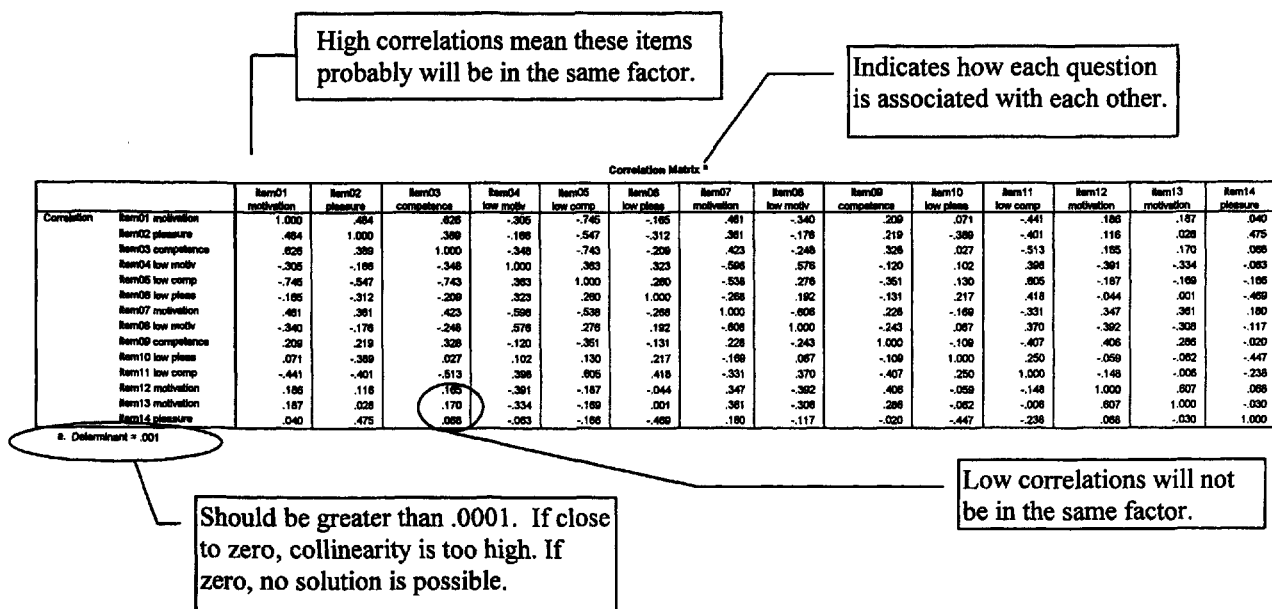


Fig. 5.5. Factor analysis: Options.

Output 5.1: Factor Analysis for Math Attitude Questions

```

FACTOR
/VARIABLES item01 item02 item03 item04 item05 item06 item07 item08 item09 item10 item11 item12
item13 item14 /MISSING
LISTWISE /ANALYSIS item01 item02 item03 item04 item05 item06 item07 item08 item09 item10 item11
item12 item13 item14
/PRINT INITIAL CORRELATION DET KMO ROTATION
/FORMAT SORT BLANK(.3)
/CRITERIA FACTORS(3) ITERATE(25)
/EXTRACTION PAF
/CRITERIA ITERATE(25)
/ROTATION VARIMAX
/METHOD=CORRELATION.
    
```



Tests of assumptions.

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.770
Bartlett's Test of Sphericity	Approx. Chi-Square	433.486
	df	91
	Sig.	.000

Should be greater than .70 indicating sufficient items for each factor.

Should be significant (less than .05), indicating that the correlation matrix is significantly different from an identity matrix, in which correlations between variables are all zero.

Communalities

	Initial
item01 motivation	.660
item02 pleasure	.542
item03 competence	.598
item04 low motiv	.562
item05 low comp	.772
item06 low pleas	.382
item07 motivation	.607
item08 low motiv	.533
item09 competence	.412
item10 low pleas	.372
item11 low comp	.591
item12 motivation	.499
item13 motivation	.452
item14 pleasure	.479

These communalities represent the relation between the variable and all other variables (i.e., the squared multiple correlation between the item and all other items).

Extraction Method: Principal Axis Factoring.

Chapter 5 – Exploratory Factor Analysis and Principal Components Analysis

Eigenvalues refer to the variance explained or accounted for.

Total Variance Explained

Percent of variance for each component before and after rotation.

Factor	Initial Eigenvalues			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4.888	34.916	34.916	3.017	21.549	21.549
2	2.000	14.284	49.200	2.327	16.621	38.171
3	1.613	11.519	60.719	1.784	12.746	50.917
4	1.134	8.097	68.816			
5	.904	6.459	75.275			
6	.716	5.113	80.388			
7	.577	4.125	84.513			
8	.461	3.293	87.806			
9	.400	2.857	90.664			
10	.379	2.710	93.374			
11	.298	2.126	95.500			
12	.258	1.846	97.346			
13	.217	1.551	98.897			
14	.154	1.103	100.000			

Extraction Method: Principal Axis Factoring.

Half of the variance is accounted for by the first three factors.

Factor Matrix^a

a. 3 factors extracted. 12 iterations required.

Rotated Factor Matrix^a

	Factor		
	1	2	3
item05 low comp	-.897		
item03 competence	.780		
item01 motivation	.777		
item11 low comp	-.572		.355
item12 motivation		.721	
item13 motivation		.667	
item08 low motiv		-.619	
item04 low motiv		-.601	
item07 motivation	.412	.585	
item09 competence		.332	
item14 pleasure			-.797
item10 low pleas			.580
item02 pleasure	.487		-.535
item06 low pleas			.515

The items cluster into these three groups defined by high loadings.

Extraction Method: Principal Axis Factoring.

Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 5 iterations.

Factor Transformation Matrix

Factor	1	2	3
1	.747	.552	-.370
2	-.162	.692	.704
3	.645	-.466	.606

We will ignore this; it was used to convert the initial factor matrix into the rotated factor matrix.

Extraction Method: Principal Axis Factoring.

Rotation Method: Varimax with Kaiser Normalization.

Interpretation of Output 5.1

The factor analysis program generates a number of tables depending on which options you have chosen. The first table in Output 5.1 is a **correlation matrix** showing how each of the 14 items is associated with each of the other 13. Note that some of the correlations are high (e.g., + or - .60 or greater) and some are low (i.e., near zero). The high correlations indicate that two items are associated and will probably be grouped together by the factor analysis.

Next, several assumptions are tested. The **determinant** (located under the correlation matrix) should be more than .00001. Here, it is .001 so this assumption is met. If the determinant is zero, then a factor analytic solution can not be obtained, because this would require dividing by zero. This would mean that at least one of the items can be understood as a linear combination of some set of the other items. The **Kaiser-Meyer-Olkin (KMO)** measure should be greater than .70, and is inadequate if less than .50. The KMO test tells one whether or not enough items are predicted by each factor. The **Bartlett** test should be significant (i.e., a significance value of less than .05); this means that the variables are correlated highly enough to provide a reasonable basis for factor analysis.

The **Total Variance Explained** table shows how the variance is divided among the 14 possible factors. Note that four factors have **eigenvalues** (a measure of explained variance) greater than 1.0, which is a common criterion for a factor to be useful. When the eigenvalue is less than 1.0, this means that the factor explains less information than a single item would have explained. Most researchers would not consider the information gained from such a factor to be sufficient to justify keeping that factor. Thus, if you had not specified otherwise, the computer would have looked for the best four-factor solution by "rotating" four factors. Because we specified that we wanted only three factors rotated, only three will be rotated.

For this and all analyses in this chapter, we will use an *orthogonal* rotation (varimax). This means that the final factors will be as uncorrelated as possible with each other. As a result, we can assume that the information explained by one factor is independent of the information in the other factors. We rotate the factors so that they are easier to interpret. Rotation makes it so that, as much as possible, different items are explained or predicted by different underlying factors, and each factor explains more than one item. This is a condition called simple structure. Although this is the *goal* of rotation, in reality, this is not always achieved. One thing to look for in the **Rotated Matrix** of factor loadings is the extent to which simple structure is achieved.

The **Rotated Factor Matrix** table, which contains these loadings, is key for understanding the results of the analysis. Note that the computer has sorted the 14 math attitude questions (*item 01* to *item 14*) into three overlapping groups of items, each which has a loading of $|\geq .30|$ or higher ($|\geq .30|$ means the absolute value, or value without considering the sign, is greater than .30). Actually, every item has some loading from every factor, but there are blanks in the matrix where weights were less than $|\geq .30|$. Within each factor (to the extent possible), the items are sorted from the one with the highest factor weight or loading

for that factor (i.e., *item 05* for factor 1, with a loading of $-.897$) to the one with the lowest loading on that first factor (*item 07*). Loadings resulting from an orthogonal rotation are correlation coefficients of each item with the factor, so they range from -1.0 through 0 to $+1.0$. A negative loading just means that the question needs to be interpreted in the opposite direction from the way it is written for that factor (e.g., *item 05* “I am a little slow catching on to new topics in math” has a negative loading from the competence factor, which indicates that the people scoring higher on this item are lower in competence). Usually, factor loadings lower than $|.30|$ are considered low, which is why we suppressed loadings less than $|.30|$. On the other hand, loadings of $|.40|$ or greater are typically considered high. This is just a guideline, however, and one could set the criterion for “high” loadings as low as $.30$ or as high as $.50$. Setting the criterion lower than $.30$ or higher than $.50$ would be very unusual.

The investigator should examine the content of the items that have high loadings from each factor to see if they fit together conceptually and can be named. Items 5, 3, and 11 were intended to reflect a perception of *competence* at math, so the fact that they all have strong loadings from the same factor provides some support for their being conceptualized as pertaining to the same construct. On the other hand, *item 01* was intended to measure *motivation* for doing math, but it is highly related to this same *competence* factor. In retrospect, one can see why this item could also be interpreted as competence. The item reads, “I practice math skills until I can do them well.” Unless one felt one could do math problems well, this would not be true. Likewise, *item 02*, “I feel happy after solving a hard problem,” although intended to measure *pleasure* at doing math (and having its strongest loading there), might also reflect competence at doing math, in that again one could not endorse this item unless one had solved hard problems, which one could only do if one were good at math. On the other hand, *item 09*, which was originally conceptualized as a competence item, had no really strong loadings. *Item 11*, as mentioned earlier, had a high loading for the first factor, as expected. However, it also had a moderate loading for Factor 3, which seems to be a (low) pleasure factor. This item reads, “I have some difficulties doing math as well as other kids my age.” Can you think of why this might be related to low *pleasure*?

Every item has a weight or loading from every factor, but in a “clean” factor analysis almost all of the loadings that are not in the circles that we have drawn on the **Rotated Factor Matrix** will be low (blank or less than $|.40|$). The fact that both Factors 1 and 3 load highly on *item 02* and fairly highly on *item 11*, and the fact that Factors 1 and 2 both load highly on *item 07* is common but undesirable, in that one wants only one factor to predict each item.

Example of How to Write About Problem 5.1

Results

Principal axis factor analysis with varimax rotation was conducted to assess the underlying structure for the fourteen items of the Math Motivation Questionnaire. Three factors were requested, based on the fact that the items were designed to index three constructs: motivation, competence, and pleasure. After rotation, the first factor accounted for 21.5% of the variance, the second factor accounted for 16.6%, and the third factor accounted for 12.7%. Table 5.1 displays the items and factor loadings for the rotated factors, with loadings less than $.40$ omitted to improve clarity.

Table 5.1
Factor Loadings for the Rotated Factors

Item	Factor Loading			Communality
	1	2	3	
Slow catching on to new topics	-.90			.77
Solve math problems quickly	.78			.60
Practice math until do well	.78			.66
Have difficulties doing math	-.57			.59
Try to complete math even if takes long		.72		.50
Explore all possible solutions		.67		.45
Do not keep at it long if problem challenging		-.62		.53
Give up easily instead of persisting		-.60		.56
Prefer to figure out problems without help	.41	.59		.61
Really enjoy working math problems			-.80	.48
Smile only a little when solving math problem			.58	.37
Feel happy after solving hard problem	.49		-.54	.54
Do not get much pleasure out of math			.52	.38
Eigenvalues	3.02	2.32	1.78	
% of variance	21.55	16.62	12.74	

Note. Loadings < .40 are omitted.

The first factor, which seems to index competence, loads most strongly on the first four items, with loadings in the first column. Two of the items indexed low competence and had negative loadings. The second factor, which seemed to index motivation, was composed of the five items with loadings in column 2 of the table. "I prefer to figure out the problem without help" had its highest loading on the second factor, but had a cross-loading over .4 on the competence factor. The third factor, which seemed to index (low) pleasure from math, comprised the four items with loadings in the third column. "I feel happy after solving a hard problem" had its highest loading from the pleasure factor, but also had a strong loading from the competence factor.

Problem 5.2: Principal Components Analysis on Achievement Variables

Principal components analysis is most useful if one simply wants to reduce a relatively large number of variables into a smaller set of variables that still captures the same information.

5.2 Run a principal components analysis to see how the five "achievement" variables cluster. These variables are *grades in h.s.*, *math achievement*, *mosaic*, *visualization*, and *scholastic aptitude test – math*.

- Click on **Analyze => Data Reduction => Factor**.
- First press **Reset**.
- Next select the variables *grades in h.s.*, *math achievement*, *mosaic*, *visualization*, and *scholastic aptitude test – math* as we did in Fig. 5.1.
- In the **Descriptives** window (Fig. 5.2), check **Univariate descriptives**, **Initial solution**, **Coefficients**, **Determinant**, and **KMO and Bartlett's test of sphericity**.
- In the **Extraction** window (Fig. 5.3), use the default **Method of Principal components**, and leave the other settings as they are (**unrotated factor solution** and **Eigenvalues over 1** checked), except

Chapter 5 – Exploratory Factor Analysis and Principal Components Analysis

request a **scree plot** to see if a different number of components would be preferable to the default of eigenvalues over 1 (see Fig. 5.3).

- In the **Rotation** window (Fig. 5.4), check **Varimax**, display **Rotated solution**, and display **Loading plots**.
- Click on **Continue** and then **OK**.

We have requested a principal components analysis for the extraction and some different options for the output for contrast with the earlier one. Compare Output 5.2 to your syntax and output.

Output 5.2: Principal Components Analysis for Achievement Scores

```
FACTOR
/VARIABLES grades mathach mosaic visual satm /MISSING LISTWISE /ANALYSIS grad
es mathach mosaic visual satm
/PRINT UNIVARIATE INITIAL CORRELATION DET KMO EXTRACTION ROTATION
/PLOT EIGEN ROTATION
/CRITERIA MINEIGEN(1) ITERATE(25)
/EXTRACTION PC
/CRITERIA ITERATE(25)
/ROTATION VARIMAX
/METHOD=CORRELATION.
```

Descriptive Statistics

	Mean	Std. Deviation	Analysis N
grades in h.s.	5.68	1.570	75
math achievement test	12.5645	6.67031	75
mosaic, pattern test	27.413	9.5738	75
visualization retest	4.5467	3.01816	75
scholastic aptitude test - math	490.53	94.553	75

Correlation Matrix^a

	grades in h.s.	math achievement test	mosaic, pattern test	visualization retest	scholastic aptitude test - math
Correlation grades in h.s.	1.000	.504	-.012	.162	.371
math achievement test	.504	1.000	.213	.465	.788
mosaic, pattern test	-.012	.213	1.000	.045	.110
visualization retest	.162	.465	.045	1.000	.436
scholastic aptitude test - math	.371	.788	.110	.436	1.000

a. Determinant = .199

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.615
Bartlett's Test of Sphericity	Approx. Chi-Square	111.440
	df	10
	Sig.	.000

SPSS for Intermediate Statistics

Communalities

	Initial	Extraction
grades in h.s.	1.000	.493
math achievement test	1.000	.869
mosaic, pattern test	1.000	.949
visualization test	1.000	.330
scholastic aptitude test - math	1.000	.748

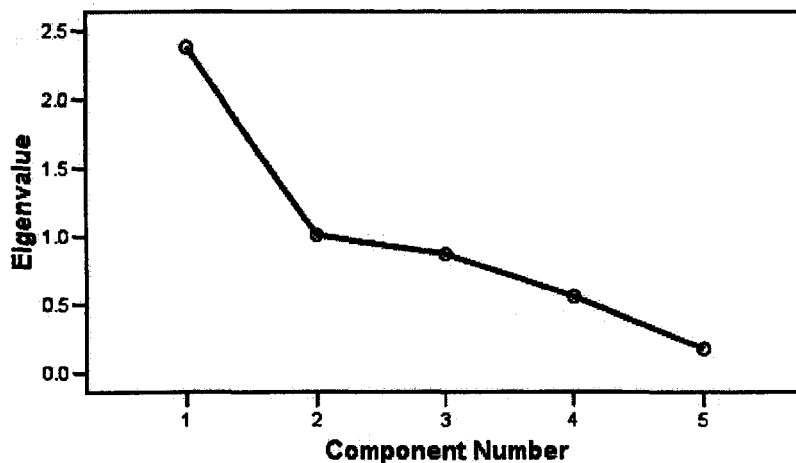
Extraction Method: Principal Component Analysis.

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	2.379	47.579	47.579	2.379	47.579	47.579	2.340	46.805	46.805
2	1.010	20.198	67.777	1.010	20.198	67.777	1.049	20.972	67.777
3	.872	17.437	85.214						
4	.560	11.197	96.411						
5	.179	3.589	100.000						

Extraction Method: Principal Component Analysis.

Scree Plot



The scree plot shows that after the first two components, increases in the eigenvalues decline, and they are less than 1.0.

Component Matrix ^a

	Component	
	1	2
grades in h.s.	.624	-.322
math achievement test	.931	4.431E-02
mosaic, pattern test	.220	.949
visualization test	.571	-5.62E-02
scholastic aptitude test - math	.865	-1.95E-02

Extraction Method: Principal Component Analysis.

a. 2 components extracted.

This unrotated matrix should not be interpreted; it provides information about how the loadings change when the solution is rotated. In addition, the first unrotated component provides the simplest summary of the variables.

Chapter 5 – Exploratory Factor Analysis and Principal Components Analysis

Rotated Component Matrix^a

	Component	
	1	2
grades in h.s.	.869	-.213
math achievement test	.911	.200
mosaic, pattern test	5.744E-02	.972
visualization test	.573	4.068E-02
scholastic aptitude test - math	.856	.126

Refer to the Interpretation of Output 5.2 and to the Component Plot.

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

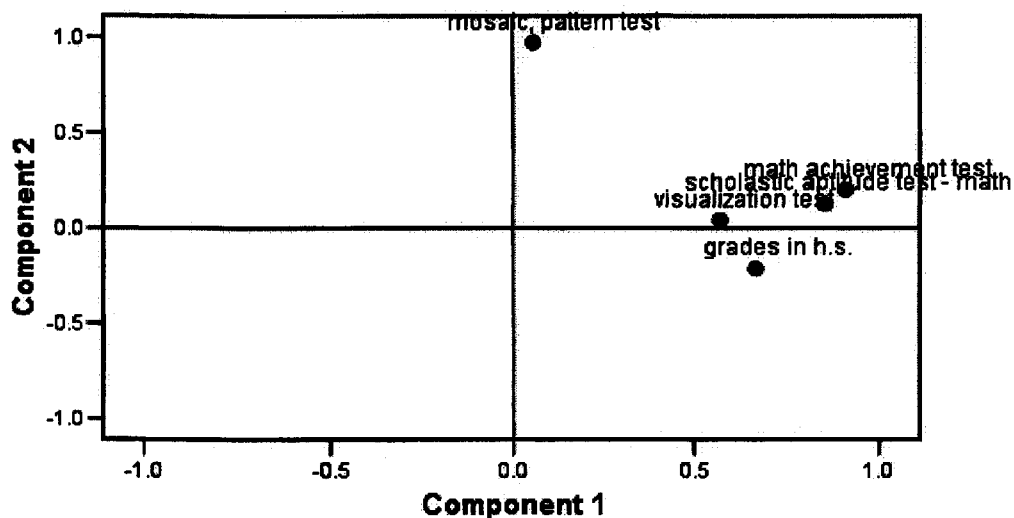
a. Rotation converged in 3 iterations.

Component Transformation Matrix

Component	1	2
1	.986	.168
2	-.168	.986

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

Component Plot in Rotated Space



Interpretation of Output 5.2

Compare Output 5.2 to your output and syntax in Output 5.1. Note that in addition to the tables in Output 5.1 you have: (a) a table of **Descriptive Statistics** for each variable, which also provides the listwise **Analysis N**, in this case 75; (b) a **Scree Plot**; (c) an unrotated **Component Matrix**, which should not be interpreted (however, if you want to compute only one variable that provides the most information about the set of variables, a linear combination of the variables with high loadings from the first unrotated component matrix would be used); (d) a **Rotated Component Matrix**, which contains the loadings for each component, in a manner comparable to those for the rotated factor matrix in 5.1; and (e) the **Component Plot** of the component loadings.

Both the **Scree Plot** and the eigenvalues support the conclusion that these five variables can be reduced to two components. However, the second component is a weak one. Note that the scree plot flattens out after the second component.

Note that the default setting we used does not sort the variables in the **Rotated Component Matrix** by magnitude of loadings and does not suppress low loadings. Thus, you have to organize the table yourself; that is, *math achievement*, *scholastic aptitude test*, *grades in h.s.*, and *visualization* in that order have high Component 1 loadings, and *mosaic* is the only variable with a high loading for Component 2.

Researchers usually give names to a component in a fashion similar to that used in EFA; however, there is no assumption that this indicates a variable that underlies the measured items. Often, a researcher will aggregate (add or average) the items that define (have high loadings for) each component, and use this composite variable in further research. Actually, the same thing is often done with EFA factor loadings; however, the implication of the latter is that this composite variable is an index of the underlying construct.

The plot of the component loadings gives one a visual representation of the loadings, plotted in space. This shows how closely related the items are to each other and to the components.

Interpretation Questions

- 5.1 Using Output 5.1: a) Are the factors in Output 5.1 close to the conceptual composites (motivation, pleasure, competence) indicated in chapter 1 ? b) How might you name the three factors in Output 5.1? c) Why did we use factor analysis, rather than principal components analysis for this exercise?
- 5.2 Using Output 5.2: a) Were any of the assumptions that were tested violated? Explain. b) Describe the main aspects of the correlation matrix, the rotated component matrix, and the plot in Output 5.2.
- 5.3 What does the plot in Output 5.2 tell us about the relation of *mosaic* to the other variables and to component 1? How does this plot relate to the rotated component matrix?

Extra SPSS Problems

- 5.1 Using the judges.sav data, do exploratory factor analysis to see if the seven variables (the judges' countries) can be grouped into two categories: former communistic block countries (Russia,

Chapter 5 – Exploratory Factor Analysis and Principal Components Analysis

China, and Romania) and non communist countries (U.S., South Korea, Italy, and France). What, if any, assumptions were violated?

- 5.2 Using the satisf.sav data, see if the six satisfaction scales can be reduced to a smaller number of variables.

CHAPTER 6

Multiple Regression

Multiple regression is one type of complex associational statistical method. Already, we have done assignments using another complex associational method, Cronbach's alpha, which, like multiple regression, is based on a correlation matrix of all the variables to be considered in a problem. In addition to multiple regression, two other complex associational analyses, **logistic regression** and **discriminant analysis** will be computed in the next chapter. Like multiple regression, logistic regression and discriminant analysis have the general purpose of predicting a dependent or criterion variable from *several* independent or predictor variables. As you can tell from examining Table 3.4 (p. 50), these three techniques for predicting one outcome measure from several independent variables vary in the level of measurement and type of independent variables and/or type of outcome variable.

For **multiple regression**, the *dependent* or outcome variable should be an interval or scale level variable, which is normally distributed in the population from which it is drawn. The *independent* variables should be mostly interval or scale level variables, but multiple regression can also have dichotomous independent variables, which are called dummy variables. **Dummy variables** are often nominal categories that have been given numerical codes, usually 1 and 0. The 0 stands for whatever the 1 is not, and is thus said to be "dumb" or silent. Thus, when we use *gender*, for instance, as a dummy variable in multiple regression, we're really coding it as 1 = female and 0 = not female (i.e., male). This gets complex when there are more than two nominal categories. In that case, we need to convert the multiple category variable to a *set* of dichotomous variables indicating presence versus absence of the categories. For example, if we were to use the ethnic group variable, we would have to code it into several dichotomous dummy variables such as Euro-American and not Euro-American, African-American and not African-American, and Latino-American and not Latino-American.

Assumptions of Multiple Linear Regression

There are many assumptions to consider but we will focus on the major ones that are easily tested with SPSS. The assumptions for multiple regression include the following: that the relationship between each of the predictor variables and the dependent variable is linear and that the error, or residual, is normally distributed and uncorrelated with the predictors. A condition that can be extremely problematic as well is **multicollinearity**, which can lead to misleading and/or inaccurate results. Multicollinearity (or collinearity) occurs when there are high intercorrelations among some set of the predictor variables. In other words, multicollinearity happens when two or more predictors contain much of the same information.

Although a correlation matrix indicating the intercorrelations among all pairs of predictors is helpful in determining whether multicollinearity is a problem, it will not always indicate that the condition exists. Multicollinearity may occur because several predictors, taken *together*, are related to some other predictors or set of predictors. For this reason, it is important to test for multicollinearity when doing multiple regression.

There are several different ways of computing multiple regression that are used under somewhat different circumstances. We will have you use several of these approaches, so that you will be able to see that the method one uses to compute multiple regression influences the information one obtains from the analysis. If the researcher has no prior ideas about which variables will create the best prediction equation and has a reasonably small set of predictors, then **simultaneous regression** is the best method

to use. It is preferable to use the **hierarchical method** when one has an idea about the *order* in which one wants to enter predictors and wants to know how prediction by certain variables *improves on* prediction by others. Hierarchical regression appropriately corrects for capitalization on chance; whereas, **stepwise**, another method available in SPSS in which variables are entered sequentially, does not. Both simultaneous regression and hierarchical regression require that you specify exactly which variables serve as predictors. Sometimes you have a relatively large set of variables that may be good predictors of the dependent variable, but you cannot enter such a large set of variables without sacrificing the power to find significant results. In such a case, stepwise regression might be used. However, as indicated earlier, stepwise regression capitalizes on chance more than many researchers find acceptable.

- Retrieve your data file: **hsbdataB.sav**

Problem 6.1: Using the Simultaneous Method to Compute Multiple Regression

To reiterate, the purpose of multiple regression is to predict an interval (or scale) dependent variable from a combination of several interval/scale, and/or dichotomous independent/predictor variables. In the following assignment, we will see if *math achievement* can be predicted better from a combination of several of our other variables, such as the *motivation scale*, *grades in high school*, and *mother's and father's education*. In Problems 6.1 and 6.3, we will run the multiple regression using alternate methods provided by SPSS. In Problem 6.1, we will assume that all seven of the predictor variables are important and that we want to see what is the highest possible multiple correlation of these variables with the dependent variable. For this purpose, we will use the method that SPSS calls **Enter** (often called **simultaneous regression**), which tells the computer to consider all the variables at the same time. In Problem 6.3, we will use the hierarchical method.

6.1. How well can you predict *math achievement* from a combination of seven variables: *motivation*, *competence*, *pleasure*, *grades in high school*, *father's education*, *mother's education*, and *gender*?

In this problem, the computer will enter/consider all the variables at the same time. Also, we will ask which of these seven predictors contribute significantly to the multiple correlation/regression.

It is a good idea to check the correlations among the predictor variables prior to running the multiple regression, to determine if the predictors are sufficiently correlated such that multicollinearity is highly likely to be a problem. This is especially important to do when one is using a relatively large set of predictors, and/or if, for empirical or conceptual reasons, one believes that some or all of the predictors might be highly correlated. If variables are highly correlated (e.g., correlated at .50 or .60 and above), then one might decide to combine (aggregate) them into a composite variable or eliminate one or more of the highly correlated variables if the variables do not make a meaningful composite variable. For this example, we will check correlations between the variables to see if there might be multicollinearity problems. We typically also would create a scatterplot matrix to check the assumption of linear relationships of each predictor with the dependent variable and a scatterplot between the predictive equation and the residual to check for the assumption that these are uncorrelated. In this problem, we will not do so because we will show you how to do these assumption checks in Problem 6.2.

- Click on **Analyze => Correlate => Bivariate**. The **Bivariate Correlations** window will appear.
- Select the variables *motivation scale*, *competence scale*, *pleasure scale*, *grades in h.s.*, *father's education*, *mother's education*, and *gender* and click them over to the **Variables** box.
- Click on **Options => Missing values => Exclude cases listwise**.

- Click on **Continue** and then click on **OK**. A correlation matrix like the one in Output 6.1a should appear.

Output 6.1a: Correlation Matrix

```

CORRELATIONS
/VARIABLES=motivation competence pleasure grades faed maed gender
/PRINT=TWOTAIL NOSIG
/MISSING=LISTWISE.
    
```

Correlations

High correlations among predictors indicate it is likely that there will be a problem with multicollinearity.

		motivation scale	competence scale	pleasure scale	grades in h.s.	father's education	mother's education	gender
motivation scale	Pearson Correlation	1	.517**	.277*	.020	.049	.115	-.178
	Sig. (2-tailed)	.	.000	.021	.872	.692	.347	.143
competence scale	Pearson Correlation	.517**	1	.413**	.216	.031	.234	-.037
	Sig. (2-tailed)	.000	.	.000	.075	.799	.053	.760
pleasure scale	Pearson Correlation	.277*	.413**	1	-.081	.020	.108	.084
	Sig. (2-tailed)	.021	.000	.	.509	.869	.378	.492
grades in h.s.	Pearson Correlation	.020	.216	-.081	1	.315**	.246*	.162
	Sig. (2-tailed)	.872	.075	.509	.	.008	.042	.182
father's education	Pearson Correlation	.049	.031	.020	.315**	1	.649**	-.266*
	Sig. (2-tailed)	.692	.799	.869	.008	.	.000	.027
mother's education	Pearson Correlation	.115	.234	.108	.246*	.649**	1	-.223
	Sig. (2-tailed)	.347	.053	.378	.042	.000	.	.065
gender	Pearson Correlation	-.178	-.037	.084	.162	-.266*	-.223	1
	Sig. (2-tailed)	.143	.760	.492	.182	.027	.065	.

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

a. Listwise N=69

The correlation matrix indicates large correlations between *motivation* and *competence* and between *mother's education* and *father's education*. To deal with this problem, we would usually aggregate or eliminate variables that are highly correlated. However, we want to show how the collinearity problems created by these highly correlated predictors affect the Tolerance values and the significance of the beta coefficients, so we will run the regression without altering the variables. To run the regression, follow the steps below:

- Click on the following: **Analyze => Regression => Linear**. The **Linear Regression** window (Fig. 6.1) should appear.
- Select *math achievement* and click it over to the **Dependent** box (dependent variable).
- Next select the variables *motivation scale*, *competence scale*, *pleasure scale*, *grades in h.s.*, *father's education*, *mother's education*, and *gender* and click them over to the **Independent(s)** box (independent variables).
- Under **Method**, be sure that **Enter** is selected.

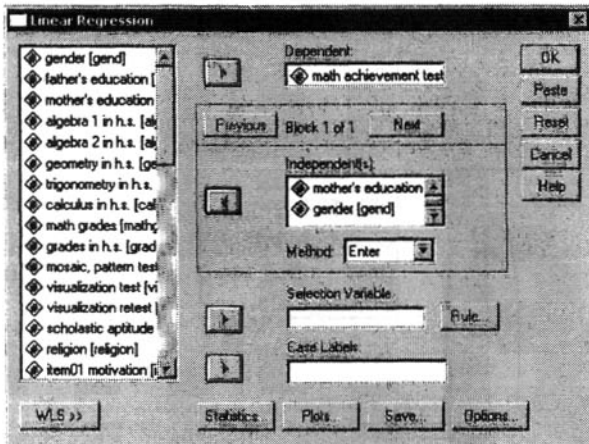


Fig. 6.1. Linear Regression.

- Click on **Statistics**, click on **Estimates** (under **Regression Coefficients**), and click on **Model fit**, **Descriptives**, and **Collinearity diagnostics**. (See Fig. 6.2.)

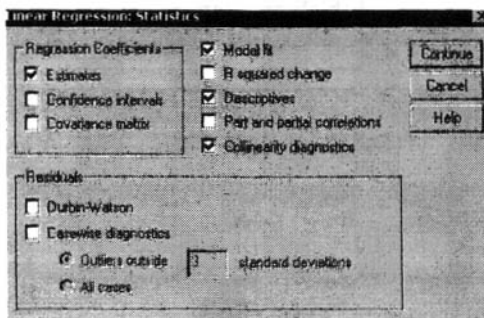


Fig. 6.2. Linear Regression: Statistics.

- Click on **Continue**.
- Click on **OK**.

Compare your output and syntax to Output 6.1b.

Output 6.1b: Multiple Linear Regression, Method = Enter

```
REGRESSION
/DESCRIPTIVES MEAN STDDEV CORR SIG N
/MISSING LISTWISE
/STATISTICS COEFF OUTS R ANOVA COLLIN TOL
/CRITERIA=PIN(.05) POUT(.10)
/NOORIGIN
/DEPENDENT mathach
/METHOD=ENTER motivation competence pleasure grades faed maed gender.
```


Regression

Descriptive Statistics

	Mean	Std. Deviation	N
math achievement test	12.7536	6.66293	69
Motivation scale	2.8913	.62676	69
Competence scale	3.3188	.62262	69
pleasure scale	3.1667	.66789	69
grades in h.s.	5.71	1.573	69
father's education	4.65	2.764	69
mother's education	4.07	2.185	69
gender	.54	.502	69

N is 69 because 6 participants have some missing data.

Correlations with *math achievement*.

This is a repeat of the correlation matrix we did earlier, indicating high correlations among predictors.

		math achievement test	Motivation scale	Competence scale	pleasure scale	grades in h.s.	father's education	mother's education	gender
Pearson Correlation	math achievement test	1.000	.256	.260	.087	.470	.416	.387	-.272
	Motivation scale	.256	1.000	.517	.274	.020	.049	.115	-.178
	Competence scale	.260	.517	1.000	.438	.216	.031	.234	-.037
	pleasure scale	.087	.274	.438	1.000	-.111	-.008	.085	.037
	grades in h.s.	.470	.020	.216	-.111	1.000	.315	.246	.162
	father's education	.416	.049	.031	-.008	.315	1.000	.649	-.266
	mother's education	.387	.115	.234	.085	.246	.649	1.000	-.223
	gender	-.272	-.178	-.037	.037	.162	-.266	-.223	1.000
Sig. (1-tailed)	math achievement test		.017	.015	.239	.000	.000	.001	.012
	Motivation scale	.017		.000	.011	.436	.346	.173	.072
	Competence scale	.015	.000		.000	.037	.400	.026	.380
	pleasure scale	.239	.011	.000		.182	.474	.244	.383
	grades in h.s.	.000	.436	.037	.182		.004	.021	.091
	father's education	.000	.346	.400	.474	.004		.000	.014
	mother's education	.001	.173	.026	.244	.021	.000		.032
	gender	.012	.072	.380	.383	.091	.014	.032	
N	math achievement test	69	69	69	69	69	69	69	69
	Motivation scale	69	69	69	69	69	69	69	69
	Competence scale	69	69	69	69	69	69	69	69
	pleasure scale	69	69	69	69	69	69	69	69
	grades in h.s.	69	69	69	69	69	69	69	69
	father's education	69	69	69	69	69	69	69	69
	mother's education	69	69	69	69	69	69	69	69
	gender	69	69	69	69	69	69	69	69

Variables Entered/Removed^a

Model	Variables Entered	Variables Removed	Method
1	gender, pleasure scale, grades in h.s., Motivation scale, mother's education, Competence scale, father's education ^a		Enter

Significance level of correlations with *math achievement*.

a. All requested variables entered.

b. Dependent Variable: math achievement test

Chapter 6 - Multiple Regression

Multiple correlation coefficient.

Model Summary^a

Indicates that 36% of the variance can be predicted from the independent variables.

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.654 ^a	.427	.362	5.32327

a. Predictors: (Constant), gender, pleasure scale, grades in h.s., Motivation scale, mother's education, Competence scale, father's education

b. Dependent Variable: math achievement test

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1290.267	7	184.324	6.505	.000 ^a
	Residual	1728.571	61	28.337		
	Total	3018.838	68			

a. Predictors: (Constant), gender, pleasure scale, grades in h.s., Motivation scale, mother's education, Competence scale, father's education

b. Dependent Variable: math achievement test

Indicates that the combination of these variables significantly ($p < .001$) predicts the dependent variable.

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-6.912	4.749		-1.455	.151		
	Motivation scale	1.639	1.233	.154	1.330	.188	.698	1.432
	Competence scale	1.424E-02	1.412	.001	.010	.992	.539	1.854
	pleasure scale	.953	1.119	.096	.852	.398	.746	1.340
	grades in h.s.	1.921	.480	.453	4.001	.000	.731	1.368
	father's education	.303	.331	.126	.915	.364	.497	2.013
	mother's education	.333	.406	.109	.920	.415	.529	1.892
	gender	-3.497	1.424	-.264	-2.455	.017	.814	1.228

a. Dependent Variable: math achievement test

Only *grades* and *gender* are significantly contributing to the equation. However, all of the variables need to be included to obtain this result, since the overall F value was computed with all the variables in the equation.

Tolerance and VIF give the same information. (Tolerance = $1/VIF$) They tell us if there is multicollinearity. If the Tolerance value is low ($< 1-R^2$), then there is probably a problem with multicollinearity. In this case, since adjusted R^2 is .36, and $1-R^2$ is about .64, then tolerances are low for *competence*, *mother's* and *father's*

This tells you how much each variable is contributing to any collinearity in the model.

Collinearity Diagnostics^a

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions							
				(Constant)	Motivation scale	Competence scale	pleasure scale	grades in h.s.	father's education	mother's education	gender
1	1	7.035	1.000	.00	.00	.00	.00	.00	.00	.00	.00
	2	.550	3.577	.00	.00	.00	.00	.00	.04	.02	.49
	3	.215	5.722	.00	.02	.01	.01	.00	.18	.09	.32
	4	8.635E-02	9.026	.00	.00	.00	.00	.06	.45	.78	.01
	5	5.650E-02	11.159	.00	.01	.00	.10	.80	.23	.04	.08
	6	2.911E-02	15.545	.01	.59	.00	.43	.05	.01	.00	.08
	7	1.528E-02	21.456	.70	.00	.46	.02	.02	.05	.04	.01
	8	1.290E-02	23.350	.29	.38	.53	.44	.28	.03	.02	.00

a. Dependent Variable: math achievement test

Interpretation of Output 6.1

First, the output provides the usual descriptive statistics for all eight variables. Note that the N is 69 because 6 participants are missing a score on one or more variables. Multiple regression uses only the participants who have complete data for all the variables. The next table is a correlation matrix similar to the one in Output 6.1a. Note that the first column shows the correlations of the other variables with *math achievement* and that *motivation*, *competence*, *grades in high school*, *father's education*, *mother's education*, and *gender* are all significantly correlated with *math achievement*. As we observed before, several of the predictor/ independent variables are highly correlated with each other; that is, *competence* and *motivation* (.517) and *mother's education* and *father's education* (.649).

The **Model Summary** table shows that the multiple correlation coefficient (R), using all the predictors simultaneously, is .65 ($R^2 = .43$) and the **adjusted R^2** is .36, meaning that 36% of the variance in *math achievement* can be predicted from *gender*, *competence*, etc. combined. Note that the adjusted R^2 is lower than the unadjusted R^2 . This is, in part, related to the number of variables in the equation. The adjustment is also affected by the magnitude of the effect and the sample size. As you will see from the coefficients table, only *father's education* and *gender* are significant, but the other five variables will always add a little to the prediction of *math achievement*. Because so many independent variables were used, a reduction in the number of variables might help us find an equation that explains more of the variance in the dependent variable. It is helpful to use the concept of parsimony with multiple regression, and use the smallest number of predictors needed.

The ANOVA table shows that $F = 6.51$ and is significant. This indicates that the combination of the predictors significantly predict *math achievement*.

One of the most important tables is the **Coefficients** table. It indicates the **standardized beta coefficients**, which are interpreted similarly to correlation coefficients or factor weights (see chapter 4). The t value and the **Sig** opposite each independent variable indicates whether that variable is significantly contributing to the equation for predicting *math achievement* from the whole set of predictors. Thus, *h.s. grades* and *gender*, in this example, are the only variables that are significantly adding anything to the prediction when the other five variables are already considered. It is important to note that all the variables are being considered together when these values are computed. Therefore, if you delete one of the predictors that is not significant, it can affect the levels of significance for other predictors.

However, as the **Tolerances** in the **Coefficients** table suggest, and as we will see in Problem 6.2, these results are somewhat misleading. Although the two parent education measures were significantly correlated with *math achievement*, they did not contribute to the multiple regression predicting *math achievement*. What has happened here is that these two measures were also highly correlated with each other, and multiple regression eliminates all overlap between predictors. Thus, neither *father's education* nor *mother's education* had much to contribute when the other was also used as a predictor. Note that tolerance for each of these variables is $< .64$ ($1 - .36$), indicating that too much multicollinearity (overlap between predictors) exists. The same is true for *competence*, once *motivation* is entered. One way to handle multicollinearity is to combine variables that are highly related if that makes conceptual sense. For example, you could make a new variable called *parents' education*, as we will for Problem 6.2.

Problem 6.2: Simultaneous Regression Correcting Multicollinearity

In Problem 6.2, we will use the combined/average of the two variables, *mother's education* and *father's education*, and then recompute the multiple regression, after omitting *competence* and *pleasure*.

We combined *father's education* and *mother's education* because it makes conceptual sense and because these two variables are quite highly related ($r = .65$). We know that entering them as two separate variables created problems with multicollinearity because tolerance levels were low for these two variables, and, despite the fact that both variables were significantly and substantially correlated with *math achievement*, neither contributed significantly to predicting *math achievement* when taken together.

When it does not make sense to combine the highly correlated variables, one can eliminate one or more of them. Because the conceptual distinction between *motivation*, *competence*, and *pleasure* was important for us, and because *motivation* was more important to us than *competence* or *pleasure*, we decided to delete the latter two scales from the analysis. We wanted to see if *motivation* would contribute to the prediction of *math achievement* if its contribution was not canceled out by *competence* and/or *pleasure*. *Motivation* and *competence* are so highly correlated that they create problems with multicollinearity. We eliminate *pleasure* as well, even though its tolerance is acceptable, because it is virtually uncorrelated with *math achievement*, the dependent variable, and yet it is correlated with *motivation* and *competence*. Thus, it is unlikely to contribute meaningfully to the prediction of *math achievement*, and its inclusion would only serve to reduce power and potentially reduce the predictive power of *motivation*. It would be particularly important to eliminate a variable such as *pleasure* if it were strongly correlated with another predictor, as this can lead to particularly misleading results.

6.2. Rerun Problem 6.1 using the *parents' education* variable (*parEduc*) instead of *faed* and *maed* and omitting the *competence* and *pleasure* scales.

First, we created a matrix scatterplot (as in chapter 2) to see if the variables are related to each other in a linear fashion. You can use the syntax in Output 6.2 or use the **Analyze => Scatter** windows as shown below.

- Click on **Graphs => Scatter...**
- Select **Matrix** and click on **Define**.
- Move *math achievement*, *motivation*, *grades*, *parent's education*, and *gender* into the **Matrix Variables:** box.
- Click on **Options**. Check to be sure that **Exclude cases listwise** is selected.
- Click on **Continue** and then **OK**.

Then, run the regression, using the following steps:

- Click on the following: **Analyze => Regression => Linear**. The **Linear Regression** window (Fig. 6.1) should appear. This window may still have the variables moved over to the **Dependent** and **Independent(s)** boxes. If so, click on **Reset**.
- Move *math achievement* into the **Dependent** box.
- Next select the variables *motivation*, *grades in h.s.*, *parent's education*, and *gender* and move them into the **Independent(s)** box (independent variables).
- Under **Method**, be sure that **Enter** is selected.
- Click on **Statistics**, click on **Estimates** (under **Regression Coefficients**), and click on **Model fit**, **Descriptives**, and **Collinearity diagnostics** (See Fig. 6.2.).
- Click on **Continue**.
- Click on **OK**.

Then, we added a plot to the multiple regression to see the relationship of the predictors and the residual. To make this plot follow these steps:

- Click on **Plots...** (in Fig. 6.1 to get Fig. 6.3.)

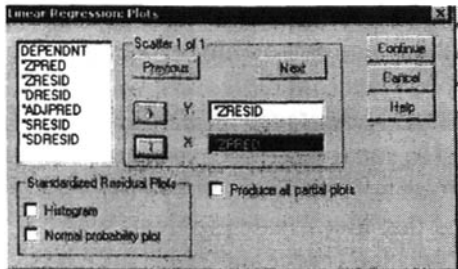


Fig. 6.3. Linear Regression: Plots.

- Move **ZRESID** to the **Y:** box.
- Move **ZPRED** to the **X:** box. This enables us to check the assumption that the predictors and residual are uncorrelated.
- Click on **Continue**.
- Click on **OK**.

Refer to Output 6.2 for comparison.

Output 6.2: Multiple Linear Regression with Parent's Education, Method = Enter

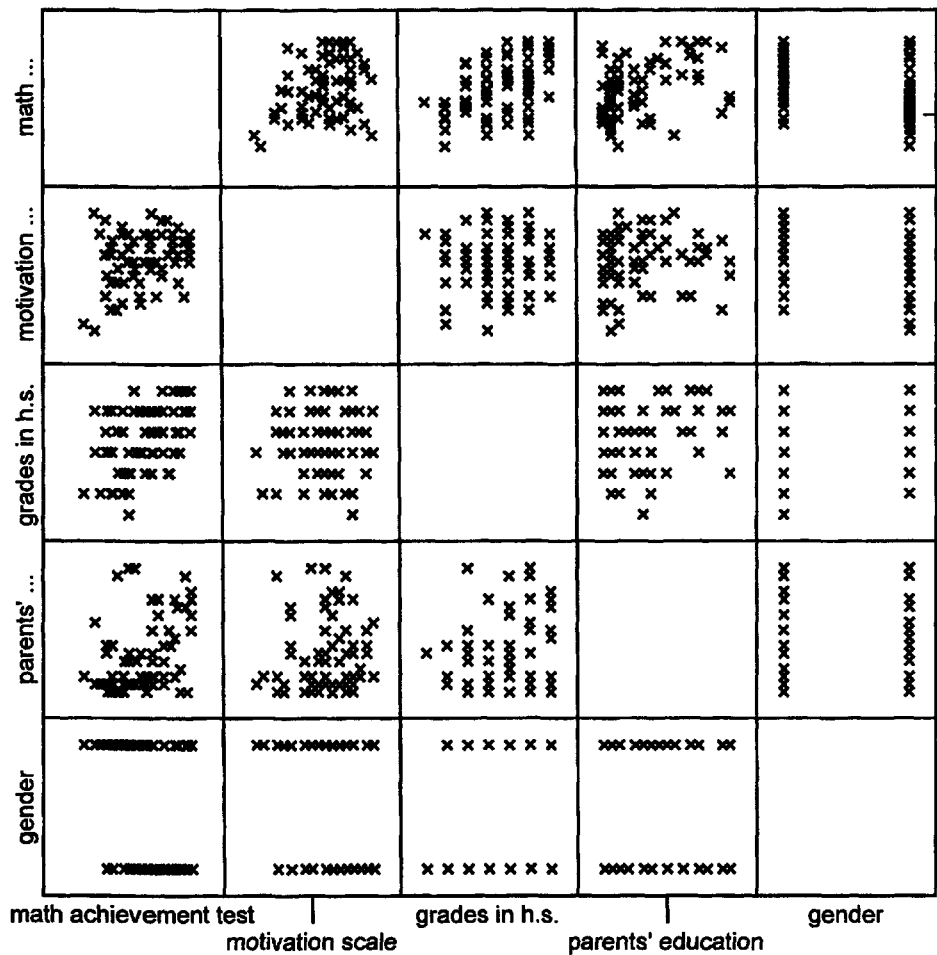
```
GRAPH
  /SCATTERPLOT(MATRIX)=mathach motivation grades parEduc gender
  /MISSING=LISTWISE .
```

```
REGRESSION
  /DESCRIPTIVES MEAN STDDEV CORR SIG N
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA COLLIN TOL
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT mathach
  /METHOD=ENTER motivation grades parEduc gender
  /SCATTERPLOT=(*ZRESID , *ZPRED) .
```

Graph

The top row shows four scatterplots (relationships) of the dependent variables with each of the predictors. To meet the assumption of linearity a straight line, as opposed to a curved line, should fit the points relatively well.

Dichotomous variables have two columns (or rows) of data points. If the data points bunch up near the top of the left column and the bottom of the right, the correlation will be negative (and vice versa). Linearity would be violated if the data points bunch at the center of one column and at the ends of the other column.



Regression

Descriptive Statistics

	Mean	Std. Deviation	N
math achievement test	12.6028	6.75676	73
motivation scale	2.8744	.63815	73
grades in h.s.	5.68	1.589	73
parents' education	4.3836	2.30266	73
gender	.55	.501	73

Note that $N = 73$, indicating that eliminating competence and pleasure reduced the amount of missing data.

Note that all the predictors are significantly related to *math achievement*.

Correlations

None of the relationships among predictors is greater than .25.

		math achievement test	motivation scale	grades in h.s.	parents' education	gender
Pearson Correlation	math achievement test	1.000	.316	.504	.394	-.303
	motivation scale	.316	1.000	.084	.090	-.209
	grades in h.s.	.504	.084	1.000	.250	.115
	parents' education	.394	.090	.250	1.000	-.227
	gender	-.303	-.209	.115	-.227	1.000
Sig. (1-tailed)	math achievement test	.	.003	.000	.000	.005
	motivation scale	.003	.	.241	.225	.038
	grades in h.s.	.000	.241	.	.016	.166
	parents' education	.000	.225	.016	.	.027
	gender	.005	.038	.166	.027	.
N	math achievement test	73	73	73	73	73
	motivation scale	73	73	73	73	73
	grades in h.s.	73	73	73	73	73
	parents' education	73	73	73	73	73
	gender	73	73	73	73	73

Variables Entered/Removed^b

Model	Variables Entered	Variables Removed	Method
1	gender, grades in h.s., motivation scale, parents' education ^a		Enter

This indicates we used simultaneous regression in this problem.

a All requested variables entered.

b Dependent Variable: math achievement test

Chapter 6 - Multiple Regression

Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.678 ^a	.459	.427	5.11249

a. Predictors: (Constant), gender, grades in h.s., motivation scale, parent's education

b. Dependent Variable: math achievement test

The Adjusted R Square indicates that we have a fairly good model, explaining about 43% of the variance in *math achievement*.

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1509.723	4	377.431	14.440	.000 ^a
	Residual	1777.353	68	26.138		
	Total	3287.076	72			

a. Predictors: (Constant), gender, grades in h.s., motivation scale, parent's education

b. Dependent Variable: math achievement test

Our model significantly predicts *math achievement*.

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-5.444	3.605		-1.510	.136		
	motivation scale	2.148	.972	.203	2.211	.030	.944	1.059
	grades in h.s.	1.991	.400	.468	4.972	.000	.897	1.115
	parent's education	.580	.280	.198	2.070	.042	.871	1.148
	gender	-3.631	1.284	-.269	-2.828	.006	.877	1.141

a. Dependent Variable: math achievement test

Here are the values to check for multicollinearity. Note that all tolerances are well over .57 ($1-R^2$).

Collinearity Diagnostics^a

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions				
				(Constant)	motivation scale	grades in h.s.	parent's education	gender
1	1	4.337	1.000	.00	.00	.00	.01	.01
	2	.457	3.082	.00	.00	.00	.07	.68
	3	.135	5.665	.02	.07	.02	.85	.17
	4	.052	9.120	.01	.20	.87	.06	.06
	5	.019	15.251	.97	.73	.11	.01	.08

a. Dependent Variable: math achievement test

SPSS for Intermediate Statistics

Casewise Diagnostics^a

Case Number	Std. Residual	math achievement test
63	-3.174	1.00

a. Dependent Variable: math achievement test

Residuals Statistics^a

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	1.5029	22.2180	12.6028	4.57912	73
Residual	-16.2254	10.3169	.0000	4.96845	73
Std. Predicted Value	-2.424	2.100	.000	1.000	73
Std. Residual	-3.174	2.018	.000	.972	73

a. Dependent Variable: math achievement test

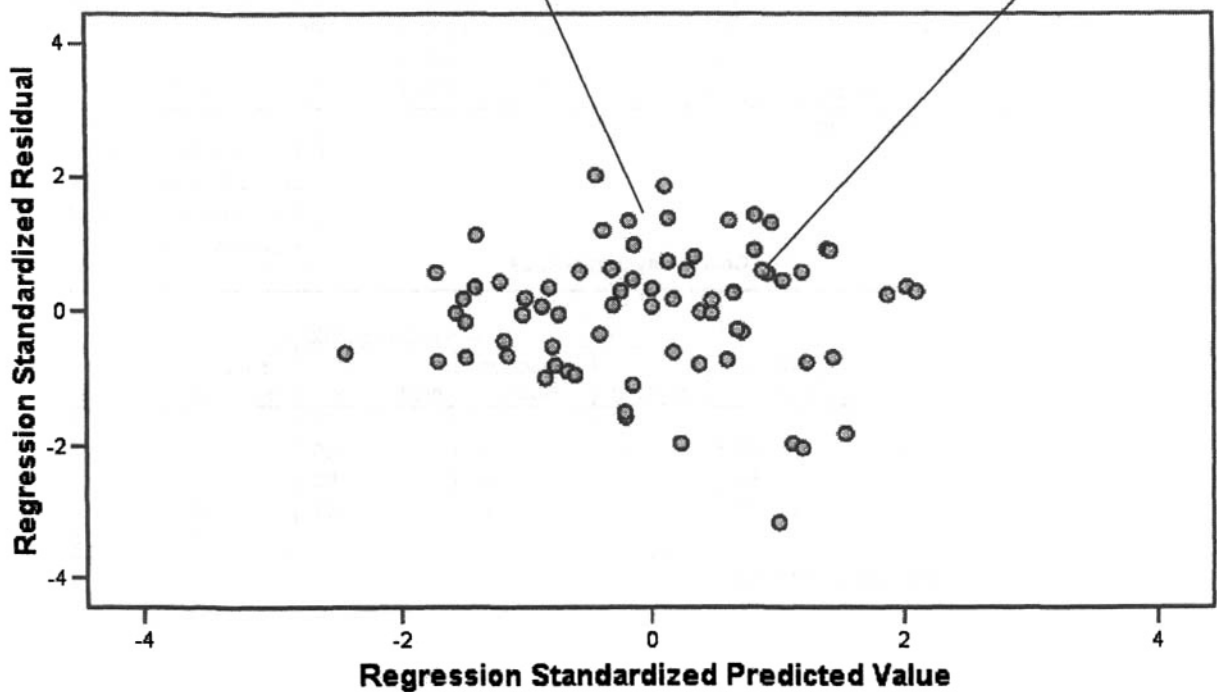
Charts

If the dots created a pattern, this would indicate the residuals are not normally distributed, the residual is correlated with the independent variables, and/or the variances of the residuals are not constant.

Scatterplot

Because the dots are scattered, it indicates the data meet the assumptions of the errors being normally distributed and the variances of the residuals being constant.

Dependent Variable: math achievement test



Interpretation of Output 6.2

This output begins with the **scatterplot matrix**, which shows that the independent variables are generally linearly related to the dependent variable of *math achievement*, meeting this assumption. One should check the matrix scatterplots to see if there are curvilinear relationships between any of the variables; in this example, there are none. If the variables had not met this assumption, we could have transformed them, aggregated some, and/or eliminated some independent variables. (See chapter 2 for how to do transformations.) There are also only low to moderate relationships among the predictor variables in the **Correlations** table. This is good. The other assumptions are checked in the **residual scatterplot** at the end of the output, which indicates that the errors are normally distributed, the variances of the residuals are constant, and the residual is relatively uncorrelated with the linear combination of predictors.

The next important part of the output to check is the **Tolerance** and **VIF** values in the **Coefficients** table for the existence of multicollinearity. In this example, we do not need to worry about multicollinearity because the Tolerance values are close to 1.

The **Model Summary** table gives the **R** (.68) and **Adjusted R square** (.43). Thus, this model is predicting 43% of the variance in *math achievement*. If we want to compare this model to that in Problem 6.1, we use the Adjusted *R* square to see which model is explaining more of the variance in the dependent variable. Interestingly, this model is predicting more of the variance in *math achievement* than the model in Problem 6.1, despite using fewer predictors.

As can be seen from the **ANOVA** table, the model of *motivation, grades in h.s., parents' education, and gender* significantly predicts *math achievement*, $F(4, 68) = 14.44, p < .001$.

We can see from the **Coefficients** table that now all of the predictors are significantly contributing to the equation (see the **Sig.** column).

How to Write About Output 6.2

Results

Multiple regression was conducted to determine the best linear combination of gender, grades in h.s., parents' education, and motivation for predicting math achievement test scores. The means, standard deviations, and intercorrelations can be found in Table 6.1. This combination of variables significantly predicted math achievement, $F(4,68) = 14.44, p < .001$, with all four variables significantly contributing to the prediction. The beta weights, presented in Table 6.2, suggest that good grades in high school contribute most to predicting math achievement, and that being male, having high math motivation and having parents who are more highly educated also contribute to this prediction. The adjusted *R* squared value was .43. This indicates that 43% of the variance in math achievement was explained by the model. According to Cohen (1988), this is a large effect.

Table 6.1

Means, Standard Deviations, and Intercorrelations for Math Achievement and Predictor Variables (N=73)

Variable	<i>M</i>	<i>SD</i>	1	2	3	4
Math Achievement	12.60	6.76	.32**	.50**	.39**	-.30**
Predictor variable						
1. Motivation scale	2.87	.64	-	.08	.09	-.21*
2. Grades in h.s.	5.68	1.59		-	.25*	.12
3. Parent's education	4.38	2.30			-	-.23*
4. Gender	.55	.50				-

* $p < .05$; ** $p < .01$.

Table 6.2

Simultaneous Multiple Regression Analysis Summary for Motivation, Grades in High School, Parent's Education, and Gender Predicting Math Achievement (N = 73)

Variable	<i>B</i>	<i>SEB</i>	β
Motivation scale	2.15	.97	.20*
Grades in h.s.	1.99	.40	.47**
Parent's education	.58	.28	.20*
Gender	-3.63	1.28	-.27**
Constant	-5.44	3.61	

Note. $R^2 = .46$; $F(4,68) = 14.44$, $p < .001$

* $p < .05$; ** $p < .01$.

Problem 6.3: Hierarchical Multiple Linear Regression

In Problem 6.3, we will use the **hierarchical** approach, which enters variables in a series of blocks or groups, enabling the researcher to see if each new group of variables adds anything to the prediction produced by the previous blocks of variables. This approach is an appropriate method to use when the researcher has a priori ideas about how the predictors go together to predict the dependent variable. In our example, we will enter *gender* first and then see if any of the other variables make an additional contribution. This method is intended to control for or eliminate the effects of *gender* on the prediction.

6.3. If we control for *gender* differences in *math achievement*, do any of the other variables significantly add anything to the prediction over and above what *gender* contributes?

Chapter 6 - Multiple Regression

We will include all of the variables from the previous problem; however, this time we will enter the variables in two separate blocks to see how *motivation*, *grades in high school*, and *parents' education* improve on prediction from *gender* alone.

- Click on the following: **Analyze => Regression => Linear.**
- Click on **Reset.**
- Select *math achievement* and click it over to the **Dependent** box (dependent variable).
- Next, select *gender* and move it over to the **Independent(s)** box (independent variables).
- Select **Enter** as your **Method.** (See Fig. 6.4.)

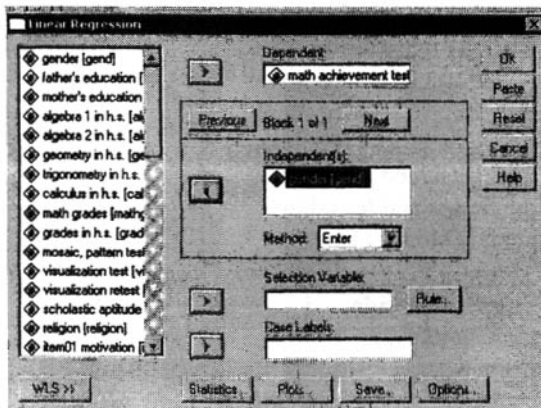


Fig. 6.4. Linear regression.

- Click on **Next** beside **Block 1 of 1.** You will notice it changes to **Block 2 of 2.**
- Then move *motivation scale*, *grades in h.s.*, and *parent's education* to the **Independent(s)** box (independent variables).
- Under **Method**, select **Enter.** The window should look like Fig. 6.5.

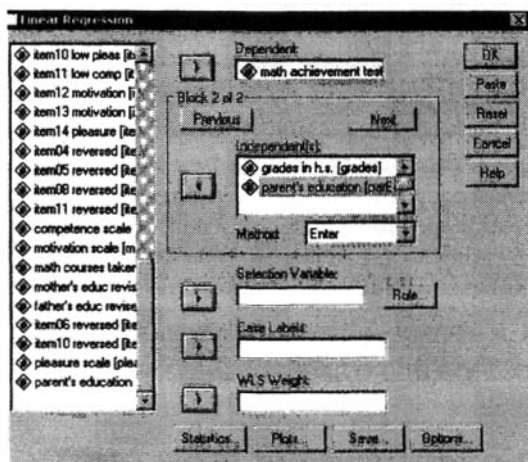


Fig. 6.5. Hierarchical Regression.

- Click on **Statistics**, click on **Estimates** (under **Regression Coefficients**), and click on **Model fit** and **R squared change.** (See Fig. 6.2.)
- Click on **Continue.**
- Click on **OK.**

Compare your output and syntax to Output 6.3.

Output 6.3: Hierarchical Multiple Linear Regression

```
REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA CHANGE
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT mathach
  /METHOD=ENTER gender /METHOD=ENTER motivation grades parEduc .
```

Regression

Variables Entered/Removed^b

Model	Variables Entered	Variables Removed	Method
1	gender ^a		Enter
2	grades in h.s., motivation scale, parents' education ^a		Enter

a. All requested variables entered.

b. Dependent Variable: math achievement test

In the first column of this table there are two models (1 and 2). This indicates that first we tested a model with *gender* as a predictor, then we added the other predictors and tested that model (Model 2).

Footnotes provide you with relevant information.

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	.303(a)	.092	.079	6.48514	.092	7.158	1	71	.009
2	.678(b)	.459	.427	5.11249	.368	15.415	3	68	.000

a Predictors: (Constant), gender

b Predictors: (Constant), gender, grades in h.s., motivation scale, parents' education

The Model Summary output shows there were two models run: Model 1 (in the first row) and Model 2 (in the second row). It also shows that the addition of *grades*, *motivation*, and *parents' education* significantly improved on the prediction by *gender* alone, explaining almost 37% additional variance.

ANOVA(c)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	301.026	1	301.026	7.158	.009(a)
	Residual	2986.050	71	42.057		
	Total	3287.076	72			
2	Regression	1509.723	4	377.431	14.440	.000(b)
	Residual	1777.353	68	26.138		
	Total	3287.076	72			

a Predictors: (Constant), gender

b Predictors: (Constant), gender, grades in h.s., motivation scale, parents' education

c Dependent Variable: math achievement test

Chapter 6 - Multiple Regression

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	14.838	1.129		13.144	.000
	gender	-4.080	1.525	-.303	-2.675	.009
2	(Constant)	-5.444	3.605		-1.510	.136
	gender	-3.631	1.284	-.269	-2.828	.006
	motivation scale	2.148	.972	.203	2.211	.030
	grades in h.s.	1.991	.400	.468	4.972	.000
	parents' education	.580	.280	.198	2.070	.042

a. Dependent Variable: math achievement test

Excluded Variables(b)

Model		Beta In	T	Sig.	Partial Correlation	Collinearity Statistics
						Tolerance
1	motivation scale	.264(a)	2.358	.021	.271	.956
	grades in h.s.	.546(a)	5.784	.000	.569	.987
	Parents' education	.343(a)	3.132	.003	.351	.949

a Predictors in the Model: (Constant), gender

b Dependent Variable: math achievement test

Interpretation of Output 6.3

We did not need to recheck the assumptions for this problem, because we checked them in Problem 6.2 with the same variables.

The **Descriptives** and **Correlations** tables would have been the same as those in Problem 6.2 if we had checked the Descriptive box in the Statistics window.

The other tables in this output are somewhat different than the previous two outputs. This difference is because we entered the variables in two steps. Therefore, this output has two models listed, Model 1 and Model 2. The information in Model 1 is for *gender* predicting *math achievement*. The information in Model 2 is *gender* plus *motivation*, *grades in h.s.*, and *parents' education* predicting *math achievement*.

We can see from the **ANOVA** table that when *gender* is entered by itself, it is a significant predictor of *math achievement*, $F(1,71) = 7.16$, $p = .009$; however, the model with the addition of the other predictor variables is a better model for predicting *math achievement* $F(4,68) = 14.44$, $p < .001$. That Model 2 is better than Model 1 can also be seen in the **Model Summary** table by the increase in the adjusted R^2 value from $R^2 = .08$ to an $R^2 = .43$, $F(3, 83) = 15.42$, $p < .001$.

Note also that results for the final model, with all of the variables entered, is identical to the model in Problem 6.2. No matter how they are entered, the same regression coefficients are selected to produce the best model with the same set of predictors and the same dependent variable.

Interpretation Questions

- 6.1 In Output 6.1: a) What information suggests that we might have a problem of collinearity? b) How does multicollinearity affect results? c) What is the adjusted R^2 and what does it mean?
- 6.2 Using Output 6.2: a) How does combining (aggregating) *mother's education* and *father's education* and eliminating *competence* and *pleasure scales* change the results from those in Output 6.1? b) Why did we aggregate *mother's education* and *father's education*? c) Why did we eliminate the *competence* and *pleasure scales*?
- 6.3 In Output 6.3: a) Compare the adjusted R^2 for Model 1 and Model 2. What does this tell you? b) Why would one enter *gender* first?

Extra SPSS Problems

- 6.1 Open the file, 'World95'. You will use this data set for all extra problems for Chapter 6. Using People living in cities (*urban*), People who read (*literacy*), Infant mortality (*babymort*), Gross domestic product (*gdp_cap*), Aids cases (*aids*), and Daily calorie intake (*calories*) as independent variables and Average male life expectancy (*lifeexpm*) as the dependent variable, conduct a linear regression.
- Is the regression significant?
 - What is the Adjusted R Square?
 - Is multicollinearity an issue for this analysis, according to the criteria we have been using? Are tolerances high? What does this imply?
 - Which variables are weighted most highly in the regression equation? Which betas are significant?
- 6.2 Rerun the analysis from problem 1, but this time omit *babymort* as a predictor.
- Is the regression significant?
 - What is the Adjusted R Square?
 - Is multicollinearity an issue for this analysis, according to the criteria we have been using? Are tolerances high? How have tolerances and multicollinearity been affected by leaving out *babymort*?
 - Which variables are weighted most highly in the regression equation? Which betas are significant?
- 6.3 Run a hierarchical regression, using Average female life expectancy (*lifeexpf*) as the dependent variable, entering Average male life expectancy (*lifeexpm*) as the first block, and then the same predictors from problem 2 for the second block. Be sure to check R square change under statistics.
- Is the regression significant?
 - What is the Adjusted R Square?
 - Is multicollinearity an issue for this analysis, according to the criteria we have been using? Are tolerances high? How have tolerances and multicollinearity been affected by including *lifeexpm* as a predictor?
 - Which variables are weighted most highly in the regression equation? Which betas are significant? Do any variables predict *lifeexpf* after *lifeexpm* is taken into account? Which variables might you want to eliminate from the equation? Why?

CHAPTER 7

Logistic Regression and Discriminant Analysis

Logistic regression and discriminant analysis, like multiple regression, are useful when you want to predict an outcome or dependent variable from a set of predictor variables. They are similar to a linear regression in many ways. However, logistic regression and discriminant analysis are more appropriate when the dependent variable is categorical. Logistic regression is useful because it does not rely on some of the assumptions on which multiple regression and discriminant analysis are based. As with other forms of regression, multicollinearity (high correlations among the predictors) can lead to problems for both logistic and discriminant analysis.

Logistic regression is helpful when you want to predict a categorical variable from a set of predictor variables. **Binary logistic regression** is similar to linear regression except that it is used when the dependent variable is dichotomous. **Multinomial logistic regression** is used when the dependent/outcome variable has more than two categories, but that is complex and less common, so we will not discuss it here. Logistic regression also is useful when some or all of the independent variables are dichotomous; others can be continuous.

Discriminant analysis, on the other hand, is useful when you have several continuous independent variables and, as in logistic regression, an outcome or dependent variable that is categorical. The dependent variable can have more than two categories. If so, then more than one discriminant function will be generated (number of functions = number of levels of the dependent variable minus 1). For the sake of simplicity, we will limit our discussion to the case of a dichotomous dependent variable here. Discriminant analysis is useful when you want to build a predictive model of group membership based on several observed characteristics of each participant. SPSS can create a linear combination of the predictor variables that provides the best discrimination between the groups.

In Problems 7.1 and 7.2, we will use logistic regression to predict a dichotomous outcome (whether or not students will take *algebra 2*) from continuous and dichotomous predictors. In Problem 7.1, we will enter all four predictors simultaneously into the equation. This is similar to the first multiple regression problem in chapter 6. In Problem 7.2, we use a hierarchical approach, in which blocks of variables are entered sequentially. In Problem 7.3, we will use discriminant analysis to do the same problem that we did with logistic regression in Problem 7.1 in order to compare the two techniques.

- Get your **hsbdataB** data file.

Problem 7.1: Logistic Regression

Assumptions of logistic regression

There are fewer assumptions for logistic regression than for multiple regression and discriminant analysis, which is one reason this technique has become popular, especially in health related fields. Binary logistic regression assumes that the dependent or outcome variable is dichotomous and, like most other statistics, that the outcomes are independent and mutually exclusive; that is, a single case can only be represented once and must be in one group or the other. Finally, logistic regression requires large samples to be accurate: Some say there should be a minimum of 20 cases per predictor, with a minimum of 60 total cases. These requirements need to be satisfied prior to doing statistical analysis with SPSS. As

with multiple regression, multicollinearity is a potential source of confusing or misleading results and needs to be assessed.

7.1. Is there a combination of *gender*, *parents' education*, *mosaic*, and *visualization test* that predicts whether students will take *algebra 2*?

Let's try a logistic regression to predict a dichotomous (two category) dependent variable when the independent variables (called covariates in SPSS) are either dichotomous or normal/scale.

First, we should check for multicollinearity. Because tolerance and VIF scores are not available through the logistic regression command, one way to compute these values is through the linear regression command, using *algebra 2* as the dependent variable and *gender*, *mosaic*, *visualization test*, and *parents' education* as independent variables. If you did this, you would find that the tolerance statistics are all above .87 so there is little multicollinearity. (See chapter 6 for how to do this.)

Use these commands to compute the logistic regression:

- **Analyze => Regression => Binary Logistic** to get to Fig. 7.1.
- Move *algebra 2 in h.s.* into the **Dependent** variable box.
- Move *gender*, *mosaic*, *visualization test*, and *parents' education* into the **Covariates** box.
- Make sure **Enter** is the selected **Method**. (This enters all the variables in the covariates box into the logistic regression equation simultaneously.)
- Click on **Options** to produce Fig. 7.2.

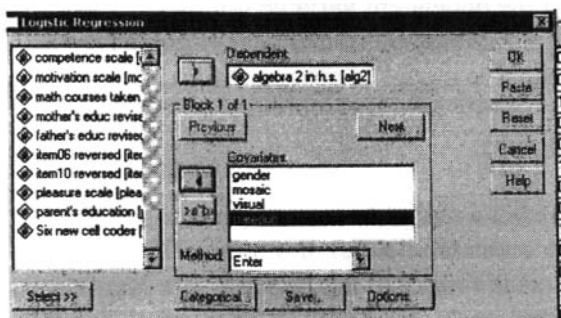


Fig. 7.1. Logistic regression.

- Check **CI for exp(B)**, and be sure **95** is in the box (which will provide confidence intervals for the odds ratio of each predictor's contribution to the equation).
- Click on **Continue**.

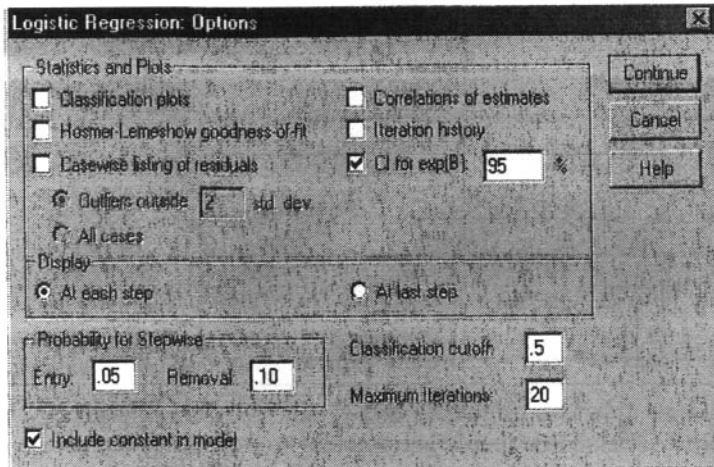


Fig. 7.2. Logistic regression: Options.

- Click on **OK**. Does your output look like Output 7.1?

Output 7.1: Logistic Regression, Method = Enter

```
LOGISTIC REGRESSION VAR alg2
/METHOD=ENTER gender mosaic visual parEduc
/PRINT=CI(95)
/CRITERIA = PIN(.05) POUT(.10) ITERATE(20) CUT(.5) .
```

Case Processing Summary

Unweighted Cases ^a		N	Percent
Selected Cases	Included in Analysis	75	100.0
	Missing Cases	0	.0
	Total	75	100.0
Unselected Cases		0	.0
Total		75	100.0

No participants have missing data.

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
not taken	0
taken	1

Algebra 2 is the dependent outcome variable and is coded 0 or 1.

Block 0: Beginning Block

Classification Table^{a,b}

Observed			Predicted		Percentage Correct
			algebra 2 in h.s.		
			not taken	taken	
Step 0	algebra 2 in h.s.	not taken	40	0	100.0
		taken	35	0	.0
Overall Percentage					53.3

a. Constant is included in the model.

b. The cut value is .500

40 students didn't take *algebra 2* and 35 did. 53% did not take *algebra 2*. Thus, if one simply guessed that no students took *algebra 2*, one would classify 53.3% of the students correctly by chance.

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 0 Constant	-.134	.231	.333	1	.564	.875

Only the constant is in the equation.

Variables not in the Equation

Gender, visual, and parents' education are each separately significantly related to *algebra 2*.

	Score	df	Sig.
Step 0 Variables			
GENDER	5.696	1	.017
MOSAIC	.480	1	.488
VISUAL	10.339	1	.001
PAREduc	12.559	1	.000
Overall Statistics	22.031	4	.000

Block 1: Method=Enter

Omnibus Tests of Model Coefficients

	Chi-square	df	Sig.
Step 1 Step	24.231	4	.000
Block	24.231	4	.000
Model	24.231	4	.000

The overall model is significant when all four independent variables are entered.

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	79.407	.276	.369

These are similar to R^2 and give a rough estimate of the variance that can be predicted from the combination of the four variables.

Note, 33/40 (83%) of those who didn't take *algebra 2* were predicted correctly with this model, but 71% of those who did were predicted correctly.

Classification Table^a

Observed			Predicted		
			algebra 2 in h.s.		Percentage Correct
			not taken	taken	
Step 1	algebra 2 in h.s.	not taken	33	7	82.5
		taken	10	25	71.4
Overall Percentage					77.3

a. The cut value is .500

Parents' education and *visualization*, but not *gender*, are significant predictors when all four variables are considered together. This suggests some correlation among predictors since *gender* was a significant predictor when used alone. *Mosaic* fails to predict *algebra2* whether used alone or with the other predictors.

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I. for EXP(B)	
								Lower	Upper
Step 1	GENDER	-.497	.577	.742	1	.389	.609	.197	1.884
	MOSAIC	-.030	.031	.892	1	.345	.971	.913	1.032
	VISUAL	.190	.075	6.428	1	.011	1.209	1.044	1.400
	PAREduc	.380	.131	8.418	1	.004	1.462	1.131	1.889
	Constant	-1.736	1.159	2.243	1	.134	.176		

a. Variable(s) entered on step 1: GENDER, MOSAIC, VISUAL, PAREduc.

Interpretation of Output 7.1

There are three main parts to this output. First, two tables provide descriptive information (see callout boxes). Next, there are three tables in **Block 0** that provide information about the baseline situation, when only the constant is in the equation or model. That is, how well can we predict *algebra 2* without using *gender*, *parents' education*, *mosaic*, or *visualization test*, if we predicted that all the students would *not take algebra 2*? The last section of four tables is below **Block 1**. They show the results when the four predictors are entered simultaneously.

The first table under **Block 0**, the initial **Classification Table**, shows the percentage of correct predictions (53%) if all of the students were predicted to be in the larger (*algebra 2 not taken*) group. The first **Variables in the Equation** table shows that if you predicted that all students would not take *algebra 2*, the odds of successful prediction would *not* be significantly different from 50-50 (i.e., no better than chance).

The **Variables not in the Equation** table shows that three of the four variables (*gender*, *parents' education*, and *visualization test*) are, individually, significant predictors of whether a student would take *algebra 2* or not. *Mosaic* is not a significant predictor.

Logistic regression in SPSS provides several ways to interpret the results of using these four variables (*gender*, *parents' education*, *mosaic* and *visualization test*) as predictors of whether or not students took *algebra 2*. The last four tables in Output 7.1 show these several ways. The **Omnibus Tests of Model Coefficients** table indicates that, when we consider all four predictors together, the **Model** or equation is significant ($\chi^2 = 24.23$, $df = 4$, $N = 75$, $p < .001$).

The **Model Summary** table includes two different ways of estimating R^2 (percent of variance accounted for) as was done in multiple regression. These “pseudo” R^2 estimates (.28 and .37) indicate that approximately 28% or 37% of the variance in whether or not students took *algebra 2* can be predicted from the linear combination of the four independent variables. The Cox & Snell R^2 (28%) is usually an underestimate.

The final **Classification Table** indicates how well the combination of variables predicts *algebra 2*. In this problem we have tried to predict, from four other variables, whether or not students would take *algebra 2*. Note from the classification table that, overall, 77% of the participants were predicted correctly. The independent/covariate variables were better at helping us predict who would *not* take *algebra 2* (83% correct) than at who would take it (71% correct).

Note, in the **Variables in the Equation** table, that only *parents' education* and *visualization test* are significant. *Gender* is not significant, which is probably due to several factors: 1) the fact that SE is quite high relative to B, which makes the Wald statistic lower, 2) the fact that *gender* is dichotomous, and 3) the fact that in this sample, *gender* is modestly (.25 and .28) but significantly related to *visualization test* and *parents' education*, so when they are already included, *gender* does not add enough to be significant ($p = .389$). Note that **Exp(B)** gives the odds ratios for each variable. The odds ratio and confidence interval for *parents' education* was 1.46 (95% CI = 1.13-1.89) and for *visualization test* was 1.21 (CI = 1.04-1.4). These indicate that the odds of estimating correctly who takes *algebra 2* improve by 46% (1.46-1) if one knows *parents' education* and by about 21% if one knows the *visualization test* score.

Example of How to Write About Problem 7.1

Results

Logistic regression was conducted to assess whether the four predictor variables, gender, parents' education, mosaic pattern test, and visualization test, significantly predicted whether or not a student took algebra 2. When all four predictor variables are considered together, they significantly predict whether or not a student took algebra 2, $\chi^2 = 24.23$, $df = 4$, $N = 75$, $p < .001$. Table 7.1 presents the odds ratios, which suggest that the odds of estimating correctly who takes algebra 2 improve by 46% if one knows parents' education and by about 21% if one knows students' visualization scores.

Table 7.1

Logistic Regression Predicting Who Will Take Algebra 2

Variable	β	SE	Odds ratio	p
Gender	-.50	.58	.61	.389
Parents' education	.38	.13	1.46	.004
Mosaic	-.03	.03	.97	.345
Visualization test	.19	.08	1.21	.011
Constant	-1.74	1.16	.18	.134

Problem 7.2: Hierarchical Logistic Regression

We will rerun Problem 7.1, but this time ask the computer to enter the background variables *gender* and *parents' education* first and then, on the second step or block, enter *mosaic* and *visualization test*.

7.2. If we control for *gender* and *parents' education*, will *mosaic* and/or *visualization test* add to the prediction of whether students will take *algebra 2*?

Now use the same dependent variable and covariates except enter *gender* and *parents' education* in **Block 1** and then enter *mosaic* and *visualization test* in **Block 2**.

- Use these commands: **Analyze => Regression => Binary Logistic.**
- Click on **Reset.**
- Move *algebra 2 in h.s.* into the **Dependent** variable box.
- Move *gender* and *parents' education* into the **Covariates** box (in this problem we actually are treating them as covariates in that we remove variance associated with them first; however, in SPSS all predictors are called **Covariates**).
- Make sure **Enter** is the selected **Method**.
- Click on **Next** to get **Block 2 of 2** (see Fig. 7.1 if you need help).
- Move *mosaic* and *visualization test* into the **Covariates** box.
- Click on **OK**. Does your output look like Output 7.2?

Output 7.2: Logistic Regression

```
LOGISTIC REGRESSION alg2
/METHOD = ENTER gender parEduc /METHOD = ENTER mosaic visual
/CRITERIA = PIN(.05) POUT(.10) ITERATE(20) CUT(.5) .
```

Case Processing Summary

Unweighted Cases ^a		N	Percent
Selected Cases	Included in Analysis	75	100.0
	Missing Cases	0	.0
	Total	75	100.0
Unselected Cases		0	.0
Total		75	100.0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
not taken	0
taken	1

Again, *algebra 2* is the dependent variable and is coded 0 for not taken and 1 for taken.

Block 0: Beginning Block

Classification Table^{a,b}

Observed			Predicted		
			algebra 2 in h.s.		Percentage Correct
			not taken	taken	
Step 0	algebra 2 in h.s.	not taken	40	0	100.0
		taken	35	0	.0
Overall Percentage					53.3

a. Constant is included in the model.

b. The cut value is .500

If we predicted that no one would take *algebra 2*, we would be correct 53% of the time, which is not significant, $p = .56$.

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
Step 0 Constant	-.134	.231	.333	1	.564	.875

Both *gender* and *parents' education* are significant predictors when entered separately.

Variables not in the Equation

	Score	df	Sig.
Step 0 Variables GENDER	5.696	1	.017
PAREduc	12.559	1	.000
Overall Statistics	14.944	2	.001

Block 1: Method=Enter

Gender and *parents' education* are entered in Block 1, the first step.

Omnibus Tests of Model Coefficients

	Chi-square	df	Sig.
Step 1 Step	16.109	2	.000
Block	16.109	2	.000
Model	16.109	2	.000

The combination of *gender* and *parents' education* significantly predicts who will take *algebra 2*.

Chapter 7 – Logistic Regression and Discriminant Analysis

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	87.530	.193	.258

These are estimates of how much knowing student's *gender* and *parents' education* helps you predict whether or not the student will take *algebra 2*.

Classification Table^a

Observed			Predicted		
			algebra 2 in h.s.		Percentage Correct
			not taken	taken	
Step 1	algebra 2 in h.s.	not taken	32	8	80.0
		taken	14	21	60.0
Overall Percentage					70.7

a. The cut value is .500

We can predict who won't take *algebra 2* (80%) better than we can predict who will (60%).

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1	GENDER	-.858	.522	2.707	1	.100	.424
	PAREduc	.374	.127	8.638	1	.003	1.454
	Constant	-1.297	.681	3.625	1	.057	.273

a. Variable(s) entered on step 1: GENDER, PAREduc.

When both *gender* and *parents' education* are entered, *gender* is no longer significant.

Block 2: Method=Enter

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	8.123	2	.017
	Block	8.123	2	.017
	Model	24.231	4	.000

Note that adding *mosaic* and *visual* to the equation increases the prediction significantly.

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	79.407	.276	.369

Note that these pseudo R^2 's and percentage correct are higher than they were when only *gender* and *parents' education* were entered, and they are the same as we found when all variables were entered simultaneously.

Classification Table^a

Observed			Predicted		
			algebra 2 in h.s.		Percentage Correct
			not taken	taken	
Step 1	algebra 2 in h.s.	not taken	33	7	82.5
		taken	10	25	71.4
Overall Percentage					77.3

a. The cut value is .500

Note that neither *gender* nor *mosaic* is significant when all of these variables are entered together.

Variables in the Equation

Step		B	S.E.	Wald	df	Sig.	Exp(B)
1	GENDER	-.497	.577	.742	1	.389	.609
	PAREduc	.380	.131	8.418	1	.004	1.462
	MOSAIC	-.030	.031	.892	1	.345	.971
	VISUAL	.190	.075	6.428	1	.011	1.209
	Constant	-1.736	1.159	2.243	1	.134	.176

a. Variable(s) entered on step 1: MOSAIC, VISUAL.

Interpretation of Output 7.2

The first four tables are the same as in Output 7.1. In this case, we have an additional step or block. **Block 1** shows the **Omnibus Chi-Square**, **Model Summary**, **Classification Table**, and **Variables in the Equation** when *gender* and *parents' education* were entered as covariates. Note that the Omnibus Test is statistically significant ($\chi^2 = 16.11, p < .001$). With only *gender* and *parents' education* entered, overall we can predict correctly 71% of the cases. Note from the last table in Block 1, that *gender* is not significant ($p = .100$) when it and *parents' education* are both in the equation.

In **Block 2**, we entered *mosaic* and *visualization test* to see if they would add to the predictive power of *gender* and *parents' education*. They do, as indicated by the **Step** in the **Omnibus Tests** table ($\chi^2 = 8.12, p = .017$). Note in the same table that the overall **Model** ($\chi^2 = 24.23, p < .001$) with all four predictors entered is significant. In this example, there is only one step in each block so the four tables in **Block 2** all say **Step 1**. The last three tables, as would be expected, are the same as those from Problem 7.1.

Problem 7.3: Discriminant Analysis (DA)

Discriminant analysis is appropriate when you want to predict which group (in this example, who took *algebra 2*) participants will be in. The procedure produces a discriminant function (or for more than two groups, a set of discriminant functions) based on linear combinations of the predictor variables that provide the best overall discrimination among the groups. The grouping or dependent variable can have more than two values, but this will greatly increase the complexity of the output and interpretation. The

codes for the grouping variable must be integers. You need to specify their minimum and maximum values, as we will in Fig. 7.3. Cases with values outside these bounds are excluded from the analysis.

Discriminant analysis (DA) is similar to multivariate analysis of variance (MANOVA, discussed in chapter 10), except that the independent and dependent variables are switched; thus, the conceptual basis for the study is typically different. In DA, one is trying to devise one or more predictive equations to maximally discriminate people in one group from those in another group; in MANOVA, one is trying to determine whether group members differ significantly on a set of several measures. The assumptions for DA are similar to those for MANOVA.

Assumptions of Discriminant Analysis

The assumptions of discriminant analysis include that the relationships between all pairs of predictors must be linear, multivariate normality must exist within groups, and the population covariance matrices for predictor variables must be equal across groups. Discriminant analysis is, however, fairly robust to these assumptions, although violations of multivariate normality may affect accuracy of estimates of the probability of correct classification. If multivariate nonnormality is suspected, then logistic regression should be used. Multicollinearity is again an issue with which you need to be concerned. It is also important that the sample size of the smallest group (35 in this example) exceed the number of predictor variables in the model (there are four in this example, so this assumption is met). The linearity assumption as well as the assumption of homogeneity of variance-covariance matrices can be tested, as we did for multiple regression in chapter 6, by examining a matrix scatterplot. If the spreads of the scatterplots are roughly equal, then the assumption of homogeneity of variance-covariance matrices can be assumed. This assumption can also be tested with Box's M.

7.3 What combination of *gender*, *parents' education*, *mosaic*, and *visualization test* best distinguishes students who will take *algebra 2* from those who do not?

This is a similar question to the one that we asked in Problem 7.1, but this time we will use discriminant analysis, so the way of thinking about the problem is a little different.

You can use discriminant analysis instead of logistic regression when you have a categorical outcome or grouping variable, if you have all continuous independent variables. It is best not to use dichotomous predictors for discriminant analysis, except when the dependent variable has a nearly 50-50 split as is true in this case.

We created scatterplots (below) after splitting the data by the grouping variable (in this case *algebra 2 in h.s.*) so we could check the assumption of homogeneity of variance-covariance matrices across groups. To split the file into groups, use **Split File**. See the Quick Reference Guide (Appendix A) if you need assistance. Then do a scatterplot, selecting "matrix," as shown in chapter 5, with *gender*, *parents' education*, *mosaic*, and *visualization test* as the variables. Don't forget to turn off the split file command before you run other analyses.

Next, follow these steps to do the Discriminant Analysis:

- Select **Analyze => Classify => Discriminant...**
- Move *algebra 2 in h.s.* into the **Grouping Variable** box (see Fig. 7.3).

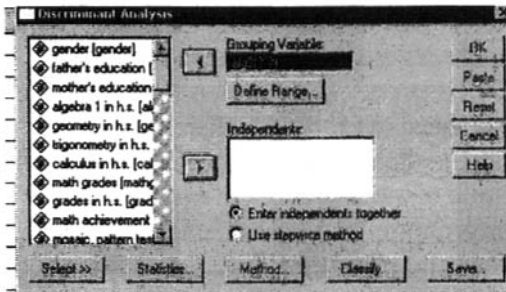


Fig. 7.3. Discriminant analysis.

- Click on **Define Range** and enter **0** for **Minimum** and **1** for **Maximum** (see Fig. 7.4).
- Click on **Continue** to return to Fig. 7.3.

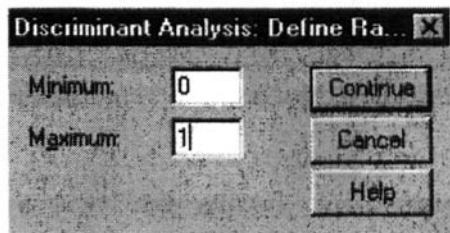


Fig. 7.4. Discriminant analysis: Define range.

- Now move *gender*, *parents' education*, *mosaic*, and *visualization test* into the **Independents** box.
- Make sure **Enter independents together** is selected.
- Click on **Statistics**.
- Select **Means**, **Univariate ANOVAs**, and **Box's M** (see Fig. 7.5). Click on **Continue**.

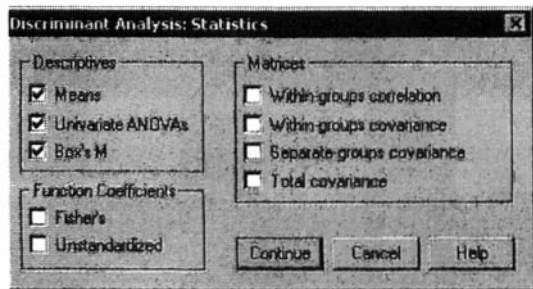


Fig. 7.5. Discriminant analysis: Statistics.

- Click on **Classify** to get Fig. 7.6.
- Check **Summary Table** under **Display**.
- Click on **Continue**.

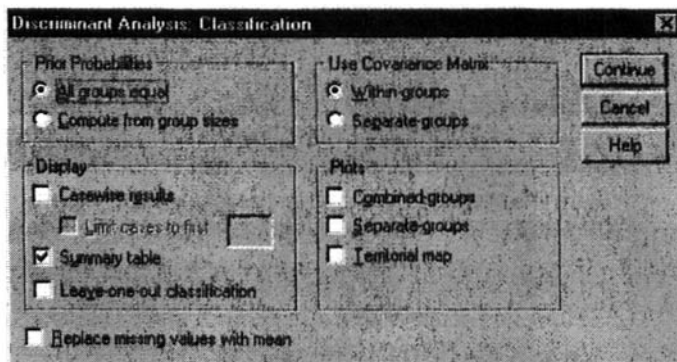


Fig. 7.6. Discriminant analysis: Classification.

- Finally, click on **OK** and compare your output to Output 7.3.

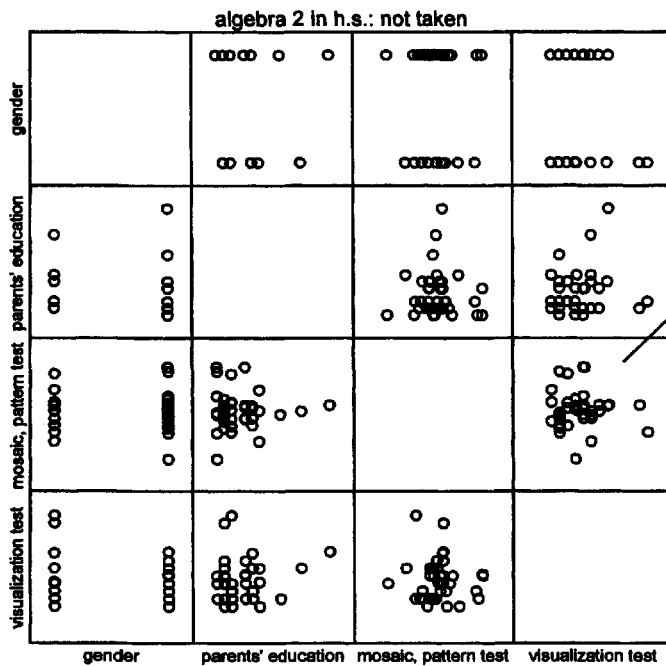
Output 7.3: Discriminant Analysis, Enter Independents Together

```

SORT CASES BY alg2 .
SPLIT FILE
  LAYERED BY alg2 .
GRAPH
  /SCATTERPLOT(MATRIX)=gender parEduc mosaic visual
  /MISSING=LISTWISE .
SPLIT FILE
  OFF.

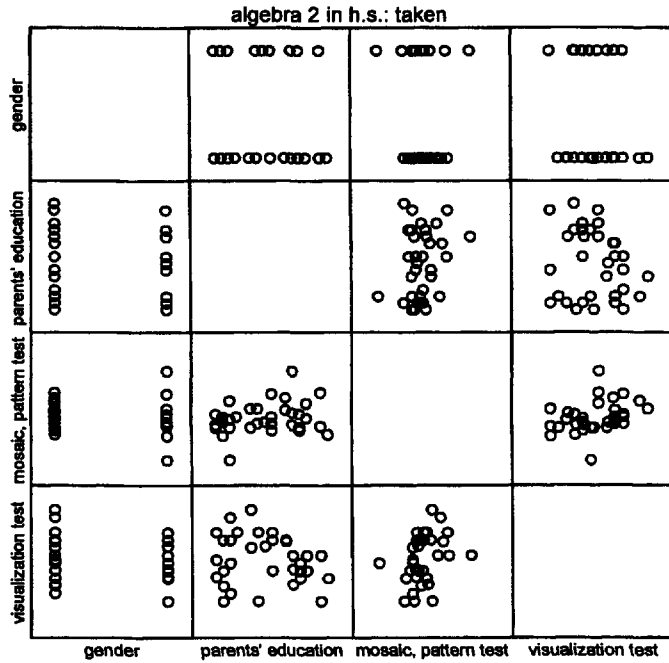
DISCRIMINANT
  /GROUPS=alg2(0 1)
  /VARIABLES=gender parEduc mosaic visual
  /ANALYSIS ALL
  /PRIORS EQUAL
  /STATISTICS=MEAN STDDEV UNIVF BOXM TABLE
  /CLASSIFY=NONMISSING POOLED .
  
```

Graph



Compare the comparable scatterplots for the two groups. The scatterplots for the same variables appear to be very similar in variability for the two groups, suggesting that the assumption of homogeneity of variance-covariance matrices is met.

SPSS for Intermediate Statistics



Analysis Case Processing Summary

Unweighted Cases		N	Percent
Valid		75	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		75	100.0

Chapter 7 – Logistic Regression and Discriminant Analysis

Group Statistics

algebra 2 in h.s.		Mean	Std. Deviation	Valid N (listwise)	
				Unweighted	Weighted
not taken	gender	.6750	.47434	40	40.000
	parents' education	3.5125	1.68891	40	40.000
	mosaic, pattern test	28.1250	12.03188	40	40.000
	visualization test	3.8938	3.42122	40	40.000
taken	gender	.4000	.49705	35	35.000
	parents' education	5.4000	2.54026	35	35.000
	mosaic, pattern test	26.6000	5.67088	35	35.000
	visualization test	6.7857	3.91037	35	35.000
Total	gender	.5467	.50117	75	75.000
	parents' education	4.3933	2.31665	75	75.000
	mosaic, pattern test	27.4133	9.57381	75	75.000
	visualization test	5.2433	3.91203	75	75.000

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
gender	.924	6.000	1	73	.017
parents' education	.833	14.683	1	73	.000
mosaic, pattern test	.994	.470	1	73	.495
visualization test	.862	11.672	1	73	.001

Gender, parents' education, and visualization are each significant predictors by themselves. Mosaic is not.

Box's Test of Equality of Covariance Matrices

Log Determinants

algebra 2 in h.s.	Rank	Log Determinant
not taken	4	6.904
taken	4	6.258
Pooled within-groups	4	7.101

The ranks and natural logarithms of determinants printed are those of the group covariance matrices.

Test Results

Box's M		36.324
F	Approx.	3.416
	df1	10
	df2	24432.818
	Sig.	.000

This indicates that the assumption of homogeneity of the covariance matrices has not been met, according to this test. However, this test is strongly influenced by nonnormality and may not be accurate. We checked this assumption with matrix scatterplots for each group, and these suggested the assumption was not badly violated. We could use logistic regression instead of discriminant analysis if we are concerned about the degree to which we met this assumption.

Tests null hypothesis of equal population covariance matrices.

Analysis 1

Summary of Canonical Discriminant Functions

Eigenvalues

Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	.416 ^a	100.0	100.0	.542

a. First 1 canonical discriminant functions were used in the analysis.

Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	.706	24.692	4	.000

This indicates that the predictors significantly discriminate the groups.

Standardized Canonical Discriminant Function Coefficients

	Function
	1
gender	-.213
parents' education	.706
mosaic, pattern test	-.220
visualization test	.624

This table indicates how heavily each variable is weighted in order to maximize discrimination of groups. In this example, *parents' education* and *visual* are weighted more than *gender* and *mosaic*.

Chapter 7 – Logistic Regression and Discriminant Analysis

Structure Matrix

	Function
	1
parents' education	.695
visualization test	.620
gender	-.445
mosaic, pattern test	-.124

These are the correlations of each independent variable with the standardized discriminant function.

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions
Variables ordered by absolute size of correlation within function.

Functions at Group Centroids

	Function
	1
algebra 2 in h.s.	
not taken	-.595
taken	.680

Unstandardized canonical discriminant functions evaluated at group means

Classification Statistics

Classification Processing Summary

Processed		75
Excluded	Missing or out-of-range group codes	0
	At least one missing discriminating variable	0
Used in Output		75

Prior Probabilities for Groups

	Prior	Cases Used in Analysis	
		Unweighted	Weighted
algebra 2 in h.s.			
not taken	.500	40	40.000
taken	.500	35	35.000
Total	1.000	75	75.000

Classification Results^a

			Predicted Group Membership		Total
			not taken	taken	
Original	Count	algebra 2 in h.s.			
		not taken	32	8	40
		taken	10	25	35
	%				
		not taken	80.0	20.0	100.0
		taken	28.6	71.4	100.0

This shows how well the model predicts who will take *algebra 2*. For example, 80% of those who did *not take algebra 2* were correctly predicted.

^a 76.0% of original grouped cases correctly classified.

Interpretation of Output 7.3

The **Group Statistics** table provides basic descriptive statistics for each of the independent/ predictor variables for each outcome group (didn't take *algebra 2* and did take it) separately and for the whole sample. The **Tests of Equality of Group Means** table shows which independent variables are significant predictors by themselves; it shows the variables with which there is a statistically significant difference between those who took *algebra 2* and those who didn't. As was the case when we used logistic regression, *gender*, *parents' education*, and *visualization test* are statistically significant.

Note, from the **Standardized Canonical Discriminant Function Coefficients** table, that only *parents' education* and *visualization test* are weighted heavily to maximize the discrimination between groups. Because *gender* correlates with *parents' education*, it has a low function coefficient. But in the **Structure Matrix** table, *gender* has a higher (-.45) correlation because it is correlated with the discriminant function that best predicts who took *algebra 2* and who did not.

The last table is similar to the classification table for logistic regression. It shows how well the combination of four independent variables classifies or predicts who will take *algebra 2*. Note that overall, 76% of the sample was classified correctly. As with logistic regression, discriminant analysis did better at predicting who did not take *algebra 2* (80% correct) than it did at predicting who would take it (71% correct).

Example of How to Write About Problem 7.3**Results**

Discriminant analysis was conducted to assess whether the four predictors, *gender*, *parents' education*, *mosaic*, and *visualization test*, could distinguish those who took *algebra 2* from those who did not. Wilks' lambda was significant, $\lambda = .71$, $\chi^2 = 24.69$, $p < .001$, which indicates that the model including these four variables was able to significantly discriminate the two groups. Table 7.3 presents the standardized function coefficients, which suggest that *parents' education* and the *visualization test* contribute most to distinguishing those who took *algebra 2* from those who did not, using these predictors. The classification results show that the model correctly predicts 80% of those who did not take *algebra 2* and 71% of those who did take *algebra 2*. The correlation coefficients in the table indicate the extent to which each variable correlates with the resulting discriminant function. Note that even though *mosaic* didn't contribute strongly to the discriminant function, it is moderately highly (negatively) correlated with the overall discriminant function.

Table 7.3

Standardized Function Coefficients and Correlation Coefficients

	Standardized Function Coefficients	Correlations between variables and discriminant function
Parents' education	.71	.62
Visualization test	.62	-.12
Gender	-.21	.70
Mosaic, pattern test	-.22	-.45

Interpretation Questions

- 7.1. Using Output 7.1: a) Which variables make significant contributions to predicting who took *algebra 2*? b) How accurate is the overall prediction? c) How well do the significant variables predict who took algebra 2? d) How about the prediction of who *didn't* take it?
- 7.2. Compare Outputs 7.1 and 7.2. How are they different and why?
- 7.3. In Output 7.3: a) What do the discriminant function coefficients and the structure coefficients tell us about how the predictor variables combine to predict who took *algebra 2*? b) How accurate is the prediction/classification overall and for who would not take *algebra 2*? c) How do the results in Output 7.3 compare to those in Output 7.1, in terms of success at classifying and contribution of different variables to the equation?
- 7.4. Comparing Outputs 7.3 and 7.1, what kind of information does the discriminant analysis provide that is not provided by the logistic regression?
- 7.5. In Output 7.2: Why might one want to do a hierarchical logistic regression?
- 7.6. a) In Output 7.1: How can one tell which variables are contributing more to the classification of participants using logistic regression? b) In Output 7.3: How can one tell which variables are contributing to the classification of participants using discriminant analysis? What is the difference between the function coefficients and the coefficients in the structure matrix?

Extra SPSS Problems

- 7.1. A marketing firm would like to understand why people do or do not buy their product. Retrieve the data set called 'satisf' and conduct a logistic regression with Made purchase (*purchase*) as the dependent variable and the variables Shopping frequency (*regular*), Price satisfaction (*price*), Variety satisfaction (*numitems*), and Organization satisfaction (*org*) as covariates.
 - a. Is the logistic regression significant?
 - b. What overall percentage of consumers were correctly classified as to whether or not they purchased the product? How accurately were purchasers predicted? How accurately were non-purchasers predicted?

- c. Which predictor(s) contribute significantly to predicting whether or not individuals purchased the product? What do the odds ratios say about how useful these predictors were?
- d. Approximately how much variance in consumers' tendency to purchase the product could be predicted from these variables? If you owned the company that sold these products and you wanted to understand who bought your product, would you be wise to find out about these predictor variables or would you use other predictors, given the results of this analysis?

7.2. Do a discriminant function analysis using these same variables.

- a. Is the discriminative function significant?
- b. Which predictor(s) contribute significantly to predicting whether or not individuals purchased the product?
- c. What overall percentage of consumers were correctly classified as to whether or not they purchased the product? How accurately were purchasers predicted? How accurately were non-purchasers predicted? How do these results differ from those for the logistic regression?
- d. Given the results of both of these analyses, what would you conclude about your understanding of these consumers' tendency to purchase the products? Describe your conclusions in non-technical terms, making sure to describe the statistical results of the discriminant function analysis.

CHAPTER 8

Factorial ANOVA and ANCOVA

In this chapter, we will introduce two complex difference statistics: factorial ANOVA and ANCOVA. Both factorial ANOVA and ANCOVA tell you whether considering more than one independent variable at a time gives you additional information over and above what you would get if you did the appropriate *basic* inferential statistics for each independent variable separately. Both of these inferential statistics have two or more independent variables and one scale (normally distributed) dependent variable. **Factorial ANOVA** is used when there is a small number of independent variables (usually two or three) and each of these variables has a small number of levels or categories (usually two to four).

ANCOVA typically is used to adjust or control for differences between the groups based on another, typically interval level, variable called the covariate. For example, imagine that we found that boys and girls differ on math achievement. However, this could be due to the fact that boys take more math courses in high school. ANCOVA allows us to adjust the math achievement scores based on the relationship between number of math courses taken and math achievement. We can then determine if boys and girls still have different math achievement scores after making the adjustment. ANCOVA can also be used if one wants to use one or more discrete or nominal variables and one or two continuous variables to predict differences in one dependent variable.

- Retrieve **hsbdataB** from your data file.

Problem 8.1: Factorial (2-Way) ANOVA

We would use a *t* test or one-way ANOVA to examine differences on a scale dependent variable between two or more groups comprising the levels of *one* independent variable or factor. These designs, in which there is only one independent variable, and it is a discrete or categorical variable, are called single factor designs. In this problem, we will compare groups formed by combining *two* independent variables. The appropriate statistic for this type of problem is called a **two factor, 2-way, or factorial ANOVA**. One can also have factorial ANOVAs in which there are more than two independent variables. If there are three independent variables, one would have a three factor or 3-way ANOVA. It is unusual, but possible, to have more than three factors as well. Factorial ANOVA is used when there are two or more independent variables (each with a few categories or values) and a between groups design.

Assumptions of Factorial ANOVA

The assumptions for factorial ANOVA are that the observations are independent, the variances of the groups are equal (homogeneity of variances), and the dependent variable is normally distributed for each group. Assessing whether the observations are independent (i.e., each participant's score is not related systematically to any other participant(s)'s score) is a design issue that should be evaluated prior to entering the data into SPSS. To test the assumption of homogeneity of variances, SPSS computes the Levene Statistic, which can be requested using the General Linear Model command. We will demonstrate it below. Factorial ANOVA is robust against violations of the assumption of the normal distributions of the dependent variable. One way to test this assumption is to compare box plots of the dependent variable for each group (cell) defined by each combination of the levels of the independent variables.

- 8.1 Are there differences in *math achievement* for people varying on *math grades* and/or *father's education revised*, and is there a significant interaction between *math grades* and *father's education* on *math achievement*? (Another way to ask this latter question: Do the "effects" of *math grades* on *math achievement* vary depending on level of *father's education revised*?)

Follow these commands:

- **Analyze => General Linear Model => Univariate.**
- Move *math achievement* to the **Dependent (variable)** box.
- Move the first independent variable, *math grades*, to the **Fixed Factor(s)** box and then move the second independent variable, *father's educ revised* (Not father's education), to the **Fixed Factor(s)** box. (See Fig. 8.1.)

Now that we know the variables we will be dealing with, let's determine our options.

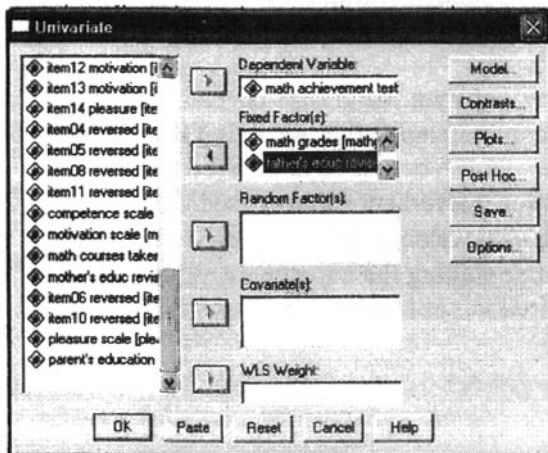


Fig. 8.1. GLM: Univariate.

- Click on **Plots** and move *faedRevis* to the **Horizontal Axis** and *mathgr* to the **Separate Lines** box in Fig. 8.2. This “profile plot” will help you picture the interaction (or absence of interaction) between your two independent variables. Note, the plots will be easier to interpret if you put *father's educ revised* with its three values on the horizontal axis.
- Then press **Add**. You will see that *mathgr* and *faedRevis* have moved to the **Plots** window as shown at the bottom of Fig. 8.2.
- Click on **Continue** to get back Fig.8.1.

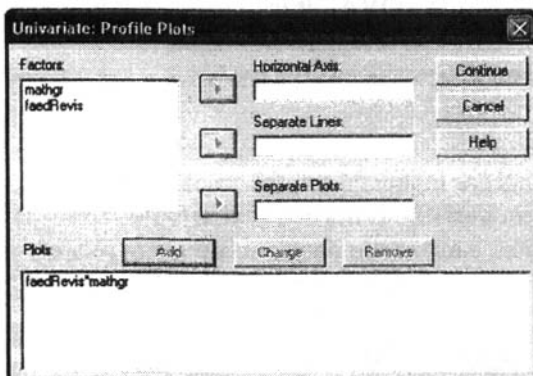


Fig. 8.2. Univariate: Profile plots.

- Select **Options** and check **Descriptive statistics**, **Estimates of effect size**, and **Homogeneity tests** in Fig. 8.3.

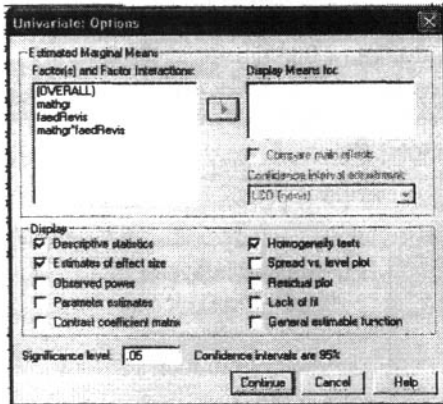


Fig. 8.3. Univariate: Options.

- Click on **Continue**. This will take you back to Fig. 8.1.
- Click on **OK**. Compare your output to Output 8.1.

Output 8.1: GLM General Factorial (2-Way) ANOVA

```
UNIANOVA
  mathach BY mathgr faedRevis
  /METHOD = SSTYPE(3)
  /INTERCEPT = INCLUDE
  /PLOT = PROFILE( faedRevis*mathgr )
  /PRINT = DESCRIPTIVE ETASQ HOMOGENEITY
  /CRITERIA = ALPHA(.05)
  /DESIGN = mathgr faedRevis mathgr*faedRevis .
```

Univariate Analysis of Variance

Between-Subjects Factors

		Value Label	N
math	0	less A-B	43
grades	1	most A-B	30
father's	1.00	HS grad or less	38
educ	2.00	Some College	16
revised	3.00	BS or More	19

Descriptive Statistics

Dependent Variable: math achievement test

math grades	father's education revised	Mean	Std. Deviation	N
less A-B	HS grad or less	9.8261	5.03708	23
	Some College	12.8149	5.05553	9
	BS or More	12.3636	7.18407	11
	Total	11.1008	5.69068	43
most A-B	HS grad or less	10.4889	6.56574	15
	Some College	16.4284	3.43059	7
	BS or More	21.8335	2.84518	8
	Total	14.9000	7.00644	30
Total	HS grad or less	10.0877	5.61297	38
	Some College	14.3958	4.66544	16
	BS or More	16.3509	7.40918	19
	Total	12.6621	6.49659	73

These six-cell means will be shown in the plot.

Levene's Test of Equality of Error Variances^a

Dependent Variable: math achievement test

F	df1	df2	Sig.
2.548	5	67	.036

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

a. Design: Intercept+MATHGR+FAEDR+MATHGR * FAEDR

This indicates that the assumption of homogeneity of variances has been violated. Because Levene's test is significant, we know that the variances are significantly different. Luckily, SPSS uses the regression approach to calculate ANOVA, so this problem is less important.

Tests of Between-Subjects Effects

Dependent Variable: math achievement test

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	1029.236(a)	5	205.847	6.863	.000	.339
Intercept	12094.308	1	12094.308	403.230	.000	.858
mathgr	325.776	1	325.776	10.862	.002	.139
faedRevis	646.015	2	323.007	10.769	.000	.243
mathgr * faedRevis	237.891	2	118.946	3.966	.024	.106
Error	2009.569	67	29.994			
Total	14742.823	73				
Corrected Total	3038.804	72				

Eta squared indicates that 24% of the variance in *math achievement* can be predicted from father's education.

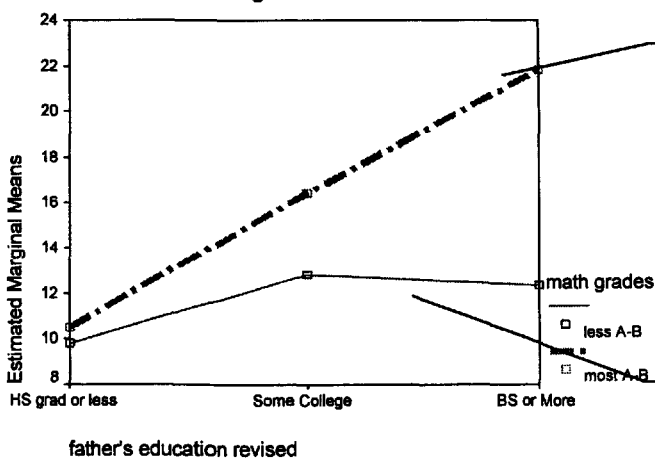
a. R Squared = .339 (Adjusted R Squared = .289)

The *R Squared* value is the percent of variance in *math achievement* predictable from both independent variables and the interaction.

Focus on these three *F*s, especially the *Mathgr x FaedRevis* interaction.

Profile Plots

Estimated Marginal Means of math achievement t



Plot for students with high *math grades*.

Note that for students whose fathers have a BS or more, there is a big difference in *math achievement* between those with high and low *math grades*. However, there is little difference for students whose fathers are high school grads or less.

Plot for students with low *math grades*.

Interpretation of Output 8.1

The GLM Univariate program allows you to print the means and counts, provides measures of effect size (eta), and plots the interaction, which is helpful in interpreting it. The first table in Output 8.1 shows that 73 participants (43 with low *math grades* and 30 high *math grades*) are included in the analysis because they had data on all of the variables. The second table, **Descriptive Statistics**, shows the cell and marginal (total) means; both are very important for interpreting the ANOVA table and explaining the results of the test for the interaction.

The ANOVA table, called **Tests of Between Subjects Effects**, is the key table. Note that the word “effect” in the title of the table can be misleading because this study was not a randomized experiment. Thus, you should not report that the differences in the dependent variable were *caused* by the independent variable. Usually you will ignore the information about the corrected model and intercept and skip down to the interaction F (*mathgr*faedRevis*). It is important to look at the interaction first because it may change the interpretation of the separate “main effects” of each independent variable.

In this case, the interaction is statistically significant, $F(2,67)=3.97, p=.024$. This means that the “effect” of *math grades* on *math achievement* depends on which *father’s education* level is being considered. If you find a significant interaction, you should examine the **profile plots** of cell means to visualize the differential effects. If there is a significant interaction, the lines on the profile plot will not be parallel. In this case, the plot indicates that *math achievement* is relatively low for both groups of students whose fathers had relatively low education (*high school grad or less*). However, for students whose fathers have a high education level (*BS or more*), differences in *math grades* seem to have a large “effect” on *math achievement*. This interpretation, based on a visual inspection of the plots, needs to be checked with inferential statistics. When the interaction is statistically significant, you should analyze the “**simple effects**” (differences between means for one variable at each particular level of the other variable). We will illustrate one method for statistically analyzing the simple effects in Problem 8.2.

Now, examine the main effects of *math grades* and of *father’s education revised*. Note that both are statistically significant, but because the interaction is significant this is somewhat misleading. The plots show that the effect of *math grades* does not seem to hold true for those whose fathers had the least education. Note also the callout boxes about the adjusted R squared and eta squared. **Eta**, the correlation ratio, is used when the independent variable is *nominal* and the dependent variable (*math achievement* in this problem) is *scale*. Eta is an indicator of the proportion of variance that is due to between groups differences. **Adjusted R^2** refers to the multiple correlation coefficient, squared and adjusted for number of independent variables, N , and effect size. Like r^2 , eta squared and R^2 indicate how much variance or variability in the dependent variable can be predicted; however, the multiple R^2 is used when there are several independent variables, and the r^2 is used when there is only one independent variable. In this problem, the eta² values for the three key F s vary from .106 to .243. Because eta and R , like r , are indexes of association, they can be used to interpret the effect size. However, the guidelines according to Cohen (1988) for eta and R are somewhat different (for eta: small = .10, medium = .24, and large = .31; for R : small = .10, medium = .36, and large = .51).

In this example, eta for *math grades* is about .37 ($\sqrt{.139} = .37$) and, thus, according to Cohen (1988) a large effect. Eta for *father’s education* (revised) is about .49, a large effect. The interaction eta is about .33, a large effect. The overall **adjusted R** is about .54, a large effect, but not really bigger than for *father’s education* alone, when you consider the different criteria for “large” for eta vs. R . Notice that the adjusted R^2 is lower than the unadjusted (.29 vs. .34). The reason for this is that the adjusted R^2 takes into account (and adjusts for) several things including the fact that not just one but three factors (*mathgr*, *faedRevis*, and the interaction), some of which have more than two levels, were used to predict *math achievement*.

An important point to remember is that statistical significance depends heavily on the sample size so that with 1,000 subjects, a much lower F or r will be significant than if the sample is 10 or even 100. Statistical significance just tells you that you can be quite sure that there is at least a tiny relationship between the independent and dependent variables. Effect size measures, which are more independent of sample size, tell how strong the relationship is and, thus, give you some indication of its importance.

How to write the results for Problem 8.1 is included in the interpretation box for Problem 8.2.

Problem 8.2: Post Hoc Analysis of a Significant Interaction

We have described, in the interpretation of Output 8.1, how to visually inspect and interpret the Profile Plots when there is a statistically significant interaction. In Problem 8.2 we will illustrate one way to test the simple main effects statistically. In the interpretation of Output 8.1, we indicated that you should examine the interaction F first. If it is statistically significant, it provides important information and means that the results of the main effects may be misleading. Figure 8.4 is a decision tree that illustrates this point and guides the analysis in this section.

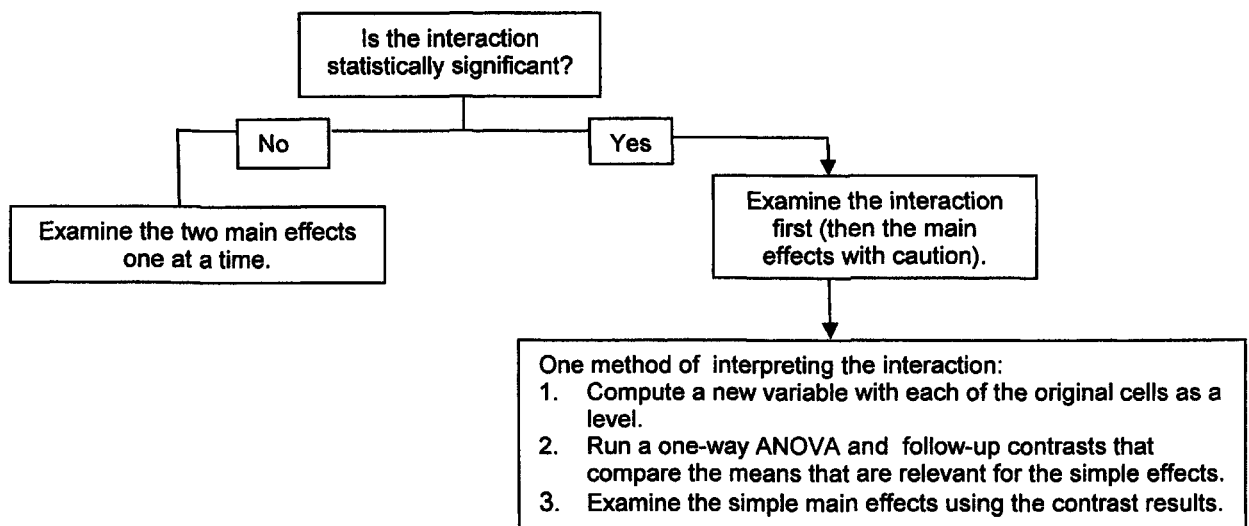


Fig. 8.4. Steps in analyzing a two-way factorial ANOVA.

8.2. Which simple main effects of *math grades* (at each level of *father's education revised*) are statistically significant?

- Select **Transform => Compute**. You will see the **Compute Variable** window. (See Fig. 8.5.)

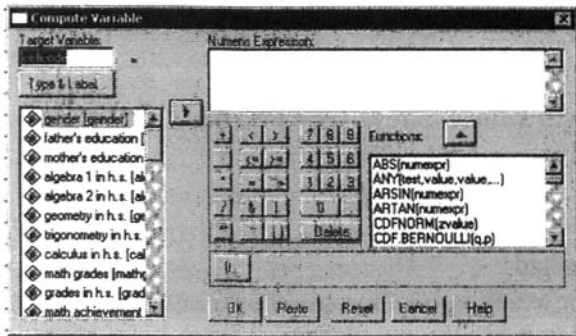


Fig. 8.5. Compute variable window.

- Under **Target Variable**, type *cellcode*. This is the name of the new variable you will compute.
- Click on **Type and Label**. You will see the **Compute Variable: Type and Label** window. (See Fig. 8.6.)
- Type *six new cell codes* in the label box.

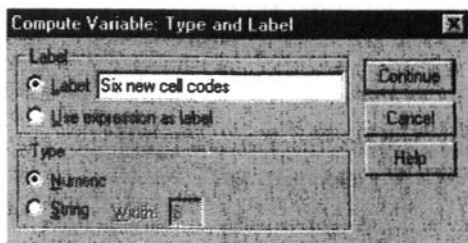


Fig. 8.6. Compute variable: Type and label.

- Click on **Continue**. You will see the **Compute Variable** window (Fig. 8.5) again.
- In the **Numeric Expression** box, type the number 1. This will be the first value or level for the new variable.
- Next, click on the **If...** button in Fig. 8.5 to produce Fig. 8.7.
- Select **Include if case satisfies condition**.
- Type *mathgr=0 & faedRevis=1* in the window. You are telling SPSS to compute level 1 of the new variable, *cellcode*, so that it combines the first level of *math grades* (0) and the first level of *father's education revised* (1). (See Fig. 8.7.)

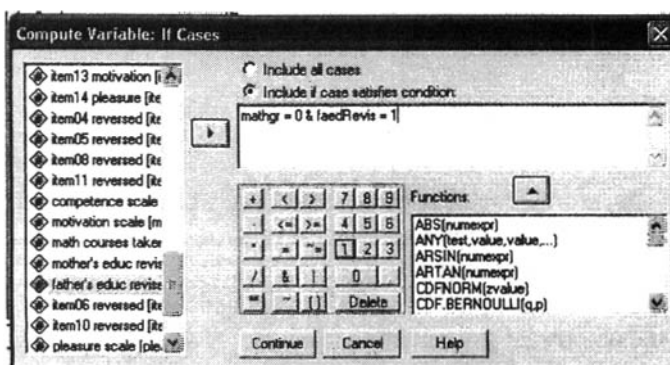


Fig. 8.7. Compute variable: If cases.

- Click on **Continue** to return to Fig. 8.5.
- Click on **OK**. If you look at your Data View or Variable View screen, you will see a new variable called *cellcode*. In the **Data View** screen, you should also notice that several of the cases now show a value of 1. Those are the cases where the original variables of *math grades* and *father's education revised* met the requirements you just computed.

You will need to repeat this process for the other five levels of this new variable. We will walk you through the process once more.

- Select **Transform=>Compute** to see the **Compute Variable** window. (See Fig. 8.5.)
- Ensure that *cellcode* is still in the target variable box.
- Type 2 in the **Numeric Expression** box.
- Click on the **If...** button.
- Ensure that **Include if case satisfies condition** is selected.
- Type *mathgr=1 & faedRevis=1* in the box. You are now telling SPSS to use the other (higher) level of *math grades* with the first level of *father's education revised*.
- Click on **Continue**.
- Click on **OK**.
- SPSS will ask you if you want to change the existing variable. Click on **OK**. This means that you want to add this second level to the *cellcode* variable. If you look at the **Data View** screen, you will notice that some of the cases now indicate a value of 2.
- Complete the above steps for the remaining four levels: Level 3 of cellcode: *mathgr=0 & faedrevi=2*; Level 4: *mathgr = 1 & faedRevis = 2*; Level 5: *mathgr = 0 & faedRevis = 3*; and Level 6: *mathgr = 1 & faedRevis = 3*.

Next, you will use this new variable, *cellcode*, to examine the statistical significance of differences among certain of the six cells using one-way ANOVA and the contrasts command. This will help us to interpret the simple effects that we discussed above.

- Select **Analyze=>Compare Means=>One Way ANOVA....** You will see the **One Way ANOVA** window. (See Fig. 8.8.)
- Move *math achievement* into the **Dependent List:** box by clicking on the top **arrow** button.
- Move *six new cell codes* into the **Factor:** box by clicking on the bottom **arrow** button.

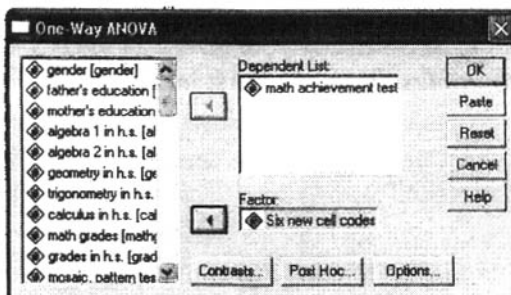


Fig. 8.8. One-way ANOVA.

Instead of doing Post Hoc tests, we will do **Contrasts**, which compare preselected pairs of means rather than all possible pairs of means. In this case, we want to compare the math achievement of high versus low math grade students at each of the three levels of father's education. This is done by telling SPSS specific cell means to contrast. For example, if we want to compare only the first and second means while ignoring the other four means, we would write a contrast statement as follows 1 -1 0 0 0 0. This can be done with the SPSS syntax or using the point and click method as shown below.

- Click on the **Contrasts...** button to see Fig. 8.9.
- Enter 1 in the **Coefficients:** window.
- Then click on **Add** which will move the 1 to the larger window.
- Next enter -1 and press **Add**; enter 0, press **Add**; enter another 0, press **Add**.

- Enter a third 0, press **Add**; enter a fourth 0. Fig 8.9 is how the window should look just before you press **Add** the final time. This compares the math achievement scores of students with low versus high *math grades* if their fathers have the lowest level of education.
- Now press **Next** so that the Fig. 8.9 says **Contrast 2 of 2**.
- Now enter the following coefficients as you did for the first contrast: 0 0 1 -1 0 0. Be sure to press **Add** after you enter each number. This compares the math achievement scores of students with low versus high grades if their fathers have the middle level of education.
- Press **Next** and enter the following **Coefficients** for **Contrast 3 of 3**: 0 0 0 0 1 -1. This compares the math achievement scores of students with low versus high grades if their fathers have the highest level of education.

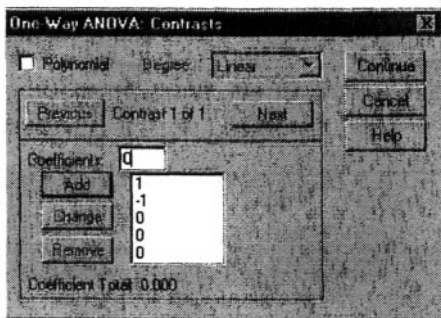


Fig. 8.9. One-way ANOVA: Contrasts.

Thus, what we have done with the above instructions is simple main effects, first comparing students with high and low *math grades* who have fathers with *less than a high school education*. Second, we have compared students with high and low *math grades* who have fathers with *some college*. Finally, we have compared students with high and low *math grades* who have fathers with a *B.S. or more*. Look back at how we computed the cellcode variable (or the syntax below) to see why this is true. Note that in cases like this it might be easier to type the syntax, which is part of the reason many experienced SPSS users prefer to use syntax. However, you must type the syntax exactly correct or the program will not run.

- Click on **Continue**.
- Click on **Options**.
- Click on **Descriptives and Homogeneity of Variance**.
- Click on **Continue**.
- Click on **OK**.

Output 8.2: One-Way ANOVA for Comparing New Cell Means Oneway

```
IF (mathgr=0 & faedrevi=1) cellcode = 1 .
VARIABLE LABELS cellcode 'Six new cell codes' .
EXECUTE .
IF (mathgr=1 & faedRevis=1) cellcode = 2 .
EXECUTE .
IF (mathgr=0 & faedRevis=2) cellcode = 3 .
EXECUTE .
IF (mathgr=1 & faedRevis=2) cellcode = 4 .
EXECUTE .
IF (mathgr=0 & faedRevis=3) cellcode = 5 .
EXECUTE .
IF (mathgr=1 & faedRevis=3) cellcode = 6 .
EXECUTE .
ONEWAY
  mathach BY cellcode
  /CONTRAST= 1 -1 0 0 0 0 /CONTRAST= 0 0 1 -1 0 0 /CONTRAST= 0 0 0 0 1 -1
  /STATISTICS DESCRIPTIVES HOMOGENEITY
  /MISSING ANALYSIS .
```

SPSS for Intermediate Statistics

Means to be compared

Descriptives

math achievement test

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
less A-B & hs grad or less	23	9.8261	5.03708	1.05030	7.6479	12.0043	2.33	21.0
most A-B & hs grad or less	15	10.4889	6.56574	1.69527	6.8529	14.1249	1.00	22.7
less A-B & some college	9	12.8148	5.05553	1.68518	8.9289	16.7009	5.00	18.7
most A-B & some college	7	16.4284	3.43059	1.29664	13.2557	19.6012	14.3	23.7
less A-B & BS or more	11	12.3636	7.18407	2.16606	7.5373	17.1900	1.00	23.7
most A-B & BS or more	8	21.8335	2.84518	1.00592	19.4549	24.2121	15.7	23.7
Total	73	12.6621	6.49659	.76037	11.1463	14.1779	1.00	23.7

Test of Homogeneity of Variances

math achievement test

Levene Statistic	df1	df2	Sig.
2.548	5	67	.036

The assumption of equal variances is violated, so we will select our contrast test accordingly (see below).

ANOVA

math achievement test

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1029.236	5	205.847	6.863	.000
Within Groups	2009.569	67	29.994		
Total	3038.804	72			

The overall F is significant at $p < .001$.

Contrast 1 looks at the difference between codes 1 and 2 (having lower and higher math grades if fathers have little education).

Contrast Coefficients

Contrast	Six new cell codes					
	1.00	2.00	3.00	4.00	5.00	6.00
1	1	-1	0	0	0	0
2	0	0	1	-1	0	0
3	0	0	0	0	1	-1

Because Levene Test was significant we will use

Note, only the third contrast is significant.

		Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
math achievement test	Assume equal variances	1	-.6628	1.81759	-.365	67	.717
		2	-3.6135	2.75997	-1.309	67	.195
		3	-9.4699	2.54478	-3.721	67	.000
	Does not assume equal variances	1	-.6628	1.99426	-.332	24.512	.742
		2	-3.6135	2.12629	-1.699	13.819	.112
		3	-9.4699	2.38826	-3.965	13.858	.001

Difference between contrasted means.

The circle in the contrasts table above indicates the three simple main effects shown in Fig. 8.10.

Interpretation of Output 8.2

This output is the result of doing the first two steps, shown in Fig. 8.4, for interpreting a statistically significant interaction. Using Output 8.2, we can examine the simple effects statistically. The first table, **Descriptives**, provides the means of the *six new cell code* groups that will be compared, two at a time, as shown in the **Contrast Coefficients** table. The second table is the **Levene** test for the assumption that the variances are equal or homogeneous. In this case, the Levene's test is significant, so the assumption is violated and the variances cannot be assumed to be equal. The third table is the **ANOVA** table and the overall F (6.86) is significant ($p < .001$), which indicates that there are significant differences somewhere. The **Contrast Tests** table helps us identify which simple main effects were statistically significant. We will focus on one set of simple main effects in our interpretation, the ones based on t tests that do not assume equal variances. Note that we have circled three **Sigs.** (the significance level or p). These correspond to the three simple main effects shown with **arrows** in our drawing of the interaction plot (Fig. 8.10). For example, the left-hand arrow (and the first contrast) compares *math achievement* scores for both high and low *math grades* of students whose fathers have a relatively low education level.

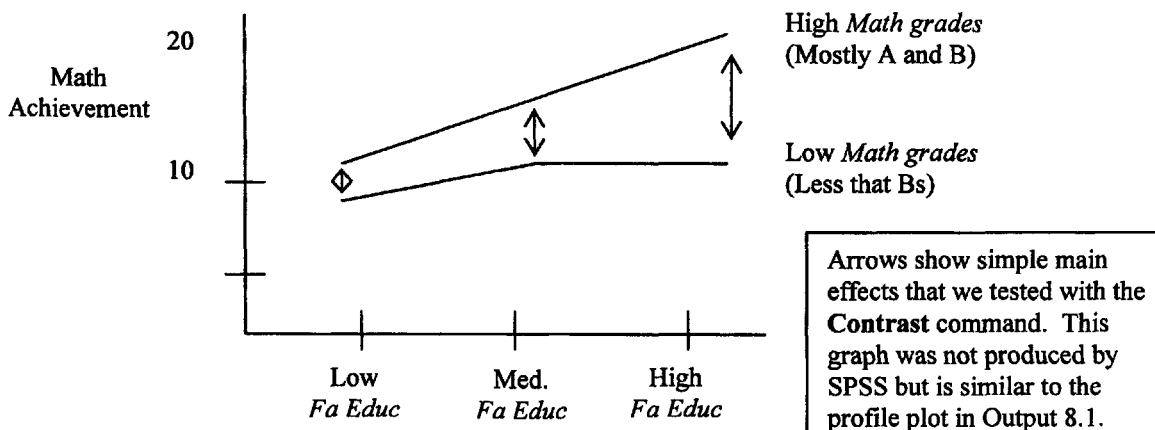


Fig. 8.10. Interaction plot showing three simple main effects.

We can now confirm statistically what we thought from visual inspection of the profile plot in Output 8.1. As you can see in Fig. 8.10 and the circled parts of the Contrast Test table, there is not a significant difference ($p = .742$) in *math achievement* between students with high and low *math grades* when their *father's education* is low. Likewise, this difference, while bigger (3.61 points), is not statistically significant ($p = .112$) when their *father's education* is medium (some college). However, when *father's education* is high, students with high (mostly As and Bs) *math grades* do much better (9.47 points) on the achievement test than those with low grades ($p = .001$). Thus, *math achievement* depends on both a student's *math grades* and their *father's education*. It would also be possible to examine the simple main effects for high and low math grades separately (the two lines), but this is usually not necessary to understand the interaction.

Example of How to Write About Problems 8.1 and 8.2

Results

Table 8.1 shows that there was a significant interaction between the effects of math grades and father's education on math achievement, $F(2,67)=3.97, p=.024$. Table 8.2 shows the number of subjects, the mean, and standard deviation of math achievement for each cell. Simple effects analyses revealed that, of students with more highly educated fathers, those who had mostly A and B grades had higher math achievement than did students who had lower grades, $t(13.86) = -3.97, p = .001$. Simple effects at the other levels of father education were not significant, indicating that for students whose fathers were less educated, students with higher and lower math grades had similar math achievement scores.

Table 8.1

Two-Way Analysis of Variance for Math Achievement as a Function of Math Grades and Father's Education

Variable and source	df	MS	F	η^2
Math Achievement				
Math Grades	1	325.78	10.86*	.139
Father's Education	2	323.00	10.77**	.243
Grades*Father's Educ	2	118.95	3.97*	.106
Error	67	29.99		

* $p < .05$, ** $p < .001$

Table 8.2

Means, Standard Deviations, and n for Math Achievement as a Function of Math Grades and Father's Education

Father's Education	Less A-B			Most A-B			Total	
	n	M	SD	n	M	SD	M	SD
HS grad or less	23	9.83	5.04	15	10.49	6.57	10.09	5.61
Some College	9	12.81	5.06	7	16.43	3.43	14.40	4.67
BS or More	11	12.36	7.18	8	21.83	2.85	16.35	7.41
Total	43	11.10	5.69	30	14.90	7.01	12.66	6.50

Problem 8.3: Analysis of Covariance (ANCOVA)

ANCOVA is an extension of ANOVA that typically provides a way of statistically controlling for the effects of continuous or scale variables that you are concerned about but that are not the focal point or independent variable(s) in the study. These scale variables are called **covariates** (or sometimes, control variables). Covariates usually are variables that may cause you to draw incorrect inferences about the ability of the independent variables to predict the dependent variable if not controlled (confounds). It is also possible to use ANCOVA when you are interested in examining a combination of a categorical (nominal) variable and a continuous (scale) variable as predictors of the dependent variable. In this latter case, you would not consider the covariate to be an extraneous variable, but rather a variable that is of interest in the study. SPSS will allow you to determine the significance of the contribution of the covariate, as well as whether the nominal variables (factors) significantly predict the dependent variable, over and above the “effect” of the covariate.

Assumptions of Analysis of Covariance (ANCOVA)

The first assumption for analysis of covariance (ANCOVA) is that the observations must be independent. This is a design issue that should be addressed prior to data collection. Using random sampling is the best way of ensuring that the observations are independent; however, this is not always possible. The most important thing to avoid is having known relationships among participants in the study (e.g., several family members or several participants obtained through “snowball” sampling included as “separate” participants). The second assumption is that the dependent variable needs to be normally distributed, which can be checked with skewness values. It is also important to have homogeneity of variances, particularly if sample sizes differ across levels of the independent variable(s). Homogeneity can be assessed through Box’s Test or Levene’s test. The fourth assumption is that there needs to be a linear relationship between the covariates and the dependent variable. This can be checked with a scatterplot (or matrix scatterplot if there is more than one covariate). The regression slopes for the covariates (in relation to the dependent variable) need to be the same for each group (this is called homogeneity of regression slopes). This assumption is one of the most important assumptions, and it can be checked with an *F* test on the interaction of the independent variables with the covariate. If the *F* test is significant, then this assumption has been violated.

In the HSB data, boys have significantly higher *math achievement* scores than girls. To see if the males’ higher math achievement scores are due to differences in the number of math courses taken by the male and female students, we will use *math courses taken* as a covariate and do ANCOVA.

8.3 Do boys have higher *math achievement* than girls if we control for differences in the number of *math courses taken*?

To answer this question, first, we need to assess the assumption of homogeneity of regression slopes:

- **Analyze => General Linear Model => Univariate.**
- Next, move *math achievement* to the **Dependent** box, *gender* to the **Fixed Factor** box, and *math courses taken* to the **Covariates** box. (See Fig. 8.11.)

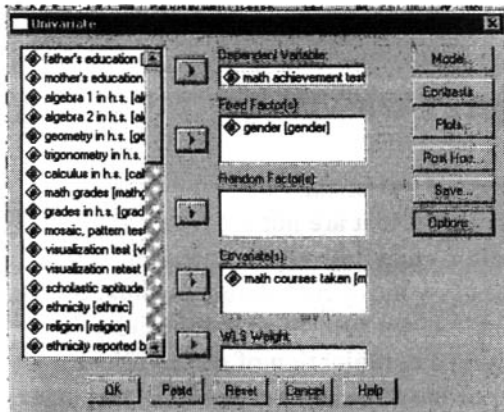


Fig. 8.11. GLM: Univariate.

- Click on **Model** and then click on the button next to **Custom** under **Specify Model**. (See Fig. 8.12.)

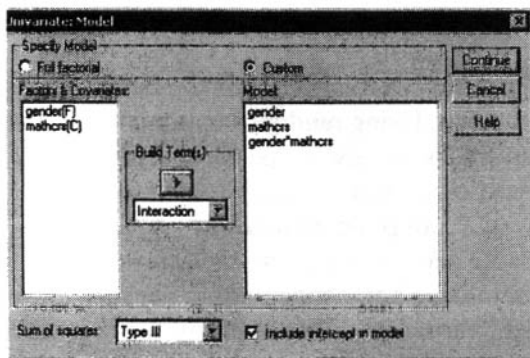


Fig. 8.12. Univariate: Model.

- Move *gender* from the **Factor & Covariates** box to the **Model** box. Do the same for *mathcrs*.
- Next highlight *gender* again, but hold down the “Shift” key and highlight *mathcrs*. This will allow you to have both *gender* and *mathcrs* highlighted at the same time. Click on the arrow to move both variables together to the **Model** box. This will make *gender*mathcrs*.
- Click on **Continue** and then **OK**. Your syntax and output should look like the beginning of Output 8.3.

Next, do the following:

- Analyze => General Linear Model => Univariate.**
- Click on **Reset**.
- Next, move *math achievement* to the **Dependent** box, *gender* to the **Fixed Factor** box, and *math courses taken* to the **Covariates** box. (See Fig. 8.11.)
- Click on **Options** to get Fig. 8.13.

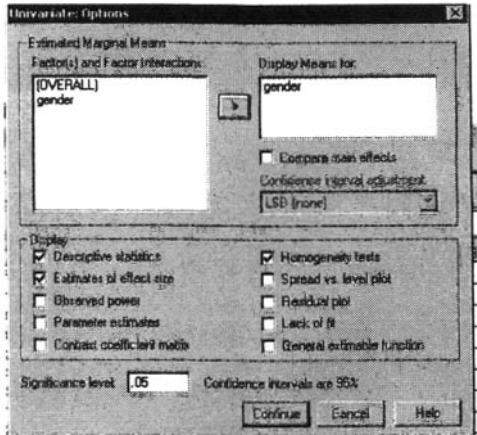


Fig. 8.13. GLM: Univariate options.

- Select **Descriptive statistics**, **Estimates of effect size**, and **Homogeneity tests**. (See Fig. 8.13.)
- Move *gender* into the box labeled **Display Means for**.
- Click on **Continue** and then **OK**. Your syntax and output should look like the rest of Output 8.3.

Output 8.3: Analysis of Covariance (ANCOVA)

```
UNIANOVA
  mathach BY gender WITH mathcrs
  /METHOD = SSTYPE(3)
  /INTERCEPT = INCLUDE
  /CRITERIA = ALPHA(.05)
  /DESIGN = gender mathcrs gender*mathcrs .
```

Tests of Between-Subjects Effects

Dependent Variable: math achievement test

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	2085.698 ^a	3	695.233	40.903	.000
Intercept	833.723	1	833.723	49.051	.000
GENDER	8.941E-05	1	8.941E-05	.000	.998
MATHCRS	1775.937	1	1775.937	104.486	.000
GENDER * MATHCRS	3.369	1	3.369	.198	.658
Error	1206.783	71	16.997		
Total	15132.393	75			
Corrected Total	3292.481	74			

a. R Squared = .633 (Adjusted R Squared = .618)

Recall that this analysis was done to check the assumption of homogeneity of regression slopes, not to test the main hypothesis. The factor and covariate do not interact, so the assumption of homogeneity of regression slopes has been met.

```
UNIANOVA
  mathach BY gender WITH mathcrs
  /METHOD = SSTYPE(3)
  /INTERCEPT = INCLUDE
  /EMMEANS = TABLES(gender) WITH(mathcrs=MEAN)
  /PRINT = DESCRIPTIVE ETASQ HOMOGENEITY
  /CRITERIA = ALPHA(.05)
  /DESIGN = mathcrs gender .
```

Univariate Analysis of Variance

Between-Subjects Factors

		Value Label	N
gender	0	male	34
	1	female	41

Descriptive Statistics

Dependent Variable: math achievement test

gender	Mean	Std. Deviation	N
male	14.7550	6.03154	34
female	10.7479	6.69612	41
Total	12.5645	6.67031	75

Note that the mean score of males was four points higher than females on math before the ANCOVA.

Levene's Test of Equality of Error Variances^a

Dependent Variable: math achievement test

F	df1	df2	Sig.
5.572	1	73	.021

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

a. Design: Intercept+MATHCRS+GENDER

This is significant; therefore, it indicates that the assumption of homogeneity of variances has been violated. However, since cell sizes are almost equal (34 and 41), this is not a big problem, especially given the way SPSS calculates the ANCOVA.

Tests of Between-Subjects Effects

Dependent Variable: math achievement test

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	2082.329 ^a	2	1041.164	61.946	.000	.632
Intercept	946.381	1	946.381	56.306	.000	.439
MATHCRS	1783.894	1	1783.894	106.136	.000	.596
GENDER	6.001	1	6.001	.357	.552	.005
Error	1210.152	72	16.808			
Total	15132.393	75				
Corrected Total	3292.481	74				

a. R Squared = .632 (Adjusted R Squared = .622)

The covariate (MATHCRS) is significant but the gender difference is not significant.

Estimated Marginal Means

gender

Dependent Variable: math achievement test

gender	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
male	12.893 ^a	.726	11.446	14.340
female	12.292 ^a	.658	10.981	13.603

a. Covariates appearing in the model are evaluated at the following values: math courses taken = 2.11.

Note that these are similar after differences in *math courses taken* were controlled.

Interpretation of Output 8.3

The ANCOVA (**Tests of Between Subject Effects**) table is interpreted in much the same way as ANOVA tables in earlier outputs. The covariate (*mathcrs*) has a highly significant “effect” on math achievement, as should be the case. However, the “effect” of gender is no longer significant, $F(1,72) = .36, p = .55$. You can see from the **Estimated Marginal Means** table that the statistically adjusted math achievement means for boys and girls are quite similar once differences in the number of *math courses taken* were accounted for.

Example of How to Write About Problems 8.3

Results

An analysis of covariance was used to assess whether boys have higher math achievement than girls after controlling for differences between boys and girls in the number of math courses taken. Results indicate that after controlling for the number of math courses taken, there is not a significant difference between boys and girls in math achievement, $F(1, 72) = .36, p = .552$. Table 8.3 presents the means and standard deviations for boys and girls on math achievement, before and after controlling for number of math courses taken. As is evident from this table, virtually no difference between boys and girls remains after differences in number of math courses taken are controlled.

Table 8.3

Adjusted and Unadjusted Gender Means and Variability for Math Achievement Using Math Grades as a Covariate

	<i>N</i>	<i>Unadjusted</i>		<i>Adjusted</i>	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SE</i>
Males	34	14.76	6.03	12.89	.73
Females	41	10.75	6.70	12.29	.66

Table 8.4

Analysis of Covariance for Math Achievement as a Function of Gender, Using Number of Math Courses Taken as a Covariate

Source	<i>Df</i>	<i>Ms</i>	<i>F</i>	<i>P</i>	η^2
Math Courses	1	1783.89	106.14	<.001	.60
Gender	1	6.00	.36	.552	.01
Error	72	16.81			

Interpretation Questions

- 8.1 In Output 8.1: a) Is the interaction significant? b) Examine the profile plot of the cell means that illustrates the interaction. Describe it in words. c) Is the main effect of *father's education* significant? Interpret the eta squared. d) Is the "effect" of *math grades* significant? e) Why did we put the word effect in quotes? f) How might focusing on the main effects be misleading?
- 8.2 Interpret in words, the contrasts shown in Output 8.2.
- 8.3 In Output 8.3: a) Are the adjusted main effects of *gender* significant? b) What are the adjusted *math achievement* means (marginal means) for males and females? c) Is the effect of the covariate (*mathcrs*) significant? d) What do a) and c) tell us about *gender* differences in *math achievement* scores?

Extra SPSS Problems

- 8.1. Using the College Student data file, do *gender* and *marital status* seem to have an effect on *student's height* and do *gender* and *marital status* interact? Run the appropriate SPSS analysis and interpret the results.
- 8.2. A county is interested in assessing the costs of their road construction (*cost*); whether it is more expensive in the South Florida District or in other districts in the county (*district*) in relation to whether the contractors were found through competitive contracts or fixed contracts (*status*). Using the Road construction bids.sav data set, conduct factorial ANOVA.
- Are these factors independently significant in the level of *cost*?
 - Is there a significant interaction effect between the *district* and the *status* on the *cost*?
 - How much total variance in *cost* can be attributed to these variables?
 - If a post-hoc test is necessary, conduct an appropriate test and discuss the results.
- 8.3. Answer the following question using the Mall rentals.sav data file. Do people with higher levels of education (*edu*) have a shorter time for getting a mall rental (*interval*) than people with lower levels of education if we control for differences in the number of *status*?

CHAPTER 9

Repeated Measures and Mixed ANOVAs

In this chapter, you will analyze a new data set that includes repeated measure data. These data allow you to compare four products (or these could be four instructional programs), each of which was evaluated by 12 consumers/judges (6 male and 6 female). The analysis requires statistical techniques for **within subjects** and **mixed designs**.

In Problem 9.1, to do the analysis, you will use the SPSS **General Linear Model** program (called GLM) and do a **repeated measures ANOVA**. In Problem 9.3, you will use the same GLM program to do a **mixed ANOVA**, one that has a repeated measures independent variable *and* a between groups independent variable. In Problem 9.2, you will use a nonparametric statistic, the **Friedman test**, which is similar to the repeated measures ANOVA. SPSS does not have a nonparametric equivalent to the mixed ANOVA.

Chapter 3 provides several tables to help you decide what statistic to use with various types of difference statistics problems. Tables 3.1 and 3.3 include the statistics used in this chapter. Please refer back to chapter 3 to see how these statistics fit into the big picture.

Assumptions of Repeated Measures ANOVA

The assumptions of repeated measures ANOVA are similar to those for the ANOVA, and they include independence of observations (unless the dependent data comprise the “within-subjects” or “repeated measures” factor), normality, and homogeneity of variances. However, in addition to variances, which involve deviations from the mean of each person’s score on one measure, the repeated measures design includes more than one measure for each person. Thus, covariances, which involve deviations from the mean of each of two measures for each person, also exist, and these covariances need to meet certain assumptions as well. The homogeneity assumption for repeated measures designs, known as sphericity, requires equal variances and covariances for each level of the within subjects variable. Another way of thinking about sphericity is that, if one created new variables for each pair of within subjects variable levels by subtracting each person’s score for one level of the repeated measures variable from that same person’s score for the other level of the within subject variable, the variances for all of these new difference scores would be equal. Unfortunately, it is rare for behavioral science data to meet the sphericity assumption, and violations of this assumption can seriously affect results. However, fortunately, there are good ways of dealing with this problem—either by adjusting the degrees of freedom or by using a multivariate approach to repeated measures. Both of these are discussed later in this chapter. One can test for the sphericity assumption using the Mauchly’s test, the Box test, the Greenhouse-Geisser test, and/or the Huynh-Feldt tests (see below). Even though the repeated measures ANOVA is fairly robust to violations of normality, the dependent variable should be approximately normally distributed for each level of the independent variable.

Assumptions of Mixed ANOVA

The assumptions for mixed ANOVA are similar to those for the repeated measures ANOVA, except that the assumption of sphericity must hold for levels of the within subjects variable at each level of between subjects variables. This can be tested using SPSS with Box’s *M*.

The Product Data Set

Open the SPSS for Windows program. Open the **Product** data set. Do not retrieve the **hsbdata** for this assignment.

In this study, each of the 12 participants (or subjects) has evaluated four products (e.g., four brands of DVD players) on 1-7 Likert scales. You will find the data presented in the SPSS data editor once you have opened the product data set.

- Click on the **Data View** tab at the bottom of the screen to get Fig 9.1.

	p1	p2	p3	p4	gender
1	7	7	6	6	1
2	7	6	6	5	1
3	6	5	5	4	1
4	6	4	4	3	1
5	5	3	3	2	1
6	4	2	2	1	1
7	6	5	5	4	2
8	5	4	5	4	2
9	4	3	4	3	2
10	3	2	3	2	2
11	2	1	2	1	2
12	1	1	1	1	2

Note that this is what the data view for your product data looks like. In most studies the *N* would be larger.

Fig. 9.1. Data view for the product data.

Figure 9.1 shows the **Data View** for 12 participants who were asked to rate four *products* (p1, p2, p3, and p4) from 1 (very low quality) to 7 (very high quality). The participants were also asked their *gender* (1=male, 2=female). Thus, subjects 1-6 are males and 7-12 are females.

- Click on the **Variable View** tab to see the names and labels of the variables as shown in Fig. 9.2.

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	p1	Numeric	1	0	Product 1	{1, Very Low}..	None	8	Right	Scale
2	p2	Numeric	1	0	Product 2	{1, Very Low}..	None	8	Right	Scale
3	p3	Numeric	1	0	Product 3	{1, Very Low}..	None	8	Right	Scale
4	p4	Numeric	1	0	Product 4	{1, Very Low}..	None	8	Right	Scale
5	gender	Numeric	8	0	Gender	{1, Male}..	None	8	Right	Nominal

Fig. 9.2. Variable view.

We have labeled the measurement of the four *products* scale because the frequency distribution of each is approximately normal. We label *gender* and other dichotomous variables as nominal; however, despite this traditional designation for dichotomous variables, dichotomous variables, unlike other types of nominal variables, provide meaningful averages (indicating percentage of participants falling in each category) and can be used in multiple and logistic regression as if they were ordered. Furthermore, many dichotomous variables (but not *gender*) even have an implied order (e.g., 0=do not have the characteristic and 1= have the characteristic; thus, a score of 1 indicates more of the characteristic than does a score of 0).

In repeated measures (also called within-subjects) analyses, SPSS creates the **Within-Subjects Factor** or independent variable from two or more existing variables (in this case *p1*, *p2*, *p3*, and *p4*). These then become levels of the new independent variable. In this example, we will call the new variable *product*, and it has four levels (*p1*, etc.), indicating which product was being rated. In order for a set of variables

to be converted into a meaningful within-subject factor, the scores on each of the existing variables (which will become levels of the new within-subjects variable) have to be comparable (e.g., ratings on the same 7-point Likert scale) and each participant has to have a score on each of the variables. The within-subject factor could be based on related or matched subjects (e.g., the ratings of a product by mother, father, and child from each family) instead of a single participant having repeated scores. The within-subjects design should be used whenever there are known dependencies in the data, such as multiple family members, that are consistent across the independent participants, such as families (i.e., there is a mother, father, and child rating for each family) but would otherwise violate the between-subjects assumption of independent observations. The **dependent variable** for the data in Fig. 9.1 could be called *product ratings* and would be the scores/ratings for each of the four products. Thus, the independent variable, *product*, indicates which product is being rated, and the dependent variable is the rating itself. Note that *gender* is a between subjects independent variable that will be used in Problem 9.3.

Problem 9.1: Repeated Measures ANOVA

The GLM repeated measures procedure provides a variety of analysis of variance procedures to use when the same measurement is made several times on each subject or the same measurement is made on several related subjects. The single-factor repeated measures ANOVA, which we will use for Problem 9.1, is appropriate when you have one independent variable with two or more levels that are repeated measures and one dependent variable. If there are only two levels, the sphericity assumption is not a problem because there is only one pair of levels. If between subjects factors are specified, they divide the sample into groups. There are no between subject or group factors in this problem. Finally, you can use a multivariate or univariate approach to testing repeated measures effects.

9.1. Are there differences between the average ratings for the four *products*?

Let's test whether there are differences between the average ratings of the four *products*. We are assuming the *product* ratings are scale/normal data. Follow these commands:

- **Analyze => General Linear Model => Repeated Measures.** (See Fig. 9.3.)
- **Delete** the factor 1 from the **Within-Subject Factor Name** box and **replace** it with the name **product**, our name for the repeated measures independent variable that SPSS will generate from the four products.
- Type **4** in the **Number of Levels** box since there are four products established in the data file.
- Click on **Add**, then click on **Define**, which changes the screen to a new menu box. (See Fig. 9.4.)

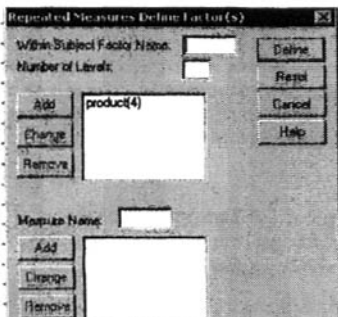


Fig. 9.3. Repeated measures GLM define factor(s).

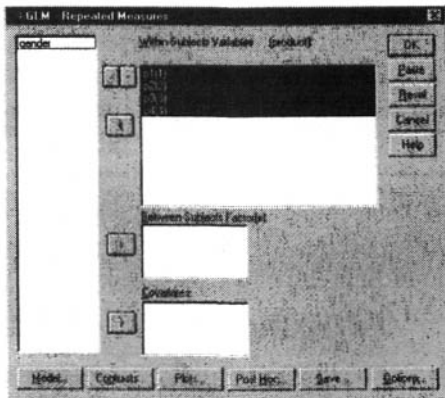


Fig. 9.4. GLM Repeated measures.

- Now move *p1*, *p2*, *p3*, and *p4* over to the **Within-Subjects Variables** box.
- Click on **Contrasts**, be sure **Polynomial** is in the parenthesis after *product*. (See Figure 9.5.) SPSS does not provide post hoc tests for the within subjects (repeated measures) effects, so we will use contrasts. If the products are ordered, let's say, by price, we can use the polynomial contrasts that are interpreted below. If we wanted to use a different type of contrast, we could change the type by clicking on the arrow under **Change Constant**.
- Click on **Continue**.

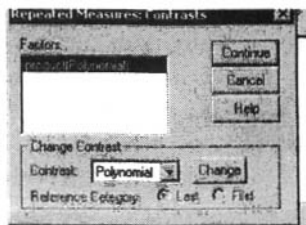


Fig 9.5. Repeated measures: Contrasts

- Click on **Options**. (See Fig. 9.6.)
- Click on **Estimates of effect size**.

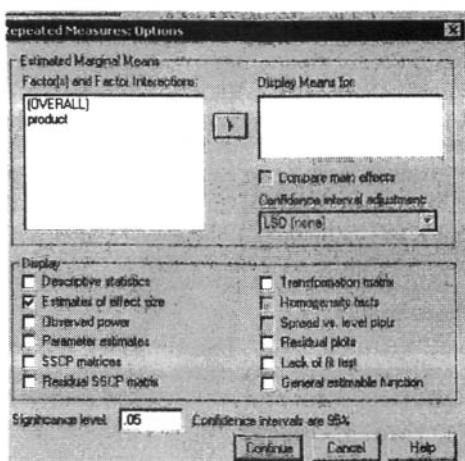


Fig 9.6. Repeated measures: Options.

- Click on **Continue**, then on **OK**.
- Compare your syntax and output to Output 9.1.

Output 9.1: Repeated Measures ANOVA Using the General Linear Model Program

```
GLM
  p1 p2 p3 p4
  /WSFACTOR = product 4 Polynomial
  /METHOD = SSTYPE(3)
  /PRINT = ETASQ
  /CRITERIA = ALPHA(.05)
  /WSDESIGN = product .
```

General Linear Model

Within-Subjects Factors

Measure: MEASURE_1

PRODUCT	Dependent Variable
1	P1
2	P2
3	P3
4	P4

This shows four similar multivariate tests of the within subjects effect. These are actually a form of MANOVA, which is discussed in the next chapter. In this case, all four tests have the same *F*s, and are significant. If the sphericity assumption is violated, a multivariate test could be used or one could use the procedure shown below, which corrects the degrees of freedom.

Multivariate Tests^b

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
product	Pillai's Trace	.864	19.065 ^a	3.000	9.000	.000	.864
	Wilks' Lambda	.136	19.065 ^a	3.000	9.000	.000	.864
	Hotelling's Trace	6.355	19.065 ^a	3.000	9.000	.000	.864
	Roy's Largest Root	6.355	19.065 ^a	3.000	9.000	.000	.864

a. Exact statistic

b.

Design: Intercept

Within Subjects Design: product

This is a test of an assumption of the univariate approach to repeated measures ANOVA. As is commonly the case, the Mauchly statistic is significant and, thus, the assumption is violated.

The epsilons, which are measures of degree of sphericity, are less than 1.0, indicating that the sphericity assumption is violated. The "lower-bound" indicates the lowest value that epsilon could be. The highest epsilon possible is always 1.0. When sphericity is violated, you can either use the multivariate results or use epsilons to adjust the numerator and denominator *dfs*. Typically, when epsilons are less than .75, use the Greenhouse-Geisser epsilon, but use Huynh-Feldt if epsilon > .75.

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
PRODUCT	.101	22.253	5	.001	.544	.626	.333

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables 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: PRODUCT

SPSS for Intermediate Statistics

Note that 3 and 33 would be the *dfs* to use if sphericity were not violated. Because it is, we will use the Greenhouse-Geisser correction, which multiplies 3 and 33 by epsilon, which in this case is .544, yielding *dfs* of 1.63 and 17.95.

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
product	Sphericity Assumed	17.229	3	5.743	23.629	.000	.682
	Greenhouse-Geisser	17.229	1.632	10.556	23.629	.000	.682
	Huynh-Feldt	17.229	1.877	9.178	23.629	.000	.682
	Lower-bound	17.229	1.000	17.229	23.629	.001	.682
Error(product)	Sphericity Assumed	8.021	33	.243			
	Greenhouse-Geisser	8.021	17.953	.447			
	Huynh-Feldt	8.021	20.649	.388			
	Lower-bound	8.021	11.000	.729			

Significance of linear, quadratic, and cubic trends.

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	product	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
product	Linear	13.537	1	13.537	26.532	.000	.707
	Quadratic	.188	1	.188	3.667	.082	.250
	Cubic	3.504	1	3.504	20.883	.001	.655
Error(product)	Linear	5.613	11	.510			
	Quadratic	.563	11	.051			
	Cubic	1.846	11	.168			

Tests of Between-Subjects Effects

Ignore this. There were no between groups/subjects variables in this problem.

Measure: MEASURE_1

Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	682.521	1	682.521	56.352	.000	.837
Error	133.229	11	12.112			

Interpretation of Output 9.1

The first table identifies the four levels of the within subjects repeated measures independent variable, *product*. For each level (*P1* to *P4*), there is a rating of 1-7, which is the dependent variable.

The second table presents four similar **Multivariate Tests** of the within-subjects effect (i.e., whether the four *products* are rated equally). **Wilks' Lambda** is a commonly used multivariate test. Notice that in this case, the *F*s, *df*, and significance are all the same: $F(3, 9) = 19.07, p < .001$. The significant *F* means that there is a difference somewhere in how the products are rated. The multivariate tests can be used whether or not sphericity is violated. However, if *epsilon*s are high, indicating that one is close to achieving sphericity, the multivariate tests may be less powerful (less likely to indicate statistical significance) than the corrected univariate repeated measures ANOVA.

The third table shows that the **Mauchly Test of Sphericity** is significant, which indicates that these data violate the sphericity assumption of the univariate approach to repeated measures analysis of variance. Thus, we should either use the multivariate approach, use the appropriate nonparametric test (Friedman), or correct the univariate approach with the **Greenhouse-Geisser** or other similar correction.

You can see in the **Tests of Within-Subject Effects** that these corrections reduce the degrees of freedom by multiplying them by **Epsilon**. In this case, $3 \times .544 = 1.63$ and $33 \times .544 = 17.95$. Even with this adjustment, the **Within-Subjects Effects** (of *product*) is significant, $F(1.63, 17.95) = 23.63, p < .001$, as were the multivariate tests. This means that the ratings of the four products are significantly different. However, this overall (*product*) *F* does not tell you which pairs of products have significantly different means.

SPSS has several tests of within subject contrasts. We have chosen to use the polynomial contrast on the assumption that the products are ordered, say from the most expensive as *P1* to the least as *P4*. The **Tests of Within-Subjects Contrasts** table shows whether the four product means are significantly like a straight line (linear effect), a line with one change in direction (quadratic), and a two bend line (cubic). You can see that there is a highly significant linear trend and a significant cubic trend. If you plot the means (shown in Output 9.2), these trends should be clear. Overall, there is a linear decline in ratings from *Product 1* (3.67) to *Product 4* (1.33). However, *Product 2* has a somewhat lower mean (2.25) than *Product 3* (2.75) producing the cubic trend.

In Output 9.1, we ignore the “**Tests of Between Subject Effects**” table because we do not have a between subject/groups variable.

How to Write about Output 9.1.

Results

A repeated measures ANOVA, with Greenhouse-Geisser correction, was conducted to assess whether there were differences between the average ratings of the four products. Results indicated that participants did rate the four products differently, $F(1.63, 17.95) = 23.63, p < .001, R^2 = .68, \eta^2 = .68$. The means and standard deviations for the products listed in order from most expensive to least expensive, are presented in Table 9.1. Examination of these means suggests that participants rated products that are more expensive more highly than less expensive products. Polynomial contrasts indicated, in support of this, there was a significant linear trend, $F(1,11) = 26.53, p < .001, \eta^2 = .71$. However, this finding was qualified by the significant cubic trend, $F(1,11) = 20.88, p = .001, \eta^2 = .66$, reflecting the higher rating for Product 3 than Product 2.

Table 9.1.

Means and Standard Deviations of the Four Product Ratings

Variable	<i>M</i>	<i>SD</i>
Product 1	4.67	1.92
Product 2	3.58	1.93
Product 3	3.83	1.64
Product 4	3.00	1.65

Problem 9.2: The Friedman Nonparametric Test for Several Related Samples

What could you do if the *product* ratings are ordinal data or the repeated measures ANOVA assumptions are markedly violated? One answer is to use a nonparametric statistic. As you can tell from Table 3.1 (p. 49), an appropriate nonparametric test for when you have more than two levels of one repeated measures or related samples (i.e., within-subjects) independent variables is the Friedman test.

9.2. Are there differences among the mean ranks of the *product* ratings?

Let's use *p1* to *p4* again with the following commands:

- **Analyze => Nonparametric tests => K Related Samples...** and move *product1[p1]* to *product4[p4]* to the **Test Variables** box (see Fig. 9.7).
- Make sure the **Friedman** test type is checked.
- Then click on **Statistics** to get Fig. 9.8.

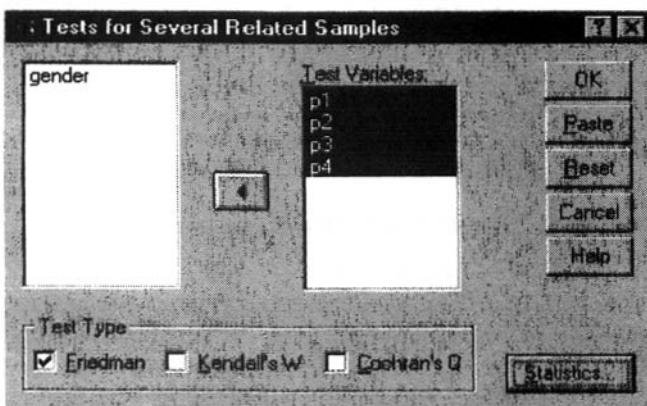


Fig. 9.7. Tests for several related samples.

- Now click on **Descriptive** statistics.
- Click on **Continue**, then **OK**. Look at your output and compare it to Output 9.2a.

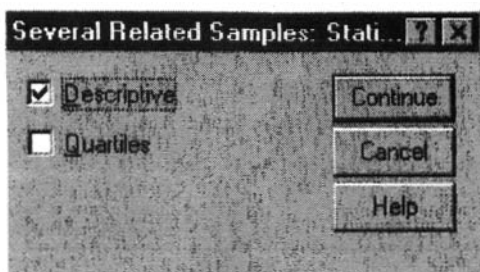


Fig. 9.8. Descriptive statistics for nonparametric tests for several related samples.

Output 9.2a: Nonparametric Tests With Four Related Samples

NPART TESTS

/FRIEDMAN = p1 p2 p3 p4
/STATISTICS DESCRIPTIVES
/MISSING LISTWISE.

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
Product 1	12	4.67	1.92	1	7
Product 2	12	3.58	1.93	1	7
Product 3	12	3.83	1.64	1	6
Product 4	12	3.00	1.65	1	6

Friedman Test

Ranks

	Mean Rank
Product 1	3.67
Product 2	2.25
Product 3	2.75
Product 4	1.33

Test Statistics^a

N	12
Chi-Square	26.170
df	3
Asymp. Sig.	.000

The overall difference between the mean ranks is significant at $p < .001$.

a. Friedman Test

Interpretation of Output 9.2a

The **Descriptive Statistics** table provides familiar statistics that were useful for interpreting the polynomial contrasts in Output 9.1 but are not, themselves used in the Friedman test. Rather, ratings are *ranked*, and the means of these ranks are used in the Friedman. The **Ranks** table shows the mean rank for each of the four *products*.

The **Friedman Test Statistic** table tests the null hypothesis that the four related variables come from the same population. For each rater/case, the four variables are ranked from 1 to 4. The test statistic is based on these ranks. The **Asymp. Sig.** (asymptotic significance, which means this is not an exact significance level) of $p < .001$ indicates that there is a significant overall difference among the four mean ranks.

Unfortunately, SPSS does not have a post hoc test built into the Friedman test program. Therefore, in order to determine which differences between mean ranks are significant, and thus the likely source of the significant Friedman test, we will need to perform a nonparametric related sample test like the Wilcoxon. See Table 3.1 (p. 49) for a more complete view of all of the different statistical tests and their uses.

- **Analyze => Nonparametric tests => 2 Related Samples** and make sure that **Wilcoxon** is checked. Then highlight *product1[1]* and *product2[2]* and click the arrow to move them over. Next, repeat for *product2[2]* and *product3[3]* and for *product3[3]* and *product4[4]*. Check to make sure your syntax and output are like those in Output. 9.2b.

Output 9.2b: Follow-up Paired Comparisons for Significant Friedman

NPAR TEST

/WILCOXON=p1 p2 p3 WITH p2 p3 p4 (PAIRED)
/MISSING ANALYSIS.

Ranks

		N	Mean Rank	Sum of Ranks
Product 2 - Product 1	Negative Ranks	10 ^a	5.50	55.00
	Positive Ranks	0 ^b	.00	.00
	Ties	2 ^c		
	Total	12		
Product 3 - Product 2	Negative Ranks	1 ^d	3.00	3.00
	Positive Ranks	4 ^e	3.00	12.00
	Ties	7 ^f		
	Total	12		
Product 4 - Product 3	Negative Ranks	10 ^g	5.50	55.00
	Positive Ranks	0 ^h	.00	.00
	Ties	2 ⁱ		
	Total	12		

- a. Product 2 < Product 1
- b. Product 2 > Product 1
- c. Product 1 = Product 2
- d. Product 3 < Product 2
- e. Product 3 > Product 2
- f. Product 2 = Product 3
- g. Product 4 < Product 3
- h. Product 4 > Product 3
- i. Product 3 = Product 4

Test Statistics^c

	Product 2 - Product 1	Product 3 - Product 2	Product 4 - Product 3
Z	-2.919 ^a	-1.342 ^b	-3.162 ^a
Asymp. Sig. (2-tailed)	.004	.180	.002

- a. Based on positive ranks.
- b. Based on negative ranks.
- c. Wilcoxon Signed Ranks Test

These two contrasts are statistically significant.

Interpretation of Output 9.2b

Given that there was a significant overall difference between the mean ranks, we followed up the Friedman with Wilcoxon tests. Since the products are ordered and there are four levels, it makes sense to do three orthogonal contrasts, contrasting adjacent pairs. A larger number of comparisons would prevent the comparisons from being completely independent of one another. Given that three post hoc comparisons were made, it would be desirable to make a Bonferroni correction on alpha, such that p would need to be $.05/3$ (about .017) to be significant. Notice that the contrasts between *products 2 and 1* and between *4 and 3* are significant at this level and that they both indicate that the higher numbered product was given a lower rating than the lower numbered product. On the other hand, the difference between ratings for *products 3 and 2* was not significant. This is not surprising when you look at the means in Output 9.2 and suggests that the slight increase in ratings for *product 3* compared to *product 2* may not be important. Remember, however, that the Wilcoxon is not performed on the rating scores themselves, but rather on the ranks of the ratings.

How to Write About Outputs 9.2a and 9.2b

Results

A Friedman test was conducted to assess if there were differences among the mean ranks of the product ratings, $\chi^2(3, N = 12) = 26.17, p = .001$. This indicates that there were differences among the four mean ranks. Three orthogonal contrasts were performed using Wilcoxon tests with the Bonferroni correction (*comparison-wise alpha* = .017). The contrasts between Product 1 and 2, and between Products 3 and 4 were found to be significant; however that between Products 2 and 3 was not significant. In both cases, the significant contrasts indicated that the more expensive product was rated more highly.

Problem 9.3: Mixed ANOVA

Using this General Linear Model procedure, you can test null hypotheses about the effects of both the between groups factors and the within subjects factors. You can investigate interactions between factors as well as the effects of individual factors on one or more dependent variables.

Repeat Problem 9.1 except add *gender* to see if there are any *gender* differences as well as *product* differences and if there is an interaction between *gender* and *product*. Is *gender* a between groups/subjects or within subjects variable? The answer is important in how you compute and interpret the analysis.

9.3. Are there *gender* as well as *product* differences? Is there an interaction between *gender* and *product*?

- Click on **Analyze => General Linear Model => Repeated Measures...**
- In the **Repeated Measures Define Factor(s)** window (Fig. 9.3), you should see, *product* (4) in the big box. If so, click on **Define** (if not repeat the steps for Problem 9.1).
- Then move *gender* to the **Between Subjects Factor(s)** box (Fig. 9.4).
- Click on **Contrasts**.
- Click on *product*(polynomial) under **Factors**.
- Be sure that **None** is listed under **Contrasts**, then click **Change**. This will make it say *product*(none).
- Click **Continue**.
- Click on **OK**.

Compare your syntax and output with Output 9.3.

Output 9.3: Mixed ANOVA: Product by Gender

GLM

```

p1 p2 p3 p4 BY gender
/WSFACTOR = product 4
/METHOD = SSTYPE(3)
/PRINT = ETASQ
/CRITERIA = ALPHA(.05)
/WSDESIGN = product
/DESIGN = gender .

```

General Linear Model**Within-Subjects Factors**

Measure: MEASURE_1

PRODUCT	Dependent Variable
1	P1
2	P2
3	P3
4	P4

Between-Subjects Factors

	Value Label	N
Gender 1	Male	6
2	Female	6

The multivariate test for product and the interaction are significant.

Multivariate Tests^b

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
product	Pillai's Trace	.897	23.152 ^a	3.000	8.000	.000	.897
	Wilks' Lambda	.103	23.152 ^a	3.000	8.000	.000	.897
	Hotelling's Trace	8.682	23.152 ^a	3.000	8.000	.000	.897
	Roy's Largest Root	8.682	23.152 ^a	3.000	8.000	.000	.897
product * gender	Pillai's Trace	.763	8.606 ^a	3.000	8.000	.007	.763
	Wilks' Lambda	.237	8.606 ^a	3.000	8.000	.007	.763
	Hotelling's Trace	3.227	8.606 ^a	3.000	8.000	.007	.763
	Roy's Largest Root	3.227	8.606 ^a	3.000	8.000	.007	.763

a. Exact statistic

b.

Design: Intercept+gender

Within Subjects Design: product

Chapter 9 – Repeated Measures and Mixed ANOVAs

Mauchly's Test of Sphericity^a

The assumption is violated because W is significant.

Measure: MEASURE_1

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon ^a		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
product	.176	15.138	5	.010	.547	.705	.333

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

- 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.
- Design: Intercept+gender
Within Subjects Design: product

Univariate test adjusted with Greenhouse-Geisser epsilons because of lack of sphericity.

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
product	Sphericity Assumed	17.229	3	5.743	41.768	.000	.807
	Greenhouse-Geisser	17.229	1.640	10.507	41.768	.000	.807
	Huynh-Feldt	17.229	2.114	8.148	41.768	.000	.807
	Lower-bound	17.229	1.000	17.229	41.768	.000	.807
product * gender	Sphericity Assumed	3.896	3	1.299	9.444	.000	.486
	Greenhouse-Geisser	3.896	1.640	2.376	9.444	.003	.486
	Huynh-Feldt	3.896	2.114	1.842	9.444	.001	.486
	Lower-bound	3.896	1.000	3.896	9.444	.012	.486
Error(product)	Sphericity Assumed	4.125	30	.138			
	Greenhouse-Geisser	4.125	16.398	.252			
	Huynh-Feldt	4.125	21.145	.195			
	Lower-bound	4.125	10.000	.413			

Univariate within subjects and interaction trends.

Tests of Within-Subjects Contrasts

Measure: MEASURE_1

Source	product	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
product	Linear	13.537	1	13.537	64.209	.000	.865
	Quadratic	.188	1	.188	5.000	.049	.333
	Cubic	3.504	1	3.504	21.345	.001	.684
product * gender	Linear	3.504	1	3.504	16.621	.002	.624
	Quadratic	.188	1	.188	5.000	.049	.333
	Cubic	.204	1	.204	1.244	.291	.111
Error(product)	Linear	2.108	10	.211			
	Quadratic	.375	10	.038			
	Cubic	1.642	10	.164			

Tests of Between-Subjects Effects

Measure: MEASURE_1

Transformed Variable: Average

The test of whether males and females rate the products differently is not significant.

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	682.521	1	682.521	65.183	.000	.867
gender	28.521	1	28.521	2.724	.130	.214
Error	104.708	10	10.471			

Interpretation of Output 9.3

Most of these tables are similar in format and interpretation to those in Output 9.1. However, the addition of *gender* as a between subjects independent variable makes the last table meaningful and adds an interaction (*product * gender*) to both the **multivariate tests** table and the univariate (**within-subjects effects**) tables. Both multivariate and univariate tables indicate that (as in Output 9.1) there are differences among the *products*. Again, there are significant linear and cubic trends, but now the quadratic trend is significant as well. In addition, the interaction of *product* and *gender* are significant with both univariate and multivariate tests. This means that the downward overall trend for all subjects is somewhat different for males and females. Moreover, the linear and quadratic trends are significant for the interaction between *product* and *gender*.

If we had three groups, instead of just males and females, for our between groups variable and if the ANOVA had been significant, we would have used a post hoc test for this effect. SPSS provides the same wide variety of post hoc multiple comparisons that were available for one-way ANOVA and factorial ANOVAs, including the Tukey and Games Howell.

How to Write About Output 9.3**Results**

A mixed ANOVA was conducted to assess whether there were gender and product differences in product ratings. Results indicated a significant main effect of product, $F(1.64, 16.4) = 41.77, p < .001$, but not of gender, $F(1, 10) = 2.72, p = .13$. However, the product main effect was qualified by a significant interaction between product and gender, $F(1.64, 16.40) = 9.44, p = .003$. This indicates that even though as a general rule, males do not rate products higher or lower than females, males do rate particular products differently than do females.

Interpretation Questions

- 9.1. In Output 9.2: a) Explain the results in nontechnical terms.
- 9.2. In Output 9.2b: a) Why did we do Wilcoxon's after the Friedman test? Why Wilcoxon's rather than paired *t* tests or Mann-Whitney U tests? b) Explain the results of the Wilcoxon's in nontechnical terms.
- 9.3. In Output 9.3: a) Is the Mauchly sphericity test significant? Does this mean that the assumption is or is not violated? If it is violated, what can you do? b) How would you interpret the *F* for *product* (within subjects)? c) Is the interaction between *product* and *gender* significant? How

Chapter 9 – Repeated Measures and Mixed ANOVAs

would you describe it in nontechnical terms? d) Is there a significant difference between the *genders*? Is a post hoc multiple comparison test needed? Explain.

- 9.4 Compare the F and significance for *product* in Output 9.3 to the same F in Output 9.1. Compare the F s for the linear, quadratic, and cubic trends for *product* in Outputs 9.1 and 9.3. Why are some things different?

Extra SPSS Problems

- 9.1 Using the Anxiety 2.sav data file find out if there are there differences between the average score for the *trials*.
- 9.2 Using the Anxiety 2.sav data file: Are there *anxiety* as well as *trial* differences? Is there an interaction between *anxiety* and *trial*? Are post-hoc tests needed?

CHAPTER 10

Multivariate Analysis of Variance (MANOVA) and Canonical Correlation

In this chapter, we introduce multivariate analysis of variance (**MANOVA**), which is a complex statistic similar to ANOVA but with multiple dependent variables analyzed together. The dependent variables should be related conceptually, and they should be correlated with one another at a low to moderate level. If they are too highly correlated, one runs the risk of multicollinearity. If they are uncorrelated, there is usually no reason to analyze them together. The **General Linear Model** program in SPSS provides you with a multivariate F based on the linear combination of dependent variables that maximally distinguishes your groups. This multivariate result is the MANOVA. SPSS also automatically prints out univariate F s for the separate univariate ANOVAs for each dependent variable. Typically, these ANOVA results are not examined unless the multivariate results (the MANOVA) are significant, and some statisticians believe that they should not be used at all.

In this chapter, you will first do a one-way or single factor MANOVA. This problem has one independent variable (*father's education revised*) and three dependent variables (*grades*, *math achievement*, and *visualization test*). One could do three separate one-way ANOVAs; however, using MANOVA, you will see how the combination of the three variables distinguishes the groups, in one analysis. The second problem is a two-way or **two-factor MANOVA**. It is two-way because there are two independent variables (*math grades* and *gender*) and MANOVA because two test scores are examined simultaneously as dependent variables. Problem 10.3 is a doubly multivariate or **mixed MANOVA**. It is mixed because there is a between groups independent variable (intervention group) and a repeated measures independent variable (time: pretest and posttest). It is MANOVA because these are two (or more) dependent variables. Mixed MANOVAs are one way to analyze intervention (experimental) studies that have more than one dependent variable.

Finally, Problem 10.4 uses **canonical correlation** for a problem with two sets of two or more variables each. Canonical correlation relates variables in set 1 to those in set 2.

Assumptions of MANOVA

The assumptions of MANOVA include: independent observations (each person's scores are independent of every other person's scores), multivariate normality, and homogeneity of variance/covariance matrices (variances for each dependent variable are approximately equal in all groups plus covariances between pairs of dependent variables are approximately equal for all groups). MANOVA is robust to violations of multivariate normality and to violations of homogeneity of variance/covariance matrices if groups are of nearly equal size (N of the largest group is no more than about 1.5 times the N of the smallest group).

- Retrieve **hsbdataB**. Note: You are to get the **hsbdataB** file, *not the product* file from the last lab assignment.

Problem 10.1: GLM Single-Factor Multivariate Analysis of Variance

Sometimes you have more than one dependent variable that you want to analyze simultaneously. The GLM multivariate procedure allows you to analyze differences between levels of one or more (usually nominal level) independent variables, with respect to a linear combination of several dependent variables. One can also include scale variables (covariates) as predictors of the linear combination of dependent

Chapter 10 – Multivariate Analysis of Variance

variables. When you include both nominal variables and scale variables as predictors, one usually refers to the analysis as **MANCOVA** (Multivariate Analysis of Covariance).

10.1. Are there differences between the three *father's education* groups on a linear combination of *grades*, *math achievement*, and *visualization test*? Also, are there differences between groups on any of these variables separately? Which ones?

Before we answer these questions, we will correlate the dependent variables, to see if they are moderately correlated. To do this:

- Select **Analyze=> Correlate=>Bivariate**.
- Move *grades in h.s.*, *math achievement test*, and *visualization test* into the **Variables** box.
- Click on **Options** and select **Exclude cases listwise** (so that only participants with all three variables will be included in the correlations, just as they will be in the MANOVA).
- Click on **Continue**.
- Click on **OK**. Compare your output to 10.1a.

Output 10.1a: Intercorrelations of the Independent Variables

```
CORRELATIONS
/VARIABLES=grades mathach visual
/PRINT=TWOTAIL NOSIG
/MISSING=LISTWISE .
```

Correlations

The three circled correlations should be low to moderate.

Correlations^a

		grades in h.s.	math achievement test	visualization test
grades in h.s.	Pearson Correlation	1	.504**	.127
	Sig. (2-tailed)	.	.000	.279
math achievement test	Pearson Correlation	.504**	1	.423**
	Sig. (2-tailed)	.000	.	.000
visualization test	Pearson Correlation	.127	.423**	1
	Sig. (2-tailed)	.279	.000	.

** . Correlation is significant at the 0.01 level (2-tailed).

a. Listwise N=75

Interpretation of Output 10.1a

Look at the correlation table, to see if correlations are too high or too low. One correlation is a bit high: the correlation between *grades in h.s.* and *math achievement test* (.504). Thus, we will keep an eye on it in the MANOVA that follows. If the correlations were .60 or above, we would consider either making a composite variable (in which the highly correlated variables were summed or averaged) or eliminating one of the variables.

Now, to do the actual MANOVA, follow these steps:

- Select **Analyze => General Linear Model => Multivariate**.

- Move *grades in h.s.*, *math achievement*, and *visualization test* into the **Dependent Variables** box.
- Move *father's education revised* into the **Fixed Factor(s)** box (see Fig. 10.1).

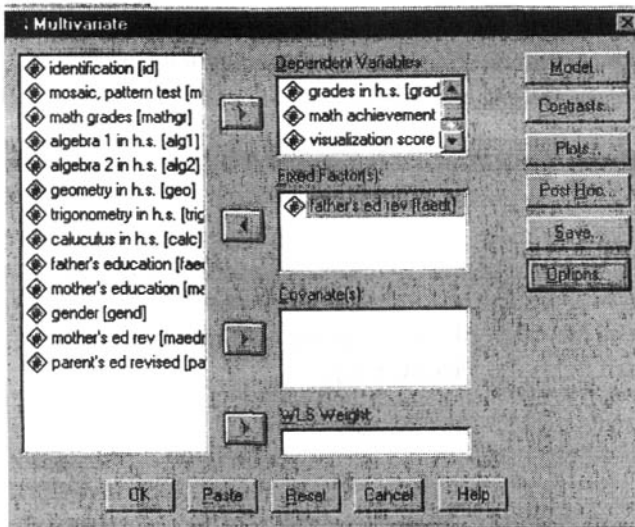


Fig. 10.1. Multivariate.

- Click on **Options**.
- Check **Descriptive statistics**, **Estimates of effect size**, **Parameter estimates**, and **Homogeneity tests** (see Fig. 10.2). These will enable us to check other assumptions of the test and see which dependent variables contribute most to distinguishing between groups.
- Click on **Continue**.

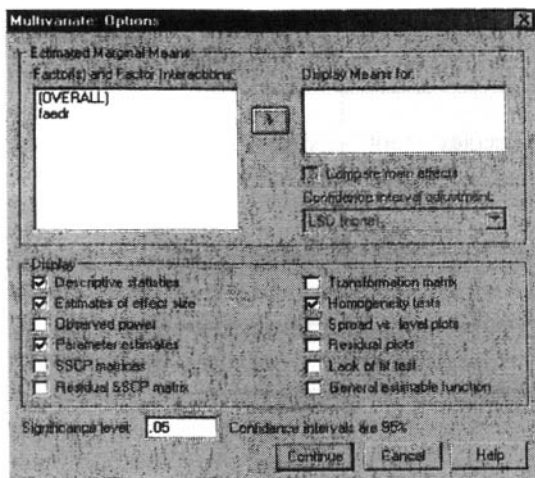


Fig. 10.2 Multivariate options.

- Click on **OK**. Compare your output to Output 10.1b.

Output 10.1b: One-Way Multivariate Analysis of Variance

GLM

```
grades mathach visual BY faedr
/METHOD = SSTYPE(3)
/INTERCEPT = INCLUDE
/PRINT = DESCRIPTIVE ETASQ PARAMETER HOMOGENEITY
/CRITERIA = ALPHA(.05)
/DESIGN = faedr .
```

General Linear Model

Between-Subjects Factors

		Value Label	N
father's education revised	1	HS grad or less	38
	2	Some College	16
	3	BS or More	19

To meet the assumptions it is best to have approximately equal cell sizes. Unfortunately, here the largest cell (38) is more than 1 ½ times the smallest. However, fortunately, Box's Test (below) indicates that that assumption is not violated.

Descriptive Statistics

	father's education revised	Mean	Std. Deviation	N
grades in h.s.	HS grad or less	5.34	1.475	38
	Some College	5.56	1.788	16
	BS or More	6.53	1.219	19
	Total	5.70	1.552	73
math achievement test	HS grad or less	10.0877	5.61297	38
	Some College	14.3958	4.66544	16
	BS or More	16.3509	7.40918	19
	Total	12.6621	6.49659	73
visualization test	HS grad or less	4.6711	3.96058	38
	Some College	6.0156	4.56022	16
	BS or More	5.4605	2.79044	19
	Total	5.1712	3.82787	73

Box's Test of Equality of Covariance Matrices

Box's M	18.443
F	1.423
df1	12
df2	10219.040
Sig.	.147

This checks the assumption of homogeneity of covariances across groups.

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

a. Design: Intercept+FAEDR

This indicates that there are no significant differences between the covariance matrices. Therefore, the assumption is not violated and Wilk's Lambda is an appropriate test to use (see the **Multivariate Tests** table).

SPSS for Intermediate Statistics

Multivariate Tests^a

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.938	341.884 ^a	3.000	68.000	.000	.938
	Wilks' Lambda	.062	341.884 ^a	3.000	68.000	.000	.938
	Hotelling's Trace	15.083	341.884 ^a	3.000	68.000	.000	.938
	Roy's Largest Root	15.083	341.884 ^a	3.000	68.000	.000	.938
FAEDR	Pillai's Trace	.229	2.970	6.000	138.000	.009	.114
	Wilks' Lambda	.777	3.040 ^a	6.000	136.000	.008	.118
	Hotelling's Trace	.278	3.106	6.000	134.000	.007	.122
	Roy's Largest Root	.245	5.645 ^b	3.000	69.000	.002	.197

This is the MANOVA using the Wilks' Lambda test.

a. Exact statistic

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

c. Design: Intercept+FAEDR

Because this is significant, we know that the assumption of homogeneity of variances is violated for *math achievement*. We could choose to transform this variable to enable it to meet the assumption. Given that the Box test is not significant and this is the only Levene test that is (barely) significant, we are opting to conduct the analysis anyway, but to use corrected follow-up tests.

Levene's Test of Equality of Error Variances^a

	F	df1	df2	Sig.
grades in h.s.	1.546	2	70	.220
math achievement test	3.157	2	70	.049
visualization test	1.926	2	70	.153

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

a. Design: Intercept+FAEDR

These are the three univariate analyses of variance.

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Eta
Corrected Model	grades in h.s.	18.143 ^a	2	9.071	4.091	.021	.105	.32 .43 .14
	math achievement test	558.481 ^b	2	279.240	7.881	.001	.184	
	visualization test	22.505 ^c	2	11.252	.763	.470	.021	
Intercept	grades in h.s.	2148.057	1	2148.057	968.672	.000	.933	
	math achievement test	11788.512	1	11788.512	332.697	.000	.826	
	visualization test	1843.316	1	1843.316	124.973	.000	.641	
FAEDR	grades in h.s.	18.143	2	9.071	4.091	.021	.105	
	math achievement test	558.481	2	279.240	7.881	.001	.184	
	visualization test	22.505	2	11.252	.763	.470	.021	
Error	grades in h.s.	155.227	70	2.218				
	math achievement test	2480.324	70	35.433				
	visualization test	1032.480	70	14.750				
Total	grades in h.s.	2544.000	73					
	math achievement test	14742.823	73					
	visualization test	3007.125	73					
Corrected Total	grades in h.s.	173.370	72					
	math achievement test	3038.804	72					
	visualization test	1054.985	72					

a. R Squared = .105 (Adjusted R Squared = .079)

b. R Squared = .184 (Adjusted R Squared = .160)

c. R Squared = .021 (Adjusted R Squared = -.007)

We have calculated the etas to help you interpret these effect sizes.

Chapter 10 – Multivariate Analysis of Variance

Parameter Estimates

Dependent Variable	Parameter	B	Std. Error	t	Sig.	95% Confidence Interval		Partial Eta Squared
						Lower Bound	Upper Bound	
grades in h.s.	Intercept	6.526	.342	19.103	.000	5.845	7.208	.839
	[FAEDR=1]	-1.184	.418	-2.830	.006	-2.019	-.350	.103
	[FAEDR=2]	-.984	.505	-1.907	.061	-1.972	4.393E-02	.049
	[FAEDR=3]	0 ^a						
math achievement test	Intercept	16.351	1.366	11.973	.000	13.627	19.075	.672
	[FAEDR=1]	-6.263	1.673	-3.745	.000	-9.599	-2.927	.167
	[FAEDR=2]	-1.955	2.020	-.968	.336	-5.983	2.073	.013
	[FAEDR=3]	0 ^a						
visualization test	Intercept	5.461	.881	6.198	.000	3.703	7.218	.354
	[FAEDR=1]	-.789	1.079	-.732	.467	-2.942	1.363	.008
	[FAEDR=2]	.555	1.303	.426	.671	-2.044	3.154	.003
	[FAEDR=3]	0 ^a						

a. This parameter is set to zero because it is redundant.

Each of the variables in brackets under **Parameter** comprises a dummy variable devised to distinguish one group from the others. For example, the circled weights (*B*) were devised to maximize differences between the first group (students whose fathers had high school education or less) and *all* other groups (those whose fathers had more education). Note that there are actually just two dummy variables, as a third would provide redundant information.

Interpretation of Output 10.1

The **GLM Multivariate** procedure provides an analysis for “effects” on a linear combination of several dependent variables of one or more fixed factor/independent variables and/or covariates. Note that many of the results (e.g., **Descriptive Statistics**, **Test of Between Subjects Effects**) refer to the univariate tests.

The first test we encounter is the **Box’s test of equality of covariance matrices**. This tests whether or not the covariances among the three dependent variables are the same for the three *father’s education* groups. The **Box test** is strongly affected by violations of normality and may not be accurate. If *N*s for the various groups are approximately equal, then the Box test should be ignored. Our largest group (*N* = 38) is 2.3 times larger than our smallest group (*N* = 16), so we should look at the Box test, which is not significant (*p* = .147). Thus, the assumption of homogeneity of covariances is *not* violated. If the Box test had been significant, we would have looked at the correlations among variables separately for the 3 groups and noted the magnitude of the discrepancies. **Pillai’s trace** is the best **Multivariate** statistic to use if there is violation of the homogeneity of covariance matrices assumption and group sizes are similar (not the case here). None of the multivariate tests would be robust if Box’s test had been significant and group sizes were very different.

MANOVA provides four multivariate tests (in the **Multivariate Tests** table). These tests examine whether the three *father’s education* groups differ on a linear combination of the dependent variables: *grades in h.s.*, *math achievement*, and *visualization test*. Under most conditions when assumptions are met, **Wilks’ Lambda** provides a good and commonly used multivariate *F* (in this case *F* = 3.04, *df* = 68, 136, *p* = .008). The “intercept” effect is just needed to fit the line to the data so skip over it. The main part of this multivariate test table to look at is the FAEDR effect. This significant *F* indicates that there are significant differences among the FAEDR groups on a linear combination of the three dependent variables.

Next, we see the **Levene’s test** table. This tests the assumption of MANOVA and ANOVA that the variances of each variable are equal across groups. If the Levene’s test is significant, as it is in this output for *math achievement*, this means the assumption has been violated. Results for *math achievement* should be viewed with caution (or the data could be transformed so as to equalize the variances).

Because the MANOVA was significant, we will now examine the univariate ANOVA results (in the

Tests of Between Subject Effects table). Note that these tests are identical to the three separate univariate one-way ANOVAs we would have performed if we opted not to do the MANOVA. Because the *grades in h.s.* and *math achievement* dependent variables are statistically significant and there are three levels or values of *father's education*, we would need to do post hoc multiple comparisons or contrasts to see which pairs of means are different.

Both multivariate and univariate (between subjects) tests provide measures of effect size (**eta squared**). For the multivariate test eta is .34 (the square root of .118), which is about a medium effect size. The univariate etas are .32, .43, .14 for *grades in h.s.*, *math achievement* and the *visualization test*, respectively. The first one is a medium effect and the second is a large effect. The eta for *visualization* indicates a small effect, and because $F(.76)$ is not significant ($p=.470$), this result could be due to chance (See Table 3.5 on page 55 for interpretation of the effect size for eta.).

In MANOVA, a linear combination of the dependent variables is created and groups are compared on that variable. To create this linear combination for each participant, the computer multiplies the participant's score on each variable by a weight (B), with the values of the weights being devised so as to maximize differences between groups. Next, we see the **Parameter Estimates**, which tell us how the dependent variables are weighted in the equation that maximally distinguishes the groups. Note that in the column under **Parameter** in this table, three variables are listed that seem new. These are the dummy variables that were used to test for differences between groups. The first one [$FAEDR = 1$] indicates differences between students whose fathers have high school education or less and the other students whose fathers have more education. The second one [$FAEDR = 2$] indicates differences between students whose fathers have some college and students in the other 2 groups. A third dummy variable would provide redundant information and, thus, is not considered; there are $k-1$ independent dummy variables, where k = number of groups. The next column, headed by **B** indicates the weights for the dependent variables for that dummy variable. For example, in order to distinguish students whose fathers have high school education or less from other students, *math achievement* is weighted highest in absolute value (-6.263), followed by *grades in h.s.* (-1.184), and then *visualization test* (-.789). In all cases, students whose fathers have less education score lower than other students, as indicated by the minus signs. This table can also tell us which variables significantly contributed toward distinguishing which groups, if you look at the **sig** column for each dummy variable. For example, both *grades in high school* and *math achievement* contributed significantly toward discriminating group 1 (high school grad or less) from the other two groups, but no variables significantly contributed to distinguishing group 2 (some college) from the other two groups (although *grades in high school* discriminates group 2 from the others at almost significant levels). *Visualization*, does not significantly contribute to distinguishing any of the groups.

We can look at the ANOVA (**Between Subjects**) and **Parameter Estimates** table results to determine whether the groups differ on each of these variables, examined alone. This will help us in determining whether multicollinearity affected results because if two or more of the ANOVAs are significant, but the corresponding variable(s) are not weighted much (examine the **B** scores) in the MANOVA, this probably is because of multicollinearity. The ANOVAs also help us understand which variables, separately, differ across groups. Note again that some statisticians think that it is not appropriate to examine the univariate ANOVAs. Traditionally, univariate F s have been analyzed to understand where the differences are when there is a significant multivariate F . One argument against reporting the univariate F is that the univariate F s do not take into account the relations among the dependent variables; thus the variables that are significant in the univariate tests are not always the ones that are weighted most highly in the multivariate test. Univariate F s also can be confusing because they will sometimes be significant when the multivariate F is not. Furthermore, if one is using the MANOVA to reduce Type I error by analyzing all dependent variables together, then analyzing the univariate F 's "undoes" this benefit, thus, increasing Type I error. One method to compensate for this is to use the Bonferroni correction to adjust the alpha used to determine statistical significance of the univariate F s.

How to Write about Output 10.1

Results

A multivariate analysis of variance was conducted to assess if there were differences between the three father's education groups on a linear combination of grades in h.s., math achievement, and visualization test. A significant difference was found, Wilk's $\Lambda = .777$, $F(68, 136) = 3.04$, $p = .008$, multivariate $\eta^2 = .12$. Examination of the coefficients for the linear combinations distinguishing father education groups indicated that grades in high school and math achievement contributed most to distinguishing the groups. In particular, both grades in high school (-1.18) and math achievement (-6.26) contributed significantly toward discriminating group 1 (high school grad or less) from the other two groups ($p = .006$ and $p < .001$, respectively), but no variables significantly contributed to distinguishing group 2 (some college) from the other two groups. Visualization did not contribute significantly to distinguishing any of the groups.

Follow up univariate ANOVAs indicated that both math achievement and grades in high school were significantly different for children of fathers with different degrees of education, $F(2, 70) = 7.88$, $p = .001$ and $F(2, 70) = 4.09$, $p = .021$, respectively.

Problem 10.2: GLM Two-Factor Multivariate Analysis of Variance

MANOVA is also useful when there is more than one independent variable and several related dependent variables. Let's answer the following questions:

- 10.2. Do students who differ in *math grades* and *gender* differ on a linear combination of two dependent variables (*math achievement*, and *visualization test*)? Do males and females differ in terms of whether those with higher and lower *math grades* differ on these two variables (is there an interaction between *math grades* and *gender*)? What linear combination of the two dependent variables distinguishes these groups?

We already know the correlation between the two dependent variables is moderate (.42), so we will omit the correlation matrix. Follow these steps:

- Select **Analyze => General Linear Model => Multivariate**.
- Click on **Reset**.
- Move *math achievement* and *visualization test* into the **Dependent Variables** box (see Fig. 10.1 if you need help).
- Move both *math grades* and *gender* into the **Fixed Factor(s)** box.
- Click on **Options**.
- Check **Descriptive statistics, Estimates of effect size, Parameter estimates, and Homogeneity tests** (see Fig. 10.2).
- Click on **Continue**.
- Click on **OK**. Compare your output to Output 10.2.

Output 10.2: Two-Way Multivariate Analysis of Variance

```
GLM
  mathach visual  BY mathgr gender
/METHOD = SSTYPE(3)
/INTERCEPT = INCLUDE
/PRINT = DESCRIPTIVE ETASQ PARAMETER HOMOGENEITY
```

SPSS for Intermediate Statistics

```
/CRITERIA = ALPHA(.05)
/DESIGN = mathgr gender mathgr*gender .
```

General Linear Model

Between-Subjects Factors

		Value Label	N
math grades	0	less A-B	44
	1	most A-B	31
gender	0	male	34
	1	female	41

Descriptive Statistics

	math grades	gender	Mean	Std. Deviation	N
math achievement test	less A-B	male	12.8751	5.73136	24
		female	8.3333	5.32563	20
		Total	10.8108	5.94438	44
	most A-B	male	19.2667	4.17182	10
		female	13.0476	7.16577	21
		Total	15.0538	6.94168	31
	Total	male	14.7550	6.03154	34
		female	10.7479	6.69612	41
		Total	12.5645	6.67031	75
visualization test	less A-B	male	5.7188	4.52848	24
		female	3.2750	2.74209	20
		Total	4.6080	3.97572	44
	most A-B	male	8.1250	4.04188	10
		female	5.2024	3.20119	21
		Total	6.1452	3.69615	31
	Total	male	6.4265	4.47067	34
		female	4.2622	3.10592	41
		Total	5.2433	3.91203	75

Box's Test of Equality of Covariance Matrices ^a

Box's M	12.437
F	1.300
df1	9
df2	12723.877
Sig.	.231

This assumption is not violated.

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

a. Design: Intercept+MATHGR+GEND+MATHGR * GEND

Chapter 10 – Multivariate Analysis of Variance

Multivariate Tests ^b

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.848	195.012 ^a	2.000	70.000	.000	.848
	Wilks' Lambda	.152	195.012 ^a	2.000	70.000	.000	.848
	Hotelling's Trace	5.572	195.012 ^a	2.000	70.000	.000	.848
	Roy's Largest Root	5.572	195.012 ^a	2.000	70.000	.000	.848
MATHGR	Pillai's Trace	.189	8.155 ^a	2.000	70.000	.001	.189
	Wilks' Lambda	.811	8.155 ^a	2.000	70.000	.001	.189
	Hotelling's Trace	.233	8.155 ^a	2.000	70.000	.001	.189
	Roy's Largest Root	.233	8.155 ^a	2.000	70.000	.001	.189
GEND	Pillai's Trace	.200	8.743 ^a	2.000	70.000	.000	.200
	Wilks' Lambda	.800	8.743 ^a	2.000	70.000	.000	.200
	Hotelling's Trace	.250	8.743 ^a	2.000	70.000	.000	.200
	Roy's Largest Root	.250	8.743 ^a	2.000	70.000	.000	.200
MATHGR * GEND	Pillai's Trace	.005	.171 ^a	2.000	70.000	.843	.005
	Wilks' Lambda	.995	.171 ^a	2.000	70.000	.843	.005
	Hotelling's Trace	.005	.171 ^a	2.000	70.000	.843	.005
	Roy's Largest Root	.005	.171 ^a	2.000	70.000	.843	.005

a. Exact statistic

b. Design: Intercept+MATHGR+GEND+MATHGR * GEND

The effects of *math grades* and *gender* on the combination dependent variable are significant but the interaction is not.

Levene's Test of Equality of Error Variances^a

	F	df1	df2	Sig.
math achievement test	1.691	3	71	.177
visualization test	2.887	3	71	.042

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

a. Design: Intercept+MATHGR+GEND+MATHGR * GEND

Because this is significant, we know that the assumption of homogeneity of variances is violated for visualization. However, since groups are nearly equal in size, the test should not be strongly affected by this violation.

These etas were computed so you would not have to compute them.

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Eta
Corrected Model	math achievement test	814.481 ^a	3	271.494	7.779	.000	.247	
	visualization test	165.986 ^b	3	55.329	4.064	.010	.147	
Intercept	math achievement test	11971.773	1	11971.773	343.017	.000	.829	
	visualization test	2082.167	1	2082.167	152.966	.000	.683	
MATHGR	math achievement test	515.463	1	515.463	14.769	.000	.172	.41
	visualization test	78.485	1	78.485	5.766	.019	.075	.27
GEND	math achievement test	483.929	1	483.929	13.866	.000	.163	.40
	visualization test	120.350	1	120.350	8.841	.004	.111	.33
MATHGR * GEND	math achievement test	11.756	1	11.756	.337	.563	.005	.07
	visualization test	.958	1	.958	.070	.792	.001	.03
Error	math achievement test	2478.000	71	34.901				
	visualization test	966.510	71	13.613				
Total	math achievement test	15132.393	75					
	visualization test	3194.438	75					
Corrected Total	math achievement test	3292.481	74					
	visualization test	1132.497	74					

a. R Squared = .247 (Adjusted R Squared = .216)

b. R Squared = .147 (Adjusted R Squared = .111)

SPSS for Intermediate Statistics

Parameter Estimates

Dependent Variable	Parameter	B	Std. Error	t	Sig.	95% Confidence Interval		Partial Eta Squared
						Lower Bound	Upper Bound	
math achievement test	Intercept	13.048	1.289	10.121	.000	10.477	15.618	.591
	[MATHGR=0]	-4.714	1.848	-2.554	.013	-8.395	-1.034	.084
	[MATHGR=1]	0 ^a
	[GEND=0]	6.219	2.270	2.740	.008	1.693	10.745	.096
	[GEND=1]	0 ^a
	[MATHGR=0] * [GEND=0]	-1.677	2.690	-.580	.563	-7.440	4.085	.005
	[MATHGR=0] * [GEND=1]	0 ^a
	[MATHGR=1] * [GEND=0]	0 ^a
visualization test	Intercept	5.202	.805	6.462	.000	3.597	6.808	.370
	[MATHGR=0]	-1.927	1.153	-1.672	.099	-4.226	.371	.038
	[MATHGR=1]	0 ^a
	[GEND=0]	2.923	1.418	2.062	.043	9.806E-02	5.749	.056
	[GEND=1]	0 ^a
	[MATHGR=0] * [GEND=0]	-.479	1.805	-.265	.792	-4.078	3.120	.001
	[MATHGR=0] * [GEND=1]	0 ^a
	[MATHGR=1] * [GEND=0]	0 ^a
	[MATHGR=1] * [GEND=1]	0 ^a

Note that the weights for *math achievement* are larger than those for *visualization test*.

a. This parameter is set to zero because it is redundant.

Interpretation of Output 10.2

Many of the tables are similar to those in Output 10.1. For the **Descriptive Statistics**, we now see means and standard deviations of the dependent variables for the groups made up of every combination of the two levels of *math grades* and the two levels of *gender*. **Box's test** is again non-significant, indicating that the assumption of homogeneity of covariance matrices is met.

The main difference in 10.2, as compared to 10.1, for both the **Multivariate Tests** table and the univariate **Tests of Between Subjects Effects** table, is the inclusion of two main effects (one for each independent variable) and one interaction (of *math grades* and *gender*: *math grades* * *gender*). The interpretation of this interaction, if it were significant, would be similar to that in Output 10.1. However, note that although both the multivariate main effects of *math grades* and *gender* are significant, the multivariate interaction is not significant. Thus, we can look at the univariate tests of main effects, but we should not examine the univariate interaction effects.

The **Levene's test** indicates that there is heterogeneity of variances for *visualization test*. Again, we could have transformed that variable to equalize the variances. However, if we consider only the main effects of *gender* and of *math grades* (since the interaction is not significant), then *Ns* are approximately equal for the groups (34 and 41 for *gender* and 31 and 44 for *math grades*), so this is less of a concern.

The **Tests of Between-Subjects Effects** table indicates that there are significant main effects of both independent variables on both dependent variables, with medium to large effect sizes. For example, the "effect" of *math grades* on *math achievement* is large ($\eta^2 = .41$) and the effect of *math grades* on *visualization test* is medium ($\eta^2 = .27$). Refer again to Table 3.5.

The **Parameter Estimates** table now has three dummy variables: for the difference between students with less A-B and more A-B (MATHGR=0), for male versus not male (GEND=0), as well as for the interaction term (MATHGR=0 GEND=0). Thus, we can see that *math achievement* contributes more than *visualization test* to distinguishing students with better and worse math grades, as well as contributing more to distinguishing boys from girls.

How to Write about Output 10.2**Results**

To assess whether boys and girls with higher and lower math grades have different math achievement and visualization test scores, and whether there was an interaction between gender and math grades, a multivariate analysis of variance was conducted. The interaction was not significant, Wilk's $\Lambda = .995$, $F(2, 70) = .17$, $p = .843$, multivariate $\eta^2 = .005$. The main effect for gender was significant, Wilk's $\Lambda = .800$, $F(2, 70) = 8.74$, $p < .001$, multivariate $\eta^2 = .20$. This indicates that the linear composite of math achievement and visualization test differs for males and females. The main effect for math grades is also significant, Wilk's $\Lambda = .811$, $F(2, 70) = 8.15$, $p = .001$, multivariate $\eta^2 = .19$. This indicates that the linear composite differs for different levels of math grades. Follow-up ANOVAs (Table 10.2) indicate that effects of both math grades and gender were significant for both math achievement and visualization. Males scored higher on both outcomes and students with higher math grades were higher on both outcomes (see Table 10.1).

Table 10.1

Means and Standard Deviations for Math Achievement and Visualization Test as a Function of Math Grades and Gender

		Math achievement		Visualization	
Group	<i>n</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Low math grades					
Males	24	12.88	5.73	5.72	4.53
Females	20	8.33	5.33	3.28	2.74
High math grades					
Males	10	19.27	4.17	8.13	4.04
Females	21	13.05	7.17	5.20	3.20

Table 10.2

Effects of Math Grades and Gender on Math Achievement and Visualization Test Scores

Source	Dependent Variable	df	F	η	p
Math Grades	Math achievement	1	14.77	.41	.001
	Visualization test	1	5.77	.27	.019
Gender	Math achievement	1	13.87	.40	.001
	Visualization test	1	8.84	.33	.004
Math Grades x Gender	Math achievement	1	.34	.07	.563
	Visualization test	1	.07	.03	.792
Error	Math achievement	71			
	Visualization test	71			

Problem 10.3: Mixed MANOVA

There might be times when you want to find out if there are differences between groups as well as within subjects; this can be answered with Mixed MANOVA.

Assumptions of Repeated and Mixed MANOVA

The assumptions of repeated or doubly multivariate (multiple measures obtained on multiple occasions) MANOVA include: linearity of relations among the dependent variables and multivariate normality; for mixed MANOVA, an additional assumption is homogeneity of variance-covariance matrices between groups. If the sample sizes in each group are approximately equal, repeated MANOVA is robust to these assumptions. Linearity can be assessed through matrix scatterplots. If sample sizes are not equal, than Box's M test can be conducted to check the homogeneity assumption, and univariate normality can be assessed using plots to help assess the normality assumption. One should also make sure that sample sizes are sufficiently large. If the number of variables times the number of levels of the within subjects variable approaches the sample size, then the doubly multivariate analysis should not be conducted. Rather, one should either use multiple MANOVAs with only one or two dependent variable each (if there are only two, three, or four dependent variables), or one should aggregate or drop some dependent variables (if there are more dependent variables).

We have created a new dataset to use for this problem (MixedMANOVAdata).

- Retrieve MixedMANOVAdata.

Let's answer the following question:

- 10.3. Is there a difference between participants in the intervention group (group 1) and participants in the control group (group 2) in the amount of change that occurs over time in scores on two different *outcome* measures?

We will not check the assumptions of linearity, multivariate normality, or homogeneity of variance-covariance matrices because the sample sizes are equal.

- **Analyze => General Linear Model => Repeated Measures** (see Fig. 10.3).
- **Delete** the factor 1 from the **Within-Subject Factor Name** box and **replace** it with the name **time**, our name for the repeated measures independent variable that SPSS will generate.
- **Type 2** in the **Number of Levels** box.
- Click on **Add**.
- In the **Measure Name** box, type **dv1**.
- Click on **Add**.
- In the **Measure Name** box, type **dv2**.
- Click on **Add**. The window should look like Fig. 10.3.

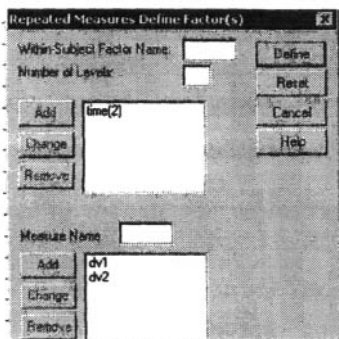


Fig. 10.3. Repeated measures define factor(s).

- Click on **Define**, which changes the screen to a new menu box (see Fig. 9.4 in chapter 9 if you need help).
- Now move *outcome 1 pretest*, *outcome 1 posttest*, *outcome 2 pretest*, and *outcome 2 posttest* over to the **Within-Subjects Variables** box.
- Move *group* to the **Between-Subjects Factor(s)** box.
- Click on **Plots**.
- Move *time* to the **Horizontal Axis** box.
- Move *group* to the **Separate Lines** box.
- Click on **Add**. This will move the variables down to the **Plots** box.
- Click on **Continue**.
- Click on **Options**. (see Fig. 9.6 in chapter 9 if you need help).
- Click on **Descriptive Statistics** and **Estimates of effect size**.
- Click on **Continue**, then on **OK**.

Compare your syntax and output to Output 10.3.

Output 10.3: Repeated Measures MANOVA Using the General Linear Model

```
GLM
  DV11 DV12 DV21 DV22 BY group
  /WSFACTOR = time 2
  /MEASURE = dv1 dv2
  /METHOD = SSTYPE(3)
  /PLOT = PROFILE( time*group )
  /PRINT = DESCRIPTIVE ETASQ
  /CRITERIA = ALPHA(.05)
  /WSDESIGN = time
  /DESIGN = group .
```

General Linear Model

Within-Subjects Factors

Measure	time	Dependent Variable
dv1	1	DV11
	2	DV12
dv2	1	DV21
	2	DV22

Between-Subjects Factors

	Value Label	N
group	1	10
	2	10
	intervention	
	comparison	

SPSS for Intermediate Statistics

Descriptive Statistics

	group	Mean	Std. Deviation	N
outcome 1 pretest	intervention	5.00	1.826	10
	comparison	5.10	1.370	10
	Total	5.05	1.572	20
outcome 1 posttest	intervention	10.00	1.414	10
	comparison	5.50	1.780	10
	Total	7.75	2.789	20
outcome 2 pretest	intervention	20.00	7.454	10
	comparison	20.50	6.852	10
	Total	20.25	6.973	20
outcome 2 posttest	intervention	20.50	6.852	10
	comparison	19.00	8.097	10
	Total	19.75	7.340	20

This table tells you what effects are significant. In this case, there are significant multivariate main effects of *group* and *time*, but these are qualified by a *group* by *time* interaction. This indicates that the difference between the intervention and control group on the linear combination of the two dependent variables is different at pretest than at posttest. Examination of the means suggests that this is because groups do not differ on either dependent variable at the time of the pretest, but they do differ, particularly on the first dependent variable, at the time of the posttest.

Multivariate Tests^b

Effect			Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Between Subjects	Intercept	Pillai's Trace	.955	182.194 ^a	2.000	17.000	.000	.955
		Wilks' Lambda	.045	182.194 ^a	2.000	17.000	.000	.955
		Hotelling's Trace	21.435	182.194 ^a	2.000	17.000	.000	.955
		Roy's Largest Root	21.435	182.194 ^a	2.000	17.000	.000	.955
	group	Pillai's Trace	.553	10.499 ^a	2.000	17.000	.001	.553
		Wilks' Lambda	.447	10.499 ^a	2.000	17.000	.001	.553
		Hotelling's Trace	1.235	10.499 ^a	2.000	17.000	.001	.553
		Roy's Largest Root	1.235	10.499 ^a	2.000	17.000	.001	.553
Within Subjects	time	Pillai's Trace	.822	39.321 ^a	2.000	17.000	.000	.822
		Wilks' Lambda	.178	39.321 ^a	2.000	17.000	.000	.822
		Hotelling's Trace	4.626	39.321 ^a	2.000	17.000	.000	.822
		Roy's Largest Root	4.626	39.321 ^a	2.000	17.000	.000	.822
	time * group	Pillai's Trace	.786	31.235 ^a	2.000	17.000	.000	.786
		Wilks' Lambda	.214	31.235 ^a	2.000	17.000	.000	.786
		Hotelling's Trace	3.675	31.235 ^a	2.000	17.000	.000	.786
		Roy's Largest Root	3.675	31.235 ^a	2.000	17.000	.000	.786

a. Exact statistic

b.

Design: Intercept+group

Within Subjects Design: time

Chapter 10 – Multivariate Analysis of Variance

Mauchly's Test of Sphericity^b

Within Subjects Effect	Measure	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon ^a		
						Greenhouse-Geisser	Huynh-Feldt	Lower-bound
time	dv1	1.000	.000	0	.	1.000	1.000	1.000
	dv2	1.000	.000	0	.	1.000	1.000	1.000

Tests the null hypothesis that the error covariance matrix of the orthonormalized transformed dependent variables 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+group
Within Subjects Design: time

Note that Mauchly's test of sphericity would be needed only in relation to the assumptions of the follow-up univariate repeated measures tests (second table below, labeled "univariate tests"); sphericity is not required for the multivariate tests. In this case, it is not needed even for univariate tests because there are only two levels of the within subjects variable, so sphericity is not an issue and all epsilons are 1.0. The follow-up repeated measures ANOVAs (see univariate effects table) and contrasts test whether these apparent findings are statistically reliable.

Tests of Within-Subjects Effects

Multivariate^{b,c}

Within Subjects Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
time	Pillai's Trace	.822	39.321 ^a	2.000	17.000	.000	.822
	Wilks' Lambda	.178	39.321 ^a	2.000	17.000	.000	.822
	Hotelling's Trace	4.626	39.321 ^a	2.000	17.000	.000	.822
	Roy's Largest Root	4.626	39.321 ^a	2.000	17.000	.000	.822
time * group	Pillai's Trace	.786	31.235 ^a	2.000	17.000	.000	.786
	Wilks' Lambda	.214	31.235 ^a	2.000	17.000	.000	.786
	Hotelling's Trace	3.675	31.235 ^a	2.000	17.000	.000	.786
	Roy's Largest Root	3.675	31.235 ^a	2.000	17.000	.000	.786

a. Exact statistic

b.

Design: Intercept+group
Within Subjects Design: time

c. Tests are based on averaged variables.

Note that this table provides information that is redundant with the previous multivariate table. You can ignore it.

SPSS for Intermediate Statistics

Univariate Tests

Source	Measure		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
time	dv1	Sphericity Assumed	72.900	1	72.900	81.000	.000	.818
		Greenhouse-Geisser	72.900	1.000	72.900	81.000	.000	.818
		Huynh-Feldt	72.900	1.000	72.900	81.000	.000	.818
		Lower-bound	72.900	1.000	72.900	81.000	.000	.818
	dv2	Sphericity Assumed	2.500	1	2.500	.240	.630	.013
		Greenhouse-Geisser	2.500	1.000	2.500	.240	.630	.013
		Huynh-Feldt	2.500	1.000	2.500	.240	.630	.013
		Lower-bound	2.500	1.000	2.500	.240	.630	.013
time * group	dv1	Sphericity Assumed	52.900	1	52.900	58.778	.000	.766
		Greenhouse-Geisser	52.900	1.000	52.900	58.778	.000	.766
		Huynh-Feldt	52.900	1.000	52.900	58.778	.000	.766
		Lower-bound	52.900	1.000	52.900	58.778	.000	.766
	dv2	Sphericity Assumed	10.000	1	10.000	.960	.340	.051
		Greenhouse-Geisser	10.000	1.000	10.000	.960	.340	.051
		Huynh-Feldt	10.000	1.000	10.000	.960	.340	.051
		Lower-bound	10.000	1.000	10.000	.960	.340	.051
Error(time)	dv1	Sphericity Assumed	16.200	18	.900			
		Greenhouse-Geisser	16.200	18.000	.900			
		Huynh-Feldt	16.200	18.000	.900			
		Lower-bound	16.200	18.000	.900			
	dv2	Sphericity Assumed	187.500	18	10.417			
		Greenhouse-Geisser	187.500	18.000	10.417			
		Huynh-Feldt	187.500	18.000	10.417			
		Lower-bound	187.500	18.000	10.417			

This table displays follow-up repeated measures ANOVAs for each dependent variable, which show that the main effect of time (change from pretest to posttest) is significant only for dependent variable 1, and that the interaction between group and time is only significant for dependent variable 1. This indicates that the change over time is associated with the intervention, but only for dependent variable 1.

This table indicates that there is a significant linear trend only for dependent variable 1 and that this time effect interacted with group. Examination of the means suggests that this interaction indicates that the time effect held only for the intervention group. The linear trend is the only polynomial contrast used because there are only two groups and only two time points, so quadratic or cubic trends are not possible. This table also provides effect sizes (partial eta squared), which indicate that the effects for dv1 were large.

Tests of Within-Subjects Contrasts

Source	Measure	time	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
time	dv1	Linear	72.900	1	72.900	81.000	.000	.818
	dv2	Linear	2.500	1	2.500	.240	.630	.013
time * group	dv1	Linear	52.900	1	52.900	58.778	.000	.766
	dv2	Linear	10.000	1	10.000	.960	.340	.051
Error(time)	dv1	Linear	16.200	18	.900			
	dv2	Linear	187.500	18	10.417			

Chapter 10 – Multivariate Analysis of Variance

Tests of Between-Subjects Effects

Transformed Variable: Average

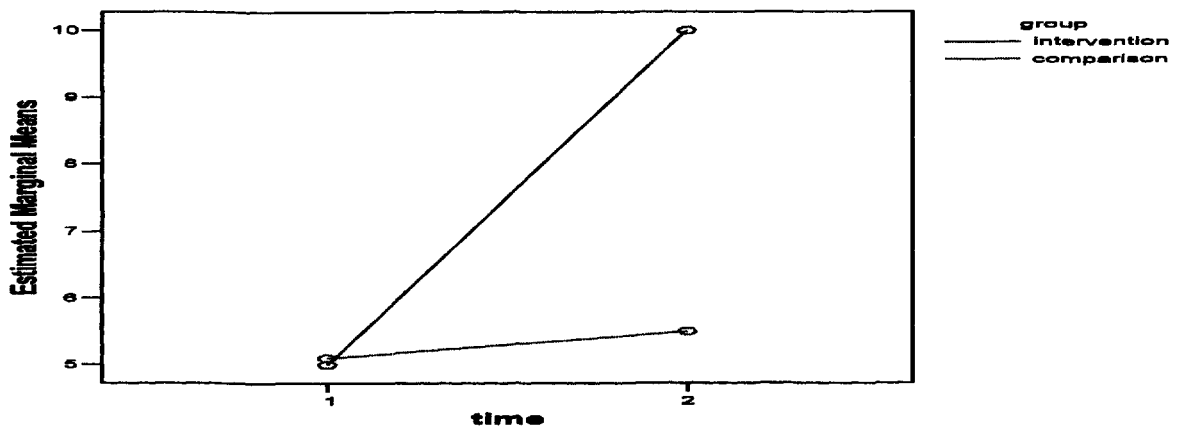
Source	Measure	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	dv1	1638.400	1	1638.400	382.010	.000	.955
	dv2	16000.000	1	16000.000	164.807	.000	.902
group	dv1	48.400	1	48.400	11.285	.003	.385
	dv2	2.500	1	2.500	.026	.874	.001
Error	dv1	77.200	18	4.289			
	dv2	1747.500	18	97.083			

This table indicates that if one averages across the within subjects variable (time), then the groups differ only on dependent variable 1. This table is misleading because we know that there is no difference at the time of the pretest. In this case, it really does not provide much useful information.

Profile Plots

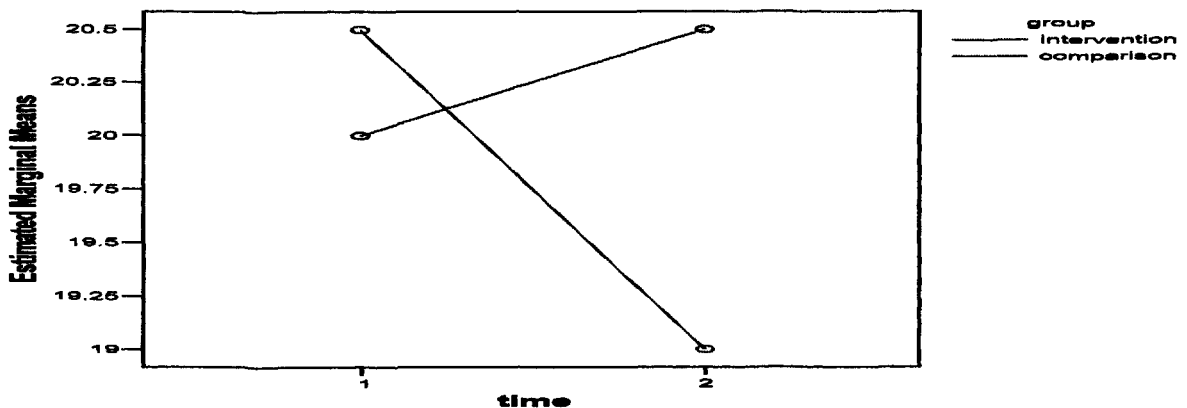
dv1

Estimated Marginal Means of dv1



dv2

Estimated Marginal Means of dv2



Interpretation of Output 10.3

This example illustrates the utility of the doubly multivariate analysis for testing a pretest-posttest design with intervention and control groups. Not only do we see from these results that the intervention seemed successful when both *outcome* measures were taken together, but that the effect was significant only for one of the dependent variables, when each was considered separately. Sphericity was not an issue in this case because there were only two levels of the within-subjects variable. If one creates difference scores by subtracting each person's score on the dependent variable at one level of the within-subjects variable from the same dependent variable at each other level of the within-subjects variable, then sphericity exists if the variances of the resulting difference scores are all equal. Because there are only two levels of the within-subjects variable, there is only one set of difference scores, so sphericity has to exist. If we had had more than two levels of the within-subjects variable, then we would have needed to be concerned about the sphericity assumption when examining the univariate results. If epsilons did not approach 1.0, then we would have used the Huynh-Feldt or Greenhouse-Geisser test results, which use an estimate of epsilon to correct the degrees of freedom.

Example of How to Write About Problem 10.3**Results**

A doubly multivariate analysis was conducted to assess if there was a difference between participants in the intervention group and participants in the control group in the amount of change in their scores on the two outcome measures. Significant multivariate effects were found for the main effects of group, $F(2,17) = 10.5, p = .001$ and time $F(2,17) = 39.3, p < .001$, as well as for the interaction between group and time, $F(2,17) = 31.2, p < .001$. This interaction effect indicates that the difference between the intervention and control group on the linear combination of the two dependent variables is different at pretest than it is at posttest. Examination of the means suggests that this is because groups do not differ on either dependent variable at the time of the pretest, but they do differ, particularly on the first dependent variable, at the time of the posttest. Follow-up ANOVAs reveal that the significant change from pretest to posttest was significant only for the first outcome variable, $F(1,18) = 81, p < .001$, and that the change in the first outcome variable was different for the two groups, $F(1,18) = 58.78, p < .001$. Examination of the means suggests that the change in the first outcome variable only held for the intervention group.

Problem 10.4: Canonical Correlation

Canonical correlation is similar to MANOVA in that it is used when you have two sets of two or more variables each and you want to see how differences in one set relate to differences in the other set of variables. With canonical correlation, however, there is no distinction between independent and dependent variables; they are called by SPSS "Set 1" and "Set 2." Moreover, usually canonical correlation is used when both sets of variables are at least interval level (scale), whereas in MANOVA usually at least some of the independent variables are dummy variables. Usually one would use canonical correlation when the variables in each set can be grouped together conceptually, but you want to see if there are particular subsets of them that relate to subsets in the other variable set, so you do not want to sum each set to make an overall score. Usually, canonical correlation is used as an exploratory technique; it is not commonly used to test specific hypotheses. Like principle components analysis, canonical correlation enables you to see which variables go together; however, it determines which subset of the "Set 1" variables best relate to which subset of the "Set 2" variables, then which other subset of the "Set 1" variables relate to another subset of the "Set 2" variables, etc. (Note: For those of you using earlier versions of SPSS, it has been reported that SPSS versions below 10.0 tend to have problems running canonical correlations. If you are using an earlier version and have problems, check the Help menu.)

Assumptions of Canonical Correlation

The assumptions of canonical correlation include: linearity (between each variable as well as between the variables and the linear composites), multivariate normality, and homoscedasticity. Because multivariate normality is difficult to assess, univariate normality can be evaluated. Multicollinearity should be assessed as well. All of the assumptions can be evaluated through a matrix scatterplot of the canonical variate scores (the scores are generated by the canonical correlation syntax). It is recommended to have at least 10 subjects per variable in order to have adequate power.

10.4. What is the pattern of relationships between the motivation items and the competence items?

We will need to check the assumptions of linearity, multivariate normality, and homoscedasticity. One way to do this is to graph the canonical variate scores that are generated by the canonical correlation syntax. Because the canonical correlation needs to be computed first, we have done that below, with the matrix scatterplot of the canonical variates afterwards. Be sure not to save your data file after running these commands, or you will save the canonical variates.

Canonical correlation must be computed through syntax. Use the following syntax to answer the question above. The first line needs to include where your SPSS syntax macro file is located on your computer (in this case, it is on the c: drive under program files\spss, but yours might be located in a different folder).

```
INCLUDE 'c:\Program files\spss\canonical correlation.sps'.
CANCORR SET1=item01 item07 item12 item13 /
        SET2=item03 item09 / .
```

Compare your output with Output 10.4a below.

Output 10.4a: Canonical Correlation Output

```
INCLUDE 'c:\Program files\spss\canonical correlation.sps'.
 32 preserve.
 33 set printback=off.
438 RESTORE.
439
441 * End of INCLUDE nesting level 01.
CANCORR SET1=item01 item07 item12 item13 /
        SET2=item03 item09 / .
NOTE: ALL OUTPUT INCLUDING ERROR MESSAGES HAVE BEEN TEMPORARILY
SUPPRESSED. IF YOU EXPERIENCE UNUSUAL BEHAVIOR THEN RERUN THIS
MACRO WITH AN ADDITIONAL ARGUMENT /DEBUG='Y'.
BEFORE DOING THIS YOU SHOULD RESTORE YOUR DATA FILE.
THIS WILL FACILITATE FURTHER DIAGNOSTICS OF ANY PROBLEMS
```

Matrix

Run MATRIX procedure:

```
Correlations for Set-1
      item01  item07  item12  item13
item01  1.0000   .4613   .1841   .1856
item07   .4613  1.0000   .3469   .3562
item12   .1841   .3469  1.0000   .5838
item13   .1856   .3562   .5838  1.0000

Correlations for Set-2
      item03  item09
item03  1.0000   .3268
item09   .3268  1.0000
```

These are the correlations among the items of each set. Remember to look on only one side of the diagonal.

SPSS for Intermediate Statistics

These are bivariate correlations between each variable in Set 1 and each variable in Set 2.

Correlations Between Set-1 and Set-2

	item03	item09
item01	.6260	.2082
item07	.4238	.2265
item12	.1662	.3944
item13	.1669	.2894

You can square these values to help with interpretation. Thus, Function1 accounts for 42% of the variance shared between Set 1 & Set 2 $(.645)^2$.

The Wilks' lambda and corresponding chi-square tests indicate that the canonical correlation on that line plus all later correlations are significantly different from zero. Thus, the first line tests whether both correlations, together, are significantly different from zero. The second line tests whether the second correlation is significant, even after removing the variance explained by the previous correlation. The degrees of freedom are calculated as the number of variables in Set 1 times number of variables in Set 2 for the first chi-square, and then df is decreased by 1 for each variable set for each remaining chi-square.

Canonical Correlations

1	.645
2	.368

Test that remaining correlations are zero:

	Wilk's	Chi-SQ	DF	Sig.
1	.505	46.175	8.000	.000
2	.865	9.808	3.000	.020

Standardized Canonical Coefficients for Set-1

	1	2
item01	-.838	-.276
item07	-.259	-.089
item12	-.047	.914
item13	-.011	.190

Raw Canonical Coefficients for Set-1

	1	2
item01	-.919	-.303
item07	-.243	-.083
item12	-.056	1.089
item13	-.013	.236

These indicate how much each item in Set 1 is weighted to create the linear combinations used in each of the two canonical correlations (canonical variates).

Standardized Canonical Coefficients for Set-2

	1	2
item03	-.972	-.418
item09	-.077	1.055

These indicate how much each item in Set 2 is weighted to create each of the two canonical variates for Set 2.

Raw Canonical Coefficients for Set-2

	1	2
item03	-1.070	-.459
item09	-.100	1.378

Canonical Loadings for Set-1

	1	2
item01	-.968	-.113
item07	-.665	.169
item12	-.297	.943
item13	-.286	.641

These indicate the correlation between each item in Set 1 and each of two canonical variates for Set 1.

Cross Loadings for Set-1

	1	2
item01	-.625	-.042
item07	-.429	.062
item12	-.192	.347
item13	-.184	.236

Canonical Loadings for Set-2

	1	2
item03	-.997	-.073
item09	-.395	.919

These indicate the correlation between each item in Set 2 and each of the two canonical variates for Set 2.

Chapter 10 – Multivariate Analysis of Variance

Cross Loadings for Set-2

	1	2
item03	-.644	-.027
item09	-.255	.338

Redundancy Analysis:

Proportion of Variance of Set-1 Explained by Its Own Can. Var.

	Prop Var
CV1-1	.387
CV1-2	.335

Proportion of Variance of Set-1 Explained by Opposite Can.Var.

	Prop Var
CV2-1	.161
CV2-2	.045

Proportion of Variance of Set-2 Explained by Its Own Can. Var.

	Prop Var
CV2-1	.575
CV2-2	.425

Proportion of Variance of Set-2 Explained by Opposite Can. Var.

	Prop Var
CV1-1	.240
CV1-2	.057

----- END MATRIX -----

The canonical scores have been written to the active file.
Also, a file containing an SPSS Scoring program has been written
To use this file GET a system file with the SAME variables
Which were used in the present analysis. Then use an INCLUDE command
to run the scoring program.
For example :

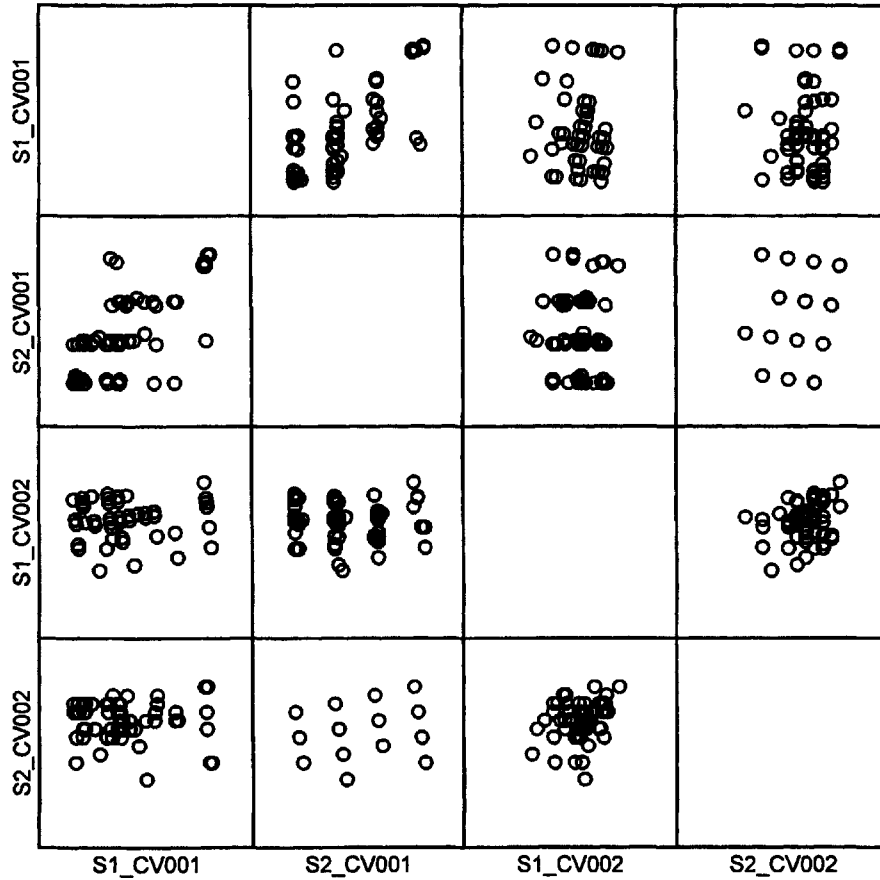
```
GET FILE anotherfilename
INCLUDE FILE "CC__.INC".
EXECUTE.
```

Output 10.4b: Matrix Scatterplot of Canonical Variates

```
GRAPH
  /SCATTERPLOT(MATRIX)=S1_CV001 S2_CV001 S1_CV002 S2_CV002
  /MISSING=LISTWISE .
```

Graph

The data appear to meet the assumptions of linearity, multivariate normality, and homoscedasticity because there does not appear to be a pattern in the scatterplots, and there are not large differences in how spread out each scatterplot is.



Interpretation of Output 10.4a and 10.4b

The first two matrices, **Correlations for Set-1** and **Correlations for Set-2**, are ordinary correlation tables. The first matrix is for the variables in Set 1, and the second matrix is for the variables in Set 2. We can see from these tables that most variables in each set are weakly to moderately correlated with each other.

The next matrix is the **Correlations Between Set-1 and Set-2**, which contains the bivariate correlations between each variable in Set 1 and each variable in Set 2.

The **Canonical Correlations** are the correlations between a linear combination of the Set 1 variables and a linear combination of the Set 2 variables (the canonical variates). Note that there are two different correlations, corresponding to two different pairs of linear combinations. By squaring and summing all canonical correlations, you can calculate a measure of R^2 indicating how much variance in one set of variables is explained by the other set of variables. Thus, $(.645^2 = .416) + (3.68^2 = .135) = .551$, so about 55% of the variance is shared.

One important part of the output to check is the **Test that remaining correlations are zero**. These chi-square goodness-of-fit tests indicate whether all correlations from that point on are significant, even after removing variance accounted for by all prior correlations.

The **Standardized Canonical Coefficients** can be interpreted much like regression weights, to show which items are weighted most heavily in the linear combination of variables for each set of variables. These weights are created so as to maximize the correlation between the two sets of variables. For example, the first canonical variate for Set 1 is created by weighting item01 most heavily (-.838), followed by item07 (-.259), item12 (-.047) and item13 (-.011). This canonical variate is then correlated with the canonical variate created by weighting item03 by -.972 and item09 by -.077 (see table of loadings in the output). The set of canonical coefficients in the second column is the set used for the second pair of canonical variates, which are correlated to produce the second canonical correlation. Thus, there are actually two linear combinations for each set of variables. These are derived, much like principal components analysis, by creating the linear combination for each set of variables that maximizes the correlation between the two sets of variables, then doing this again with what remains after the variance associated with the first correlation is removed, and so on. These coefficients are often called pattern coefficients.

Canonical Loadings are the correlations between each item and the linear combinations of variables for that same set (canonical variates). They are often called structure coefficients. These loadings aid in the interpretation of the correlations, much as the loadings in principal components analysis are interpreted to give meaning to the different components.

Cross Loadings for each set indicate the correlation between each item in one variable set and the canonical variate for the other variable set. In the **Cross Loadings for Set-2**, we can see that both items for Set 2 are negatively correlated with the Set 1 canonical variate for the first canonical correlation.

Example of How to Write About Problems 10.4

Results

Canonical correlation analysis was performed to assess the pattern of relationships between the motivation items and the competence items. The first canonical correlation was .65 (42% overlapping variance); the second was .37 (14% overlapping variance). With both canonical correlations included, $\chi^2(8) = 46.18, p < .001$, and with the first removed, $\chi^2(3) = 9.81, p = .02$. The correlations and canonical coefficients are included in Table 10.3. Examination of the loadings suggests that first canonical

correlation seems to involve a relation between practicing math a lot, independently, and becoming efficient (quick) at math; whereas, the second seems to capture a relation between persistence and thoroughness and being very competent at math.

Table 10.3

Correlation and Standardized Canonical Coefficients Between Motivation and Competence Variables

Item Content	First Canonical Correlation		Second Canonical Correlation	
	Loading	Coefficient	Loading	Coefficient
Motivation				
Practice till do well	-.97	-.84	-.11	-.28
Figure out math without help	-.67	-.26	.17	-.09
Complete math even if it takes a long time	-.30	-.05	.94	.91
Explore all solutions	-.29	-.01	.64	.19
Competence				
Solve math quickly	-1.00	-.97	-.07	-.42
Very competent in math	-.40	-.08	.92	1.06

Interpretation Questions

- 10.1 In Output 10.1b: a) Are the multivariate tests statistically significant? b) What does this mean? c) Which individual dependent variables are significant in the ANOVAs? d) How are the results similar and different from what we would have found if we had done three univariate one-way ANOVAs?
- 10.2 In Output 10.2: a) Are the multivariate tests statistically significant? What does this mean? b) If the interaction effect had been significant, what would that mean? c) For which individual/univariate dependent variables are the genders significantly different? d) Write up sentences summarizing the multivariate results, being sure to include the *F* values and degrees of freedom as these would be reported in a journal article.
- 10.3 In Output 10.3: a) What makes this a “doubly multivariate” design? b) What information is provided by the multivariate tests of significance that is not provided by the univariate tests? c) State in your own words what the interaction between time and group tells you. What implications does this have for understanding the success of the intervention?
- 10.4 In Output 10.4: a) What is the difference between canonical coefficients (pattern coefficients) and canonical loadings (structure coefficients)? What information do each of these sets of coefficients tell you? b) Give an example of a research problem that could be addressed using canonical correlation. c) What do the canonical correlations tell you? d) Why is there more than one canonical correlation?

Extra SPSS Problems

- 10.1 A company is interested in how consumers of different age groups like their DVD players. Open the dvdplayer.sav dataset and conduct a MANOVA using Age group and sex (male or female) as fixed factors and *price, ease, support, and func* as dependent variables. Request that descriptives,

Chapter 10 – Multivariate Analysis of Variance

measures of effect size, parameters, and tests of homogeneity are printed, and specify that you want a polynomial contrast on *agegroup*.

- a. How many participants of each gender are in each group? How might this affect results?
- b. Which results are significant, using Wilk's lambda?
- c. What are the eta squared values for each effect using Wilk's lambda? What do these mean?
- d. Were homogeneity assumptions met?
- e. What do the B values listed under Parameter Estimates (and their significance levels) tell you? What are the parameters?
- f. Which polynomial contrast(s) are significant for each dependent variable? Interpret these results.

- 10.2 Now, conduct the same analysis as in problem 1, but omit *sex* as an independent variable?
- a. Why might you choose to omit *sex* as a factor?
 - b. Was the effect of *agegroup* significant (using Wilk's lambda)? What does this mean?
 - c. Were the homogeneity assumptions met?
 - d. Which univariate effect(s) were significant? Interpret this.
 - e. Which group differences were significant, for which variable(s)?
 - f. Which polynomial contrast(s) were significant, for which variable(s)? Describe the results of this MANOVA, making sure to include information about the significance tests.
- 10.3 Open the World95.sav dataset, and conduct a canonical correlation, using *literacy*, *fertility*, *gdp_cap*, and *calories* as set 1 and *birth_rt*, *lifeexpf*, *death_rt*, and *aids_rt* as set 2. Be sure to check your syntax carefully!
- a. Which of the canonical correlations were significant? How do you know?
 - b. Which variables in set 1 were weighted most heavily for function 1? Which set 2 variables were weighted most heavily for function 1? Which set 1 and set 2 variables were weighted most for function 2?
 - c. Interpret the meaning of functions 1 and 2, based on the canonical loadings

APPENDIX A

Quick Reference Guide to SPSS Procedures

Joan Naden Clay

Introduction

The Quick Reference Guide is intended as a supplemental resource to use while the SPSS program is open. It provides a brief overview of commonly utilized procedures in SPSS as well as the statistical procedures covered in Morgan, Leech, Gloeckner, and Barrett (2004). This guide presents one basic method for accomplishing each command; however, there are often multiple ways to accomplish any given task in SPSS. The intent of this guide is to provide a brief description of each subject area in addition to giving step-by-step directions for performing specific actions for each subject.

Throughout the Quick Reference Guide, each subject area (e.g., variable, cell, file) is organized alphabetically and written in ***ALL CAPITAL, BOLD ITALICS***. Common actions (e.g., cut, paste, insert) are itemized for each subject and are indicated with an arrow. Stepwise instructions for functions are designated with a round bullet. Items within the stepwise directions that are bolded either represent window items that require you to make a selection or indicate that the user should select the item with the click of the mouse. Each step of a function is sequentially represented with arrows. For example, the instructions to ***Open a File*** would be: **File** (select the word “file” in the header) => **Open** (select the word “open” in the drop down window) => **Data** (select the word “data” in the subsequent drop-down menu) => using the **Files of Type** pull-down menu (click the arrow on the “Files of Type” window to get the drop-down menu), choose the type of data file to bring into SPSS=>locate the file=>double click on the file name.

SPSS Commands

ANALYSIS OF VARIANCE (ANOVA) is used to compare two or more independent groups on the dependent variable. Sample means are compared in order to infer whether the means of the corresponding population distributions also differ. ANOVA allows for several means to be compared.

- ***One-Way ANOVA*** (single factor analysis of variance) is used when you have one independent variable with a few nominal levels and one normally distributed dependent variable.
 - **Analyze => Compare Means => One-Way ANOVA** => select one or more dependent variables with a left click and move them into the **Dependent List** box by clicking the top arrow in the middle of the dialog box => select the independent variable with a left click and move it into the **Factor** (independent variable) box by clicking the bottom arrow in the middle of the dialog box => **Options** => choose **Descriptives** and **Homogeneity of variance test** => **Continue** => **OK**.
- ***Two-Way (Factorial) ANOVA*** is used when there are two different independent variables with a few nominal levels, a between groups design, and one normally distributed dependent variable.

- **Analyze => General Linear Model => Univariate** => select the dependent variable with a left click and move it into the **Dependent Variable** box by clicking the top arrow in the middle of the dialog box => select the independent variables with a left click and move them into the **Fixed Factor(s)** box by clicking the arrow in the middle of the dialog box => **Options** => select **Descriptive statistics, Estimates of effect size, and Homogeneity tests** => **Continue** => **Plots** => select one of the factors (independent variable) with a left click and move it to the **Horizontal Axis** box by clicking the top arrow in the middle of the dialog box => select the remaining factor (independent variable) with a left click and move it to the **Separate Lines** box by clicking on the middle arrow in the center of the dialog box => **Continue** => **OK**.

BAR CHARTS Bar charts are useful for displaying frequency information for categorical data, and SPSS allows several options for creating them. Types of bar charts available include: **simple** (presents a single bar for each category, case, or variable); **clustered** (presents a group of bars on the category axis, and bars within a cluster can represent groups of cases, separate variables, or individual cases); and **stacked** (there is one bar stack for each category, case, or variable on the category axis and segments of each stack can represent groups of cases, separate variables, or individual cases).

- **Graphs => Bar** => select type of chart (simple, clustered, or stacked) => choose if the **Data in Chart** are either summaries for groups of cases, summaries of separate variables, or values of individual cases => **Define** => at this point, the dialog box varies based on the type of chart and type of data to be displayed. Highlight variables and move them to the appropriate boxes (**bars represent, category axis, define clusters by, define stacks by, etc.**) by clicking the arrows. For summaries of groups of cases, choose if each **Bar Represents** the number of cases, cumulative cases, percent of cases, cumulative percent of cases, or other summary function => **Titles** => type in title, subtitles, and footnotes => **Continue** => **OK**.

BOX PLOTS This type of graphic provides a visual display of the distribution of data by presenting the median, quartiles, and extreme values for the category or variable. Box plots can be **simple** (with a single box plot for each category or variable) or **clustered** [with 2 or more plots for each category or variable on the category axis, and each plot within a cluster being defined by another definition (grouping) variable]. Box plots are automatically generated by the **Explore** command or can be custom generated with the **Graphs** command.

- **Graphs => Boxplot** => select type of plot (simple or clustered) => choose if the **Data in Chart** are either summaries for groups of cases or summaries of separate variables => **Define** => at this point, the dialog box varies based on the type of chart and type of data to be presented. Highlight variables and move them to the appropriate boxes (variable, category axis, define clusters by, label cases by, or boxes represent) => **OK**.

CASES are the individual records for each subject and are organized by rows for each record/case. SPSS numbers each case down the left side of the Data View window. All of the data for each participant should be entered as one row of information.

- **Insert Cases** This command allows data to be added for new subjects or cases anywhere in the data set.

- Click on the case number below where the case is to be inserted => **Data => Insert Cases**
=> enter data for the new case.
- **List Cases (Case Summaries)** This command allows the user to list either an entire dataset or a subset of the data. Case summaries are especially helpful for ensuring that data are computed or entered correctly. Various options allow the user to select one or all variables, create an order for listing the cases using a grouping variable, limit the number of cases shown, and conduct basic descriptive statistics on the cases selected.
 - **Analyze => Reports => Case Summaries** => select variables to be summarized with a left click => click the top arrow in the middle of the dialog box to move the selected variable to the **Variables** box => if desired, select categorical variables with a left click => click the bottom arrow in the middle of the dialog box to move the selected variable to the **Grouping Variables** box => utilize check boxes in lower left corner to display individual cases, limit number of cases shown, show valid cases or show case numbers => **Statistics** => highlight desired statistic with a left click => click the arrow in the middle of the dialog to move the selected statistic to the **Cell Statistics** box => **Continue** => **OK**.
- **Select Cases** The select cases command permits the analysis of a specific subset of the data. Once a subset is selected, the user can either revert back to the entire dataset or delete the unselected cases to create a subset data file. Users can **Select** data in a variety of ways including: **If condition is satisfied** (a conditional expression is used to select cases); **Random sample of cases** (cases are selected randomly based upon a percent or number of cases); **Based on time or case range** (case selection is based upon a range of case numbers or a range of dates/time); or, **Use filter variable** (a numeric variable can be used as the filter – any cases with a value other than 0 or missing are selected). **Unselected cases** may be **Filtered** (remain in data file, but excluded in the analysis) or **Deleted** (removed from the working data file and cannot be recovered if the data file is saved after the deletion). If you want to do the same analysis separately on all subsets of data, then **Split File** should be used instead of **Select Cases** (see below)
 - **Data => Select Cases => Select** (choose method: all cases, if condition is satisfied, random sample of cases, range, filter variable) => **Unselected cases are** (choose either filtered or deleted) => **OK** (Save your work before deleting cases, just in case you change your mind!).
- **Sort Cases** This procedure allows the user to rearrange the order in which data are displayed using one or more variables of interest. Data can be sorted in **ascending** (small to large numbers or alphabetical) or **descending** (large to small numbers or reverse alphabetical) order.
 - **Data => Sort Cases** => select a variable with a left click => click the arrow in the middle of the dialog box to move the variable to the **Sort By list** => choose either the **Ascending** or **Descending** display => **OK**.

CELLS A cell is a single data point where a row and column intersect within the dataset. Cells can be manipulated individually or in blocks. Data can be edited using **Copy** (puts a duplicate of the cell value(s) on the clipboard, but leaves original cells in place), **Cut** (puts the cell value(s) on the clipboard and deletes the original cell values), or **Paste** (inserts cell value(s) from the clipboard into cell(s) at a new location).

Appendix A – Quick Reference Guide

- **Copy and Paste Cells** This command makes a copy of data in a cell while leaving the original cell in place. The copy can be inserted in another location.
 - Highlight the data (cell, row, column, etc.) to copy => **Edit => Copy** => click on the individual cell or the upper left cell of the group of cells to paste to (If data are pasted over already existing cells, the data in those cells will be erased! It may be necessary to create space by inserting variables or cases. Pay close attention to cell alignment.) => **Edit => Paste** (Save your work or experiment on a copy of your data—just in case the unexpected happens!).
- **Cut and Paste Cells** This command removes the data from a cell so you can insert it in a different location.
 - Highlight data (cell, row, column, etc.) that you want to cut => **Edit => Cut** => click on the individual cell or the upper left cell of the group of cells you want to paste to (If you paste data over already existing cells, the data in those cells will be erased! You may need to create space by inserting variables or cases. Pay close attention to your alignment.) => **Edit => Paste**.

CHART EDITOR Existing charts may be edited using the SPSS chart editor. Common actions include adding titles, changing labels, altering color and fill patterns, and adjusting axis intervals.

- Double click on a chart in the output (chart appears in the **Chart Editor** window) => double click on the element of the chart to edit (**Palettes** window opens) => select appropriate tab of the palettes window => utilize various dialog boxes, check boxes, etc. to edit the element of interest => **Apply** (repeat process on all the various elements of the chart to edit).

CHI-SQUARE allows the user to determine whether there is a statistically significant relationship between two nominal variables. This test compares the observed and expected frequencies in each category to test whether all the categories contain the same proportion of values. The chi-square test does not indicate the strength (effect size) of a statistically significant relationship. The optional Phi and Cramer's V tests can be used as a measure of effect size.

- **Analyze => Descriptive Statistics => Crosstabs** => select the first nominal variable with a left click and move it to the **Row(s)** box by clicking the top arrow in the middle of the dialog box => select the second nominal variable with a left click and move it to the **Column(s)** box by clicking the arrow in the middle of the dialog box => **Statistics** => check **Chi-square** and **Phi and Cramer's V** => **Continue** => **Cells** => check **Observed, Expected and Total** => **Continue** => **OK**.

CODEBOOK This feature allows information for all of the variables in the dataset to be printed including: variable names, measurement level, column width, number of decimal places, values, and value labels.

- **Utilities => File Info** (the codebook is printed into the output).

COHEN'S KAPPA is used to assess interobserver reliability (the agreement between the evaluations of two raters when both are rating the same object). Kappa is used for situations

where both variables use the same category values and both variables have the same number of categories.

- **Analyze => Descriptive Statistics => Crosstabs** => select the variable representing the first rater's observation with a left click and move it to the **Row(s)** box by clicking the top arrow in the middle of the dialog box => select the variable representing the second rater's observation with a left click and move it to the **Column(s)** box by clicking the center arrow in the middle of the dialog box => **Statistics** => check box for **Kappa** => **Continue** => **Cells** => check **Observed** in the **Counts** box => check **Total** in the **Percentages** box => **Continue** => **OK**.

COPY DATA PROPERTIES WIZARD provides the ability to use an external SPSS data file as a template for defining file and variable properties for a working data file. This process replaces the Apply Data Dictionary function, which was formerly located on the File menu.

- **Data => Copy Data Properties** => identify the source file for the data properties => **Next** => choose source and target variables => **Next** => choose variable properties to copy => **Next** => choose dataset properties to copy => **Next** => **Finish**.

COUNT creates a new variable that counts the occurrences of a given value in a list of variables for each case. This is often used to devise a variable indicating how many responses a participant checked (e.g., when asked to check which of the following were reasons, the person selected a particular brand or product) or how many times a person said, "yes" to a series of related questions (e.g., questions asking if the participant experienced particular risk factors in development).

- **Transform => Count** => type the new variable name in the **Target Variable** box => type a description for the variable in the **Target Label** box => select the variables you want to be included in the count with a left click and move to the **Variables** box by clicking the arrow in the middle of the dialog box => **Define Values** => determine the value or range of values to be counted using the radio buttons and dialog boxes in the **Value** box => **Add** => **Continue** => **OK**.

CORRELATIONS are inferential statistics that are used to assess the association or relationship between two variables.

- **Pearson Product Moment Correlation (r)** is a bivariate parametric statistic, also known as Pearson correlation coefficient, used when both variables are approximately normally distributed.
- **Analyze => Correlate => Bivariate** => select the scale variables to be correlated with a left click and move them to the **Variables** box by clicking the arrow in the middle of the dialog box => check the **Pearson** box => select the **two-tailed** radio button => check the **Flag significant correlations** box => **Options** => check **Means and standard deviations** box => select the **Exclude cases listwise** radio button => **Continue** => **OK**.
- **Spearman Rho (r_s)** is the nonparametric equivalent of the Pearson correlation coefficient. This statistic should be selected when the data are ordinal or when the assumptions for the Pearson correlation coefficient are markedly violated.

Appendix A – Quick Reference Guide

- **Analyze => Correlate => Bivariate =>** select the variables to be correlated with a left click and move them to the **Variables** box by clicking the arrow in the middle of the dialog box => check the **Spearman** box => select the **two-tailed** radio button => check the **Flag significant correlations** box => **Options** => check **Means and standard deviations** box => select the **Exclude cases listwise** radio button => **Continue** => **OK**.

CRONBACH'S ALPHA is a commonly used procedure to assess the internal consistency reliability of several items or scores that are going to be used to create a summated scale score.

- **Analyze => Scale => Reliability Analysis =>** select variables (items) to be included in the summated scale with a left click and move to the **Items** box by clicking the arrow in the middle of the dialog box => use the drop-down arrow to designate the **Model** as **Alpha** => **Statistics** => check the **Correlations** box => **Continue** => **OK**.

CROSSTABULATION see **CHI-SQUARE**.

DATA are the values entered into the cells that are created by the intersection of the rows (cases) and columns (variables).

- **Copy Data** (see **CELLS** – copy and paste or cut and paste).
- **Enter (Edit) Data** Values can be entered into blank cells, or existing values may be changed.
 - Left click the cell of interest => type (or edit) value => **Enter**.

Export Data Datasets can be saved in other file formats for use in other applications (e.g. other versions of SPSS, Excel, dBASE, SAS).

- **File => Save As** => click the drop-down arrow in the **Save as type** dialog box => select the file type => type the file name in the **File name** dialog box => click the drop down arrow in the **Save in** dialog box => select a drive or folder for the file => **Save**.
- **Import Data** This command copies data from a word document and pastes it into SPSS.
 - In the word document, highlight the data to be pasted into SPSS => in the Data View of SPSS, set the mouse on the first cell to receive the incoming data => **Edit** => **Paste**.
 - Alternatively, in the word document, highlight the data to be pasted into SPSS => in the Data View of SPSS, set the mouse on the first cell to receive the incoming data => right click on **Paste** to import data from an excel, dbase, or other file, use **Open Data** file (below).
- **Open Data** See **FILES** (Open Data File).
- **Print Data** This command allows some or all of the data to be printed.
 - **Print All Data** With the database open go to Data View => **File** => **Print** => in the **Print Range** dialog box use the radio button to select **All** => **OK**.
 - **Print Selected Data** With the database open go to Data View => highlight the cells to print => **File** => **Print** => in the **Print Range** dialog box use the radio button to select **Selection** => **OK**.

➤ **Save Data** This command should be used often to avoid losing work!

- **Initially Saving a New Data Set** **File** => **Save** => (**Save Data As** dialog box appears) select the appropriate drive and folder using the **Save in** drop down menu => type the file name in the **File Name** dialog box => select SPSS (*.sav) in the **Save as type** dialog box => **Save**.
- **Resaving an Existing Data Set Under the Same File Name** **File** => **Save** (SPSS automatically resaves the data using the existing file name).
- **Saving Existing Data Set Under a Different File Name** **File** => **Save As** => select the appropriate drive and folder using the **Save in** drop-down menu => type the file name in the **File Name** dialog box => select SPSS (*.sav) in the **Save as type** dialog box => **Save**.

DATABASE INFORMATION DISPLAY provides dictionary information about the selected data file, including variable names, variable formats, and descriptive variable and value labels.

- **File** => **Display Data Info** => use the drop-down arrow for the **Look in** box to select the appropriate drive => left click on the file name => **Open**.

DESCRIPTIVE STATISTICS displays summary statistics for several variables in a single table and can also calculate standardized values (z scores).

- **Analyze** => **Descriptive Statistics** => **Descriptives** => select variables with a left click and move them to the **Variable(s)** box by clicking the arrow in the middle of the dialog box => **Options** => check boxes for desired descriptives => select preference for **Display Order** => **Continue** => **OK**.

EXPLORE DATA This command produces summary statistics and graphical displays for entire datasets or selected subsets of data. This command is useful for screening data, generating descriptive information, checking assumptions and looking for differences among groups of cases.

- **Analyze** => **Descriptive Statistics** => **Explore** => highlight one or more dependent variables to explore and move them to the **Dependent List** box by clicking the arrow => if desired, move one or more categorical variables to the **Factor List** box in order to obtain separate analyses for groups of cases based upon the factor => **Statistics** => check desired statistics => **Continue** => **Plots** => select types of plots desired (box plots, stem-and-leaf, histogram, normality plots) => **Continue** => **OK**.

FILES There are three different common file types in SPSS (data, output, and syntax) that are differentiated with unique file extensions (data = .sav; output = .spo; syntax = .sps). Script is a fourth type of SPSS file and is designated with a .sbs extension; however, discussion of script files is beyond the scope of this appendix. Each of the different file types will open in its own type of window. Data files are presented in the **SPSS Data Editor** window, output files are displayed in the **SPSS Viewer**, and syntax is viewed in the **SPSS Syntax Editor**. (Only one data file can be open in any one session.)

Appendix A – Quick Reference Guide

- **Create a New File** This command will open a blank file. Remember to save files frequently as you generate data, output, and syntax!
- **New Data File (.sav)** File => New => Data => see chapter 2 for information about entering data. (Remember to save your work frequently!)
 - **New Output File (.spo)** File => New => Output => this provides a blank SPSS Viewer. (SPSS will automatically open the SPSS Viewer when the first analysis is run; however, more than one SPSS viewer/output file can be open at any given time.)
 - **New Syntax File (.sps)** File => New => Syntax => this provides a blank syntax editor. Syntax can be typed directly into the editor or pasted in from a word processing program (e.g., Word). Information about writing syntax can be found in the SPSS Syntax Reference Guide in the Help Menu – Help => Syntax Guide => Base => this opens an Adobe portable document format (PDF) file that can be printed or viewed online (Adobe Reader software is required to open this file and is available free online at www.adobe.com).
- **Merge Files** allows the working data file to be combined with a second data file that contains the same variables but different cases or the same cases but different variables. Before you add variables to a file using this method, you should first make sure that each participant is identified by the same ID number in both files. Then, you should use **Sort Cases** to sort each file (by itself), sequentially in ascending order, saving each file once it is sorted. Then, you should make sure that you open the data file first that has the correct values of any variables that exist in both data sets. SPSS will save only one copy of each variable, and that will be the one that is in the first (working data) file. Now, you are ready to merge files.
- Open the first data file => Data => Merge Files => Add Cases (or Add Variables) => use the drop-down arrow for the **Look in** box to select the appropriate drive => select the data file to merge with the open file with a left click => Open => remove any variables you do not want from the **Variables in the New Working Data File** list. If you are adding variables, you should make sure that they are being added to the correct participants by checking **Match cases on key variables in sorted files**. (In order to do this, both files must have data that are arranged sequentially by the key variable, usually ID, as indicated above) move ID or other key variable to key variables box => OK.
- **Open File** This command helps locate and open existing files (data, output, syntax). The dialog box provides a pull-down menu that allows for the selection of the type of data file to import/open.
- **Open Data File** File => Open => Data => use the **Look In** drop-down menu to select the appropriate directory/folder => using the **Files of Type** pull-down menu, choose the type of data file to bring into SPSS. Common file types and file name extensions include SPSS (.sav), Excel (.xls), and dBase Files (.dbf). In searching for the data file to open, only those files of the specified type will be shown in the dialog box => locate the desired file, select the file name with a left click => Open. (Only one data file can be open in any given session.)
 - **Open Output File** File => Open => Output => use the **Look In** drop-down menu to select the appropriate directory/folder => using the **Files of Type** pull-down menu, choose .spo => when you locate the file you want, select the file name with a left click => Open.

- **Open Syntax File** File => Open => Syntax => use the **Look In** drop-down menu to select the appropriate directory/folder => using the **Files of Type** pull-down menu, choose **.sps** => locate the desired file, select the file name with a left click => **Open**.
- **Save File** This command should be performed frequently to prevent unwanted loss of work! The first time a file is saved, the **Save As** dialog box will appear to allow the selection of a directory/folder and allow the user to name the file. If a file is saved that is already named, the program copies over the existing file. The directions presented below represent saving a file for the first time.
 - **Save Data File** See **Save Data** (above)
 - **Save Output File** In the **Output - SPSS Viewer** => **File** => **Save** => use the **Save in** drop-down menu to select the appropriate directory/folder => using the **Save as type** drop-down menu, choose **Viewer Files (*.spo)** => type in a name in the **File name** dialog box => **Save**.
 - **Save Syntax File** In the **Syntax – SPSS Syntax Editor** => **File** => **Save** => use the **Save in** drop-down menu to select the appropriate directory/folder => using the **Save as type** drop-down menu, choose **SPSS Syntax Files (*.sps)** => type in a name in the **File name** dialog box => **Save**.
- **Split File** splits the data file into separate groups for analysis based upon the values of one or more grouping variables. Data can be displayed for group comparisons, or data can be displayed separately for each group. This is a very useful tool, but be sure to reset **Split File** after doing the analyses you want split, or all further analyses will be split in the same way.
 - **Data** => **Split File** => select appropriate radio button for desired display option (**Compare groups** or **Organize output by groups**) => select grouping variable(s) by left clicking and move to them to the **Groups Based on** box by clicking the arrow in the center of the dialog box => **OK**.

FISHER'S EXACT TEST This test is included in the chi-square output and should be used to interpret results for small (2x2) crosstabulations (see **CHI-SQUARE**).

FREQUENCIES provides statistics and graphical displays that are useful for obtaining summaries of individual variables. The frequencies command can be used to determine typical values of the variables, check assumptions for statistical procedures and check the quality of the data.

- **Analyze** => **Descriptive Statistics** => **Frequencies** => select variables with a left click and move to the **Variable(s)** box by clicking the arrow in the middle of the dialog box => check box for **Display frequency tables** as desired => **Statistics** => check boxes for desired statistics => **Continue** => **Charts** => select **Chart Type** as desired => select type of **Chart Values** desired => **Continue** => **OK**.

HELP SPSS provides online help in a variety of ways including the help menu, dialog box help, and tutorials. Every window has a Help menu on the menu bar.

- **Help Menu** This command is accessed from the menu bar in any window.

Appendix A – Quick Reference Guide

- **Help => Topics** => click the **Index** tab => type in a keyword or simply scroll down through the topic list => double click on the topic of interest => information and links to related topics will be displayed.
 - **Help => Topics** => click the **Search** tab => type in a word or phrase to search for => click **List Topics** => browse results using the scroll bar => double click on the topic of interest => information and links to related topics will be displayed.
- **Dialog Box Help Button** Dialog boxes offer a context-sensitive help button that brings up a standard help window that contains information on the current dialog box.
- **Dialog Box Quick Help** Right click on any control in a dialog box and a brief description of the item will appear.

HISTOGRAM displays the distribution of a quantitative variable in a graphical format. A normal curve can be displayed on the histogram.

- **Graphs => Histogram** => select the variable with a left click => move it to the **Variable** box by clicking the arrow in the middle of the dialog box => check box for **Display normal curve** if desired => **Titles** => type in titles, subtitles, and footnotes if desired => **Continue** => **OK**.

INTERACTIVE CHARTS/GRAPHS provide special editing features that are not available in charts generated from statistical procedures. The chart editor allows customization of charts for displaying data. Some of the features of the chart editor include: direct text editing, moving and resizing chart frames, changing styles of chart elements, and adding chart elements.

- **Graphs => Interactive** => select type of chart/graph desired => create chart by working through the dialog boxes under the various tabs for creating the chart.

KENDALL'S TAU-B is a crosstabs strength of association statistic that can be selected if the variables are ordinal (see **CHI-SQUARE**).

KRUSKAL-WALLIS H TEST is the nonparametric equivalent of a one-way analysis of variance and tests whether several independent samples are from the same population. This test should be selected instead of a one-way ANOVA if the data are ordinal or if the homogeneity of variance assumption is seriously violated and group sizes differ.

- **Analyze => Nonparametric Tests => K Independent Samples** => select the dependent variable(s) with a left click and move them to the **Test Variable List** box by clicking the top arrow in the middle of the dialog box => select the independent variable with a left click and move it to the **Grouping Variable** box by clicking the bottom arrow in the middle of the dialog box => **Define Range** => type in the **Minimum** and **Maximum** values of the independent variable => **Continue** => check **Kruskal-Wallis H** in the **Test Type** box => **OK**.

LAYERS can be used to add depth to tables by creating three-dimensional cubes. Layers nest or stack information; however, only one layer for a given category is visible at a time. A single variable can be used to create stacked layers (one layer per category of the variable). A combination of variables can create nested (crosstabulated) layers. The total number of layers is the product of the number of categories for each layer variable.

- **Analyze => Tables => Custom Tables =>** highlight variables with a left click and drag them to the row or column areas of the canvas pane => utilize dialog boxes under the various tabs to customize the table => **OK**.

LINE CHARTS allows the user to make selections to create a variety of line charts. Line charts can be created to summarize one or more separate variables, individual cases, or groups of cases.

- **Graphs => Line =>** select the type of chart desired (**simple, multiple, drop-line**) => make a selection in the **Data in Chart Are** dialog box (**Summaries for groups of cases, Summaries of separate variables, Values of individual cases**) => **Define** => the **Define** dialog boxes vary based upon the type of chart and type of data, make selections for dialog boxes => **Titles** => type in **Title, Subtitle, and Footnote** text as desired => **Continue** => **OK**.

LINEAR REGRESSION estimates the coefficients of the linear equation by using one or more independent variables to predict the value of the dependent variable.

- **Bivariate (Simple Linear) Regression** is used to predict scores of a normal/scale dependent variable from one normal/scale independent variable.
 - **Analyze => Regression => Linear =>** select the dependent variable with a left click and move to the **Dependent** box by clicking the top arrow in the middle of the dialog box => select the independent variable with a left click and move to the **Independent(s)** box => use the drop-down arrow to select **Enter** as the **Method** => **OK**.
- **Multiple Regression** is used to predict scores of a normal/scale dependent variable from a combination of several interval, scale, and/or dichotomous independent (predictor) variables.
 - **Analyze => Regression => Linear =>** select the dependent variable with a left click and move to the **Dependent** box by clicking the top arrow in the middle of the dialog box => select the independent variables with a left click and move to the **Independent(s)** box => use the drop-down arrow to select **Enter** as the **Method** => **Statistics** => check **Estimates** in the **Regression Coefficients** box => check **Model fit** and **Descriptives** => **Continue** => **OK**.

MANN-WHITNEY U TEST is a nonparametric test similar to the independent samples *t* test, that tests whether or not two sampled populations are equivalent. This test should be used if the dependent variable is ordinal or if the assumptions for the independent samples *t* test are markedly violated.

- **Analyze => Nonparametric Tests => 2 Independent Samples =>** select dependent variable(s) with a left click and move to the **Test Variable List** box by clicking the top arrow in the middle of the dialog box => select the independent variable with a left click and move it to the **Grouping Variable** box by clicking the bottom arrow in the middle of the dialog box => check the **Mann-Whitney U** box in the **Test Type** box => **OK**.

MEANS This procedure calculates subgroup means and other descriptive statistics for dependent variables within categories of one or more independent variables. Options within the means procedure include running a one-way analysis of variance as well as tests for linearity.

- **Analyze => Compare Means => Means** => select one or more dependent variables with a left click and move them to the **Dependent List** box by clicking the top arrow in the middle of the dialog box => select one or more independent variables with a left click and move them to the **Independent List** box by clicking the bottom arrow in the middle of the dialog box => **Options** => select desired statistics with a left click and move to the **Cell Statistics** box by clicking the arrow in the middle of the dialog box => if desired, check boxes for **Anova table and eta** or **Test for linearity** in the **Statistics for First Layer** box => **Continue** => **OK**.

MISSING VALUES SPSS defines missing values as “system-missing” versus “user-missing.” System-missing values are omissions in the dataset and are designated with a period in the cell. A user-missing value is an actual value that the user defines to indicate that a value is known to be missing. Up to three different discrete values or a range of values can be coded to define reasons for missing data.

- **Define User-Missing Values** This procedure allows the user to specify data values that indicate why information is missing.
 - Select the **Data Editor** as the active window => click the **Variable View** tab => click on the cell in the **Missing** column that corresponds to the variable of interest => click on the shaded box that appears in the right of the cell to get the **Missing Values** dialog box => use the radio buttons to select either **discrete missing values** or **range plus one optional discrete missing value** => type in the values to be assigned in the appropriate boxes => **OK**.
- **Replace Missing Values** Missing values may make datasets difficult to work with and may hinder certain types of analyses. For these reasons, it may be appropriate to replace missing values using techniques such as interpolation, multiple imputation, inserting a mean or median of nearby values, etc.
 - **Transform => Replace Missing Values** => highlight the variable(s) of interest and move them to the **New Variable(s)** box by clicking the arrow in the center of the dialog box => highlight one of the variables in the **New Variable(s)** box => the variable name will appear in the **Name** box => click on the down arrow for the **Method** box to display the techniques for replacing the missing value => if either **mean of nearby points** or **median of nearby points** is chosen, the span of points to be used must be designated => **OK**.

OUTPUT This term refers to the display of results of analyses, as well as of commands if display of these was requested. Output is generated into its own window and can be saved as a separate file that has a .spo filename extension.

- **Copy and Paste Output** This allows the user to move output to another place while leaving the original where it is.
 - Left click on the output item of interest (or highlight the item in the outline view) to select the item => right click to get the drop-down menu => **Copy** => left click to place the cursor over the area to copy to => right click => **Paste**.

➤ **Delete Output**

- Left click on the output to select => right click to **Delete**.

➤ **Display Syntax (Command Log) in the Output** This procedure will write the syntax for everything a user does in SPSS and will post the syntax in the output window. The user can then paste that syntax into a new syntax file enabling the user to repeat procedures without having to use a point-and-click approach. This can help to ensure consistency across analyses and eliminate errors. It is also useful because it shows which file is used for each analysis so that results in the output can be easily identified.

- **Edit => Options => Viewer => check the box to Display Commands in the Log => Apply => OK.**

➤ **Edit Output** Charts, tables, and text can all be edited in a variety of ways. Double clicking on any item will activate chart editors, pivot tables, and text boxes to allow editing.

➤ **Export Output (to Word, HTML, etc)**

- **File =>Export => click the drop-down arrow under file type to indicate format => type in new file name => click on OK**
- Or you can **Save as**, as in the instructions above for data files, if you want to save as other versions of SPSS, etc.
- To save a portion of the output to an existing Word file, left click to highlight table or graphic to copy => right click to **Copy** (NOT copy object) => proceed to the word document => **Edit => Paste => OK.**

➤ **Format Output** SPSS users may adjust the size, shape, color, or other features of any displayed output.

- Double click on the portion of the output to format => double click on the specific output cell to format => right click and select the appropriate formatting option. To edit the text of an output cell, instead of right clicking just double click again to begin typing.

➤ **Hide Results Within an Output File** Results in the output can be hidden if the user does not want them displayed or printed but still wants them in the file. Conversely, results previously hidden can be shown again.

- In the SPSS Viewer, go to the outline view on the left side of the screen => scroll down to the output item to hide => click on the minus sign (-) to the left of the output name to hide the item (a plus sign will appear to indicate the item has been collapsed and is hidden) => click on the plus sign (+) to unhide the output.

➤ **Insert Text/Title to Output** This command allows the user to insert descriptive titles or detailed notes directly into the output.

- In the output or in the output outline, click where the text will be inserted => **Insert => NewText (or NewTitle) => type desired text => right click.**

Appendix A – Quick Reference Guide

➤ **Open Output** See **FILES** (Open Output File).

➤ **Print Output**

- **File => Print** => choose to print **All Visible Output** or just a **Selection** (when printing a selection, one or more output items in the outline must be highlighted before accessing the print command).

➤ **Resize/Rescale Output** This allows larger tables to be fit onto a single printed page.

- Double click on the table to be resized to enter editing mode => right click => **Table Properties** => select the **Printing** tab=> check to **Rescale Wide Table to Fit Page** and/or **Rescale Long Table to Fit Page** => **OK**.

➤ **Save Output** See **FILES** (Save Output File).

PIVOT TABLES When using the SPSS viewer, this feature can be used to quickly rearrange rows, columns, and layers in output tables in order to present results more clearly.

- Double click on the table to be formatted to enter editing mode=> right click => **Pivoting Trays** => left click on the icon of the row, column, or layer to be moved => drag the icon to the desired row, column, or layer and release.

PRINT PREVIEW This feature allows the user to see what will print on each page of output. Items that may not be visible in the viewer, but can be seen in print preview, include: page breaks, hidden layers of pivot tables, breaks in wide tables, headers, and footers.

- Highlight all or some of the output in the output outline => **File** => **Print Preview** => use buttons at the top of the print preview window to navigate the document, print, close, etc.

RESTRUCTURE DATA WIZARD helps the user restructure or transpose all or some cases into variables or vice versa.

- **Data** => **Restructure** => follow the steps in the wizard to complete the desired restructuring tasks.

RESULTS COACH explains how to interpret specific output.

- Double click on the output table or graphic => right click and select **Results Coach** => use arrows at the bottom right of the screen to page through the explanation of the output.

SCATTERPLOT graphs the correspondence between two variables to determine how individuals' scores on one variable are associated with their scores on another variable. The scatterplot graph results in a visual picture of the correlation between the two variables and provides information about bivariate outliers.

- **Graphs** => **Scatter** => click on **Simple** => **Define** => select the first variable with a left click and move it to the **Y axis** box by clicking the top arrow in the middle of the dialog box => select the second variable with a left click and move it to the **X Axis** box by

clicking the second arrow in the middle of the dialog box => **Titles** => type text for the **Title**, **Subtitle**, and **Footnote** as desired => **Continue** => **OK**.

SPEARMAN RHO See **CORRELATIONS**.

STATISTICS COACH This feature prompts the novice user with questions and examples to help in the selection of basic statistical and charting features that are most appropriate for the data. The Statistics Coach only covers the basic, most commonly used statistical procedures.

- **Help** => **Statistics Coach** => complete the wizard to step through the coaching process.

STEM AND LEAF PLOT gives a visual impression of the distribution of scores for a variable. This type of plot looks somewhat like a histogram or frequency distribution turned on its side, but it includes the individual numbers comprising the data. It enables one to display the distribution of scores for each of the levels of a grouping variable, if desired.

- **Analyze** => **Descriptive Statistics** => **Explore** => select a dependent variable with a left click and move it to the **Dependent List** box by clicking the top arrow in the middle of the dialog box => if desired, select a categorical independent variable with a left click and move it to the **Factor List** box by clicking the center arrow in the middle of the dialog box => use the radio button to select **Both** in the **Display** box => **OK**.

SUMMARIZE This procedure produces report tables containing summary statistics and/or individual case listings for one or more scale variables. A wide variety of statistics are available to describe the data. Statistics and case displays can be calculated within levels of a categorical variable and may include all observations or a limited number of cases.

- **Analyze** => **Reports** => **Case Summaries** => select variables to summarize with a left click and move to the **Variables** box by clicking the top arrow in the middle of the dialog box => if desired, select a grouping variable with a left click and move it to the **Grouping Variable(s)** box by clicking the bottom arrow in the middle of the dialog box => select **Display cases** if individual case listings are desired => **Statistics** => select statistics desired with a left click and move them to the **Cell Statistics** box by clicking on the arrow in the middle of the dialog box => **Continue** => **Options** => type in a **Title** and a **Caption** => **Continue** => **OK**.

SYNTAX SPSS provides a command language that can be used instead of the point-and-click (windows) method to run SPSS. A syntax file is simply a text file that contains commands that can be used to repeat an analysis at a later time. Use of syntax files can help to ensure consistency across analyses and eliminate errors.

- **Create syntax** See **FILES** (Create a New File – New Syntax File). One can also click on **Paste** within any statistics program to produce a file containing the syntax for that procedure.
- **Run syntax** The user can choose to run some or all of the commands in the syntax file.
 - **All Syntax** With the appropriate syntax file as the active window => **Run** => **All**.
 - **Selected Syntax** With the appropriate syntax file as the active window => highlight the portion of syntax to be run => **Run** => **Selection**.

➤ **Print Syntax**

- With the appropriate syntax file as the active window => **File** => **Print** => select printing options => **OK**.

➤ **Save Syntax** See **FILES** (Save File – Save Syntax File).

TABLES SPSS allows the user to build a variety of tables using the custom, basic, and general tables commands. Only the basic tables command will be presented.

- **Analyze** => **Tables** => **Basic Tables** => select one or more variables to be summarized within the cells and move to the **Summaries** dialog box using the arrow => select categorical variables to create row subgroups and move to the **Subgroups – Down** dialog box using the arrow => select categorical variables to create column subgroups and move to the **Subgroups – Across** dialog box using the arrow => select categorical variables whose values create separate tables and move to the **Subgroups – Separate Tables** dialog box using the arrow => **Statistics** => choose options for presentation of statistics => **Continue** => **Layout** => select layout options => **Totals** => select totals options => **Continue** => **Titles** => type in titles as desired => **Continue** => **OK**.

T TEST The *t* test is a procedure that is used to compare sample means to determine if there is evidence that the corresponding population distributions differ.

➤ **Independent Samples *t* test** is used to compare two independent or unrelated groups (between groups design) on an approximately normal dependent variable.

- **Analyze** => **Compare Means** => **Independent Samples T Test** => select one or more dependent variables with a left click and move them to the **Test Variable(s)** box by clicking the top arrow in the middle of the dialog box => select a dichotomous independent variable with a left click and move it to the **Grouping Variable** box by clicking the bottom arrow in the middle of the dialog box => **Define Groups** => select **Use specified values** => type the values of the independent variable that will designate **Group 1** and **Group 2** => **Continue** => **OK**.

➤ **One-Sample *t* test** allows the mean of a sample to be compared to a hypothesized population mean.

- **Analyze** => **Compare Means** => **One-Sample T Test** => select the dependent variable with a left click and move it to the **Test Variable(s)** box by clicking the arrow in the middle of the dialog box => type the hypothesized population mean in the **Test Value** box => **OK**.

➤ **Paired Sample *t* test** is used when the two scores being compared are paired or matched in some way (not independent of one another) or if the two scores are repeated measures.

- **Analyze** => **Compare Means** => **Paired Samples T Test** => select the two variables that make up the pair and move them simultaneously to the **Paired Variable** box by clicking the arrow in the middle of the dialog box => **OK**.

VARIABLES These are characteristics of the cases (e.g., participants) in the study that are able to vary or have different values.

- **Compute Variable** This procedure allows the user to create new variables from existing variables by either using numeric expressions to mathematically generate the new values or by using an “If” statement to conditionally select values to compute into a new variable.
 - **Numeric Expression Procedure:** Transform => Compute => type name of the new variable in the **Target Variable** box => click on **Type & Label** => type in label to describe the new variable in the **Label** box => choose whether the variable is numeric or string in the **Type** section => **Continue** => create a numeric expression by highlighting variables and moving them to the **Numeric Expression** box with the arrow in conjunction with the calculator pad to create mathematical expressions (or alternatively highlight a preset mathematical function from the **Functions** box and move it up to the **Numeric Expression** box with the arrow) => **OK**.
 - **“IF” Procedure:** Transform => Compute => type name of the new variable in the **Target Variable** box => click on **Type & Label** => type in label to describe the new variable in the **Label** box => choose whether the variable is numeric or string in the **Type** section => **Continue** => click **If...** => select whether the condition should “**include all cases**” versus “**include if case satisfies condition**.” => create a numeric expression by highlighting variables and moving them to the computation box with the arrow in conjunction with the calculator pad to create mathematical expressions (or alternatively highlight a preset mathematical function from the **Functions** box and move it up to the **Numeric Expression** box with the arrow) => **Continue** => enter a value or computation in the **Numeric Expression** box that creates the value of the new variable when the conditional statement is met => **OK**.
- **Copy/Paste Variable** To avoid pasting over an already defined variable, space for the new variable first must be inserted.
 - In Variable View, set the mouse where the new variable will be => **Data** => **Insert Variable**. Then, highlight with a left click the variable to copy => right click to **Copy Variable** => highlight with a left click to paste the variable => right click to **Paste Variable**.
- **Cut/Paste Variable** To avoid pasting over an already defined variable, space for the new variable must be created (see Insert Variable).
 - Highlight with a left click the variable to cut => right click to **Cut Variable** => highlight with a left click to paste the variable => right click to **Paste Variable**.
- **Information on Variables** This function presents a dialog box that displays a variety of information about the selected variable including: data format, variable label, user-missing values, and value labels.
 - **Utilities** => **Variables** => highlight the variable of interest in the **Variables** dialog box => Variable Information is listed on the right side of the dialog box.

Appendix A – Quick Reference Guide

- **Insert Variable** A new variable can be added to the dataset either in Data View or Variable View.
 - Select any cell in the variable to the right of (Data View) or below (Variable View) the position where the new variable will be inserted => **Data => Insert Variable**.
- **Move Variable** This process allows you to relocate a variable.
 - Click the variable name in Data View once and release (this highlights the variable column) => click and drag the variable or case to a new location => drop the variable in the new location (to place the variable between two existing variables, drop the variable on the variable column to the right of where the variable will be placed).
- **Recode Into Different Variable** This process reassigns the values of an existing variable into a new set of values for a new variable.
 - **Transform => Recode => Into Different Variables** => select the variable to recode and move it to the **Numeric Variable – Output Variable** dialog box using the arrow => type in the new variable name and label in the **Output Variable** dialog box => **Change => Old and New Values** => select and define the **Old Value** to change => define the **New Value** it will become => **Add => Continue => OK**.
- **Recode Into Same Variables** This process assigns new values in place of the old values of an existing variable. (Be very cautious when using this procedure to collapse data because the original data will be lost unless an original copy of the dataset is saved.)
 - **Transform => Recode => Into Same Variables** => select the variable to recode and move it to the **Variables** dialog box using the arrow => **Old and New Values** => select and define the **Old Value** to change => define the **New Value** it will become => **Add => Continue => OK**.

WILCOXON SIGNED-RANKS TEST is a nonparametric test that is similar to the paired samples *t* test and tests that two sample populations are equivalent. This test should be used for a repeated measures design when the dependent variable is ordinal or if the assumptions for the paired samples *t* test are markedly violated.

- **Analyze => Nonparametric Tests => 2 Related Samples** => select the two variables that make up the pair with left clicks and move them simultaneously to the **Test Pair(s) List** box by clicking the arrow in the middle of the dialog box => check the **Wilcoxon** box in the **Test Type** box => **OK**.

Z SCORES This procedure standardizes data to a standard score that has a mean of zero and a standard deviation of one.

- **Analyze => Descriptive Statistics => Descriptives** => select a variable with a left click => click the arrow in the middle of the dialog box to move the variable to the **Variables** box => check the box to **Save Standardized Values as Variables** => **OK**.

APPENDIX B

Getting Started With SPSS

This section includes step-by-step instructions for five procedures related to getting started with SPSS: a) a list of data files included on the compact disc, (b) opening and starting SPSS and getting familiar with the **SPSS Data Editor**, (c) setting your computer so that SPSS will print the **syntax** or **log** along with each output, (d) defining and labeling variables and their values, and (e) printing a dictionary or codebook.

The Data Files on the Compact Disk

The files included on the compact disk (CD) provided with this book are listed below. It is wise to make a working copy of the files.

hsbdataB.sav
college student data.sav
product data.sav
mixedMANOVAdata.sav
Anxiety 2.sav
bids.sav
dvdplayer.sav
judges.sav
Mall rentals.sav
Satisf.sav
site.sav
World95.sav

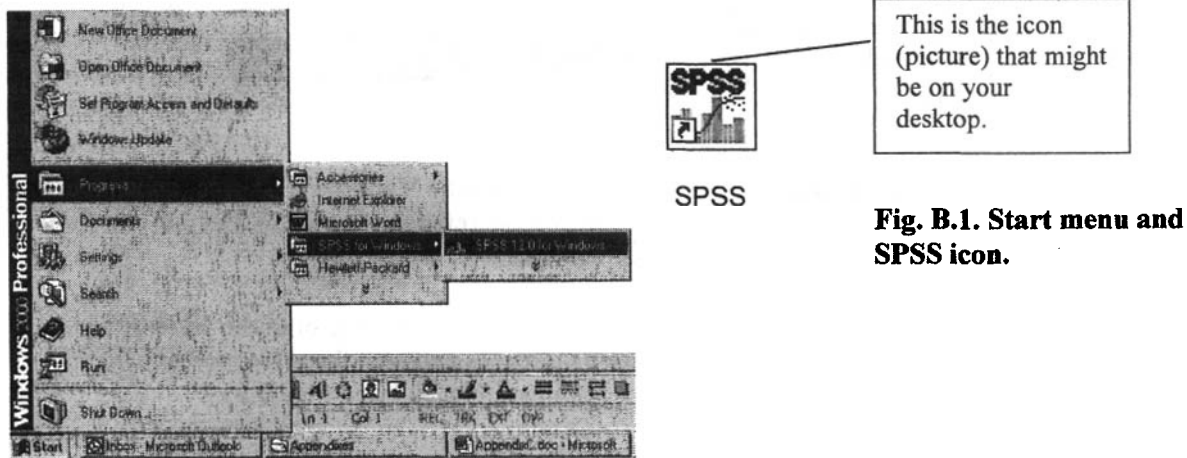
Note: You may not see the file extension (.sav) depending on your computer setup.

Open and Start the SPSS Application

- Begin at the **Start** button (bottom left of the Windows Desktop).
- Click **Start** => **Programs** => **SPSS for Windows** (See Fig. B.1).

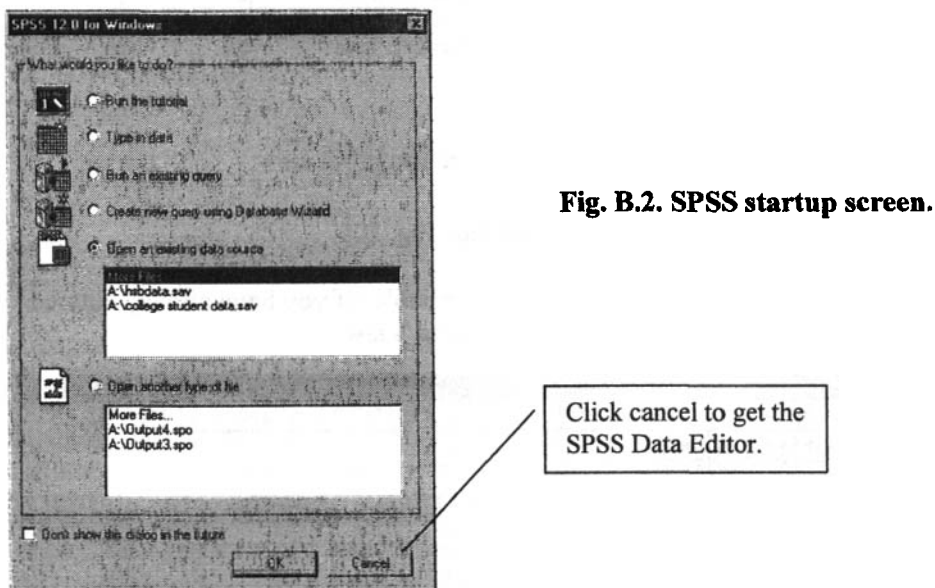
Alternatively, if an **SPSS icon** is available on the desktop, double click on it (see Fig. B.1). If **SPSS for Windows** is not listed in the **Programs** menu, it will need to be installed on your computer; it is not part of the Microsoft Windows package or the CD that accompanies this book and must be loaded separately.

Appendix B – Getting Started and Printing Syntax



After you start SPSS, you will see the **SPSS Startup** screen shown in Fig. B.2. In the Startup screen there is a list of all SPSS files that have recently been used on your computer.

- Click on the SPSS file you wish to use *or* click on **OK** to browse to find other SPSS files. If no files are listed, click **OK** to bring up the **Open File** dialogue box to search for the file you want to open. In any one session, only one data file can be open at time. If you want to create a new data file, click the **Cancel** button shown in Fig. B.2, which will bring up a new **SPSS Data Editor** as shown in Figs. B.4 and B.5.



Existing files as well as new data files are opened in the **SPSS Data Editor** screen. In this screen, there are two tabs at the bottom left side of the screen: the **Data View** tab and the **Variable View** tab (see Fig. B.3).

When you first open the SPSS data editor, you will have the Variable View or the Data View highlighted.

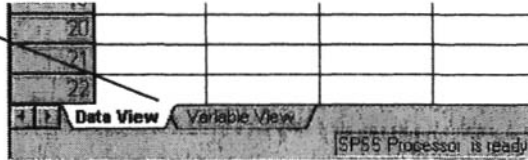


Fig. B.3. View tabs.

If you have SPSS 9.0 or a lower version, you will not have the Variable View screen option and will have to define and label variables differently. Please refer to your **SPSS Help menu** for further information on how to do this in earlier versions.

Although the toolbar at the top of the data editor screen is the same for both the **Variable** and **Data View** screens, it is important to notice the differences between these two screens found within the data editor in features and functions (compare Fig. B.4 and Fig. B.5).

- Click on the **Variable View** tab in the data editor screen to produce Fig. B.4.

Notice the column names in Fig. B.4. You define and label new variables using the **Variable View**.

h:\dataB.sav - SPSS Data Editor										
File Edit View Data Transform Analyze Graphs Utilities Window Help										
	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure
1	gend	Numeric	1	0	gender	{0, male}...	None	8	Right	Nominal
2	faed	Numeric	2	0	father's education	{2, < h.s. grad}	None	8	Right	Ordinal
3	maed	Numeric	2	0	mother's education	{2, < h.s.}...	None	8	Right	Ordinal
4	alg1	Numeric	1	0	algebra 1 in h.	{0, not taken}..	None	8	Right	Nominal
5	alg2	Numeric	1	0	algebra 2 in h.	{0, not taken}..	None	8	Right	Nominal
6	geo	Numeric	1	0	geometry in h.	{0, not taken}..	None	8	Right	Nominal

Fig. B.4. SPSS Data Editor: Variable view.

- Click on the **Data View** tab in the data editor to produce Fig. B.5.

Notice the columns change to **var** or to the names of your variables if you have already entered them. (See Fig. B.5.) You enter (input) the data using the **Data View**.

h:\dataB.sav - SPSS Data Editor													
File Edit View Data Transform Analyze Graphs Utilities Window Help													
	gend	faed	maed	alg1	alg2	geo	trig	calc	mathgr	grades	mathach	mosaic	visual
1	1	10	10	0	0	0	0	0	0	4	9.00	31.0	8.75
2	1	2	2	0	0	0	0	0	0	5	10.33	56.0	4.75
3	1	2	2	0	0	0	0	0	1	6	7.67	25.0	4.75
4	0	3	3	1	0	0	0	0	0	3	5.00	22.0	1.00
5	1	3	3	0	0	0	0	0	0	3	-1.67	17.5	2.25
6	1	3	2	0	0	0	0	0	1	5	1.00	23.5	1.00
7	0	9	6	1	1	1	0	0	0	6	12.00	26.5	2.50

Fig. B.5. SPSS Data Editor: Data view.

Set Your Computer to Print the SPSS Syntax (Log)

In order to have your computer print the SPSS commands on your output, as shown throughout the book, you will need to set your computer using the following:

- Click on **Edit => Options**.
- Click on the **Viewer** tab window to get Fig. B.6.

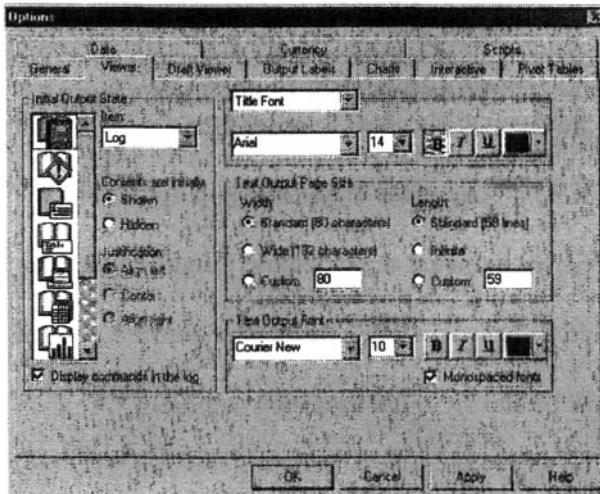


Figure B.6. Options.

- Check **Display commands in the log** near the lower left of the window.
- Leave the other defaults as is.
- Click on **OK**. Doing this will always print the syntax on your output until this box is unchecked.

Define the Variables

The following section will help you name and label variables. To do this, we need to use the **Variable View** screen. If you are in **Data View**, there are two ways to get to Variable View:

- Click on the **Variable View** tab at the bottom left of your screen (see Fig. B.3). In SPSS 10-12, this will bring up a screen similar to Fig. B.7. OR *double* click on **var** above the blank column to the far left side of the data matrix.

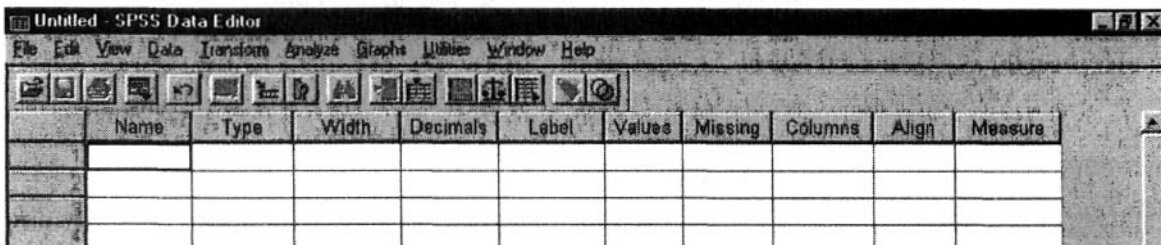


Fig. B.7. Blank variable view screen.

In this window, you will see ten sections that will allow you to input the variable name, type, width, decimals, label, values, missing (data), columns (width), align (data left or right), and measurement type.

Important: This is new to SPSS version 10. If you have SPSS 9 or a lower version, you will not have the Variable View screen option and will have to enter the variable information differently. Please refer to your **SPSS Help** menu for further information in lower versions.

To label a variable, follow these commands:

- Click in the blank box directly under **Name**.
- Type the name of the variable in the box directly beneath **Name** (see Fig. B.7.). Notice the number 1 to the left of this box. This indicates that you are entering your first variable. SPSS 11.5 and earlier versions allow only eight letters for the name. SPSS 12 allows as many as you wish, although we recommend keeping the name fairly short in order to make the outputs easier to read.
- Press enter. This will move your cursor to the next box, under **Type**, and pop up the information for **Type**, **Width**, etc.
- The **Type** indicates whether the variable levels are numbers or are letters. Usually, you will use numeric.
- **Width** indicates how wide the columns should be based on the number of characters allowed in the column.
- **Decimals** are the number of decimals that will be shown in the data view window.
- **Labels** gives a longer name for the variable.
- **Values** indicate labels for the levels of the variable.
- **Missing** indicates any special values for missing data. The SPSS system missing uses a blank. If only blanks are used for missing, this column will say “none” for no missing values.
- **Columns** defines the number of spaces for the variable in the data view window.
- **Align** indicates how you want the data aligned in the data view window.
- **Measure** shows the type of variable: nominal, ordinal, or scale.

Label the Levels for the Variables

For variables that are nominal or dichotomous it is important to label the levels so that you will know the group referred to by each number. For example, with the variable of *gender*, the levels are male and female, so you might label males as “1” and females as “2”. To do this, follow these commands:

- Under **Values**, click on the small gray box with ... (three dots) to get Fig. B.8.
- In the **Value Labels** window, type 1 in the **Value** box, *male* in the **Value Label** box, and then click **Add**. Do the same for 2 = female. The **Value Labels** window should resemble Fig. B.8 just before you click **Add** for the second time.

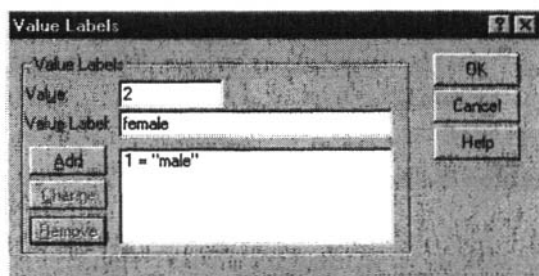


Fig. B.8. Value labels window.

Print a Dictionary or Codebook

Now that you have defined and labeled your variables, you can print a codebook or dictionary of your variables. It is a very useful record of what you have done. Notice that the information in the codebook is essentially the same as that in the variable view (Fig. B.4) so you do not really have to have both, but the codebook makes a more complete printed record of your labels and values.

- Select **File => Display Data File Information => Working File**. Your **Dictionary or Codebook** should look like Output B.1, without the callout boxes.

Output B.1: Codebook

DISPLAY DICTIONARY.

List of variables on the working file

Name	Position	The variable number
RECOMMEN I recommend course	1	
Measurement Level: Scale		
Column Width: 8 Alignment: Right		
Print Format: F8.2		
Write Format: F8.2		
Value Label		
1.00 strongly disagree		
5.00 strongly agree		
WORKHARD I worked hard	2	
Measurement Level: Scale		
Column Width: 8 Alignment: Right		
Print Format: F8.2		
Write Format: F8.2		
Value Label		
1.00 strongly disagree		
5.00 strongly agree		
COLLEGE college	3	
Measurement Level: Nominal		
Column Width: 8 Alignment: Right		
Print Format: F8		
Write Format: F8		
Missing Values: 98, 99		
Value Label		
1 arts & science		
2 business		
3 engineering		
98 M other, multiple ans.		
99 M blank		

Callout Boxes:

- Short variable name**: Points to the variable name in the list.
- Long variable label**: Points to the label in the list.
- This means the data for this variable can be up to eight digits including two decimal places.**: Points to the Print Format (F8.2) for the RECOMMEN variable.
- This variable has three unordered values, so it is called nominal.**: Points to the Measurement Level (Nominal) for the COLLEGE variable.
- These are missing values for this variable. Most variables use blanks, which are the SPSS system-missing value. However, here we used two missing value codes.**: Points to the Missing Values (98, 99) for the COLLEGE variable.
- The M after 98 and 99 shows that SPSS considers these numbers to be missing data.**: Points to the 'M' in the Value Label for the missing values of the COLLEGE variable.

SPSS for Intermediate Statistics

GENDER gender 4
Measurement Level: Nominal
Column Width: 8 Alignment: Right
Print Format: F8
Write Format: F8

Value Label

0 male
1 female

These are the values or levels of the gender variable and their labels.

GPA grade point average 5
Measurement Level: Scale
Column Width: 8 Alignment: Right
Print Format: F8.2
Write Format: F8.2

Value Label

.00 All F's
4.00 All A's

This variable has many ordered values; the possible values are equally spaced and probably normally distributed so we called it scale.

READING I did the reading 6
Measurement Level: Nominal
Column Width: 8 Alignment: Right
Print Format: F8
Write Format: F8

Value Label

0 not checked/blank
1 check

We call dichotomous (two level) variables nominal, but they are a special case, as discussed in chapter 1.

HOMEWORK I did the homework 7
Measurement Level: Nominal
Column Width: 8 Alignment: Right
Print Format: F8
Write Format: F8

Value Label

0 not check/blank
1 check

EXTRACRD I did the extra credit 8
Measurement Level: Nominal
Column Width: 8 Alignment: Right
Print Format: F8
Write Format: F8

Value Label

0 not checked
1 checked

APPENDIX C

Making Tables and Figures

Don Quick

Tables and figures are used in most fields of study to provide a visual presentation of important information to the reader. They are used to organize the statistical results of a study, to list important tabulated information, and to allow the reader a visual method of comparing related items. Tables offer a way to detail information that would be difficult to describe in the text.

A figure may be just about anything that is not a table, such as a chart, graph, photograph, or line drawing. These figures may include pie charts, line charts, bar charts, organizational charts, flow charts, diagrams, blueprints, or maps. Unless the figure can dramatically illustrate a comparison that a table cannot, use a table. A good rule is to use a table for numbers and text and to use figures for visual presentations.

The meaning and major focus of the table or figure should be evident to the readers without them having to make a thorough study of it. A glance should be all it takes for the idea of what the table or figure represents to be conveyed to the reader. By reading only the text itself, the reader may have difficulty understanding the data; by constructing tables and figures that are well presented, the readers will be able to understand the study results more easily.

The purpose of this appendix is to provide guidelines that will enhance the presentation of research findings and other information by using tables and figures. It will highlight the important aspects of constructing tables and figures using the *Publication Manual of the American Psychological Association, Fifth Edition* (2001) as the guide for formatting.

General Considerations Concerning Tables

Be selective as to how many tables are included in the total document. Determine how much data the reader needs to comprehend the material, and then decide if the information would be better presented in the text or as a table. A table containing only a few numbers is unnecessary; whereas a table containing too much information may not be understandable. Tables should be easy to read and interpret. If at all possible, combine tables that repeat data.

Keep a consistency to all of your tables throughout your document. All tables and figures in your document should use a similar format, with the results organized in a comparable fashion. Use the same designation measure or scale in all tables, figures, and the text.

In a final manuscript such as a thesis or dissertation, adjust the column headings or spacing between columns so the width of the table fits appropriately between the margins. Fit all of one table on one page. Reduce the data, change the type size, or decrease line spacing to make it fit. A short table may be on a page with text, as long as it follows the first mention of it. Each long table is on a separate page immediately after it is mentioned in the text. If the fit and appearance would be improved, turn the table sideways (landscape orientation, with the top of table toward the spine) on the page.

Each table and figure must be discussed in the text. An informative table will supplement but not duplicate the text. In the text, discuss only the most important parts of the table. Make sure the table can be understood by itself without the accompanying text; however, it is never independent of the text. There must be a reference in the text to the table.

Construction of the Table

Table C.1 is an example of an APA table for displaying simple descriptive data collected in a study. It also appears in correct relation to the text of the document. (Fig. C.1 shows the same table with the table parts identified.) The major parts of a table are: the number, the title, the headings, the body, and the notes.

Table C.1. *An Example of a Table in APA Format for Displaying Simple Descriptive Data*

Table 1

Means and Standard Deviations on the Measure of Self-Direction in Learning as a Function of Age in Adult Students

Age group	<i>n</i>	Self-directed learning inventory score	
		<i>M</i>	<i>SD</i>
20-34	15	65	3.5
35-40	22	88	6.3
50-64	14	79	5.6
65-79	7	56	7.1
80+	— ^a	—	—

Note. The maximum score is 100.

^a No participants were found for the over 80 group.

Table Numbering

Arabic numerals are used to number tables in the order in which they appear in the text. Do NOT write in the text “the table on page 17” or “the table above or below.” The correct method would be to refer to the table number like this: (see Table 1) or “Table 1 shows...” Left-justify the table number (see Table C.1). In an article, each table should be numbered sequentially in the order of appearance. Do not use suffix letters with the table numbers in articles. However, in a book table numbers may be numbered within chapters; e.g. Table 7.1. If the table appears in an appendix, identify it with the letter of the appendix capitalized, followed by the table number; for instance Table C.3 is the third table in Appendix C.

Table Titles

Include the variables, the groups on whom the data were collected, the subgroups, and the nature of the statistic reported. The table title and headings should concisely describe what is contained in the table. Abbreviations that appear in the body of the table can sometimes be explained in the title, however, it may be more appropriate to use a general note (see also comments below on *Table Headings*). The title must be italicized. Standard APA format for journal submission requires double spacing throughout. However, tables in student papers may be partially single spaced for better presentation.

The diagram illustrates the components of an APA table. Labels with arrows point to specific parts of the table structure:

- Table Number:** Points to "Table 1".
- Title:** Points to the italicized title "Means and Standard Deviations on the Measure of Self-Direction in Learning as a Function of Age in Adult Students".
- Column Spanner:** Points to the horizontal line above the column headings.
- Stub Column:** Points to the "Age group" column header.
- Headings:** Points to the row containing "n", "M", and "SD".
- Cell:** Points to the value "6.3" in the "SD" column for the "35-40" age group.
- Body:** Points to the rows of data below the headings.
- Notes:** Points to the "Note" section at the bottom.

A box on the left states: "Use horizontal lines under the title, headings and the body; but no vertical lines."

Age group	n	M	SD
20-34	15	65	3.5
35-40	22	88	6.3
50-64	14	79	5.6
65-79	7	56	7.1
80+	-- ^a	--	--

Note. The maximum score is 100.

^a No participants were found for the over 80 group.

Fig. C.1. The major parts of an APA table.

Table Headings

Headings are used to explain the organization of the table. You may use abbreviations in the headings; however, include a note as to their meaning if you use mnemonics, variable names, and scale acronyms. Standard abbreviations and symbols for non technical terms can be used without explanation (e.g., *no.* for *number* or *%* for *percent*). Have precise title, column headings, and row labels that are accurate and brief. Each column must have a heading, including the *stub column*, or leftmost column. Its heading is referred to as the *stubhead*. The stub column usually lists the significant independent variables or the levels of the variable, as in Table C.1.

The *column heads* cover one column, and the *column spanners* cover two or more columns -- each with its own column head (see Table C.1 and Fig. C.1). Headings stacked in this manner are called *decked heads*. This is a good way to eliminate repetition in column headings but try to

avoid using more than two levels of decked heads. **Column heads, column spanners, and stubheads** should all be singular, unless referring to a group (e.g., children). Table spanners, which cover the entire table, may be plural. Use sentence capitalization in all headings.

Notice that there are no vertical lines in an APA style table. The horizontal lines can be added by using a “draw” feature or a “borders” feature for tables in the computer word processor, or they could be drawn in by hand if typed. If translating from an SPSS Table or box, the vertical lines must be removed.

The Body of the Table

The body contains the actual data being displayed. Round numbers improve the readability and clarity more than precise numbers with several decimal places. A good guideline is to report two digits more than the raw data. A reader can compare numbers down a column more easily than across a row. Column and row averages can provide a visual focus that allows the reader to inspect the data easily without cluttering the table. If a cell cannot be filled because the information is not applicable, then leave it blank. If it cannot be filled because the information could not be obtained, or was not reported, then insert a dash and explain the dash with a note to the table.

Notes to a Table

Notes are often used with tables. There are three different forms of notes used with tables: a) to eliminate repetition in the body of the table, b) to elaborate on the information contained in a particular cell, or c) to indicate statistical significance:

- A **general note** provides information relating to the table as a whole, including explanations of abbreviations used:

Note. This could be used to indicate if the table came from another source.

- A **specific note** makes a reference to a specific row or column or cell of the table and is given superscript lowercase letters, beginning with the letter “a”:

^a*n* = 50. Specific notes are identified in the body with superscript.

- A **probability note** is to be included when one or more inferential statistic has been computed and there isn’t a column showing the probability, *p*. Asterisk(s) indicate the statistical significance of findings presented within the table. Try to be consistent across all tables in a paper. The important thing is to use the fewest asterisks for the largest *p* value. It is common to use one asterisk for .05 and two for .01. For example:

p* < .05. *p* < .01.

Notes should be listed with general notes first, then specific notes, and concluded with probability notes, without indentation. They may be single spaced for better presentation. Explain all uses of dashes and parentheses. Abbreviations for technical terms, group names, and those of a similar nature must be explained in a note to the table.

Constructing a Table in Microsoft Word XP or 2000

For this step-by-step example the ANOVA was chosen from previous examples in the book. See Fig. C.2. The data are transferred from the standard SPSS output to an APA table.

ANOVA

grades in h.s.

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	18.143	2	9.071	4.091	.021
Within Groups	155.227	70	2.218		
Total	173.370	72			

Fig. C.2. An example of the type of default table generated from a SPSS ANOVA output.

The finished table should look like Table D.2. This explanation is accomplished using MS Word XP but using MS Word 2000 will be similar. Any earlier version will have some problems with line spacing. You will also need to adjust the number of columns and rows for the type and amount of data that you need to present.

Table C.2. *An Example of an ANOVA Table in APA Format*

Table 2

The Table Number is double spaced but the Table Title is single spaced. The Title is in italics but the Table Number is not.

One-Way Analysis of Variance of Grades in High School by Father's Education

Source	df	SS	MS	F	p
Between groups	2	18.14	9.07	4.09	.02
Within groups	70	155.23	2.22		
Total	72	173.37			

The **Headings** and **Body** of the table are actually built using Word's table function. Type your **Table Number** and **Title**. Then on the next line after the title, insert a 6x4 table:

- **Table** => **Insert** => **Table...** (See Fig. C.3).
- For our example of the ANOVA set it to 6 columns and 4 rows. Leave everything else as is. See Fig. C.4.
- Click **OK**.

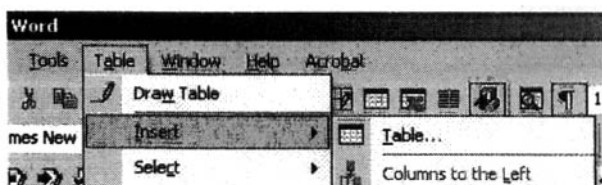


Fig. C.3. Using MS Word to make a table.

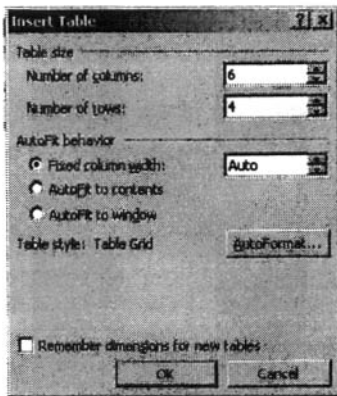


Fig. C.4. Inserting a 6x4 table.

This is for the single ANOVA table. You will need to adjust the number of columns and rows for the type and amount of data that you need to present.

Compare your table to Table C.3.

Table C.3.

APA uses no vertical and just a few horizontal lines so it is best to remove them all and then put back the ones that are needed:

- Select the table by clicking anywhere inside the table, then: **Table => Select => Table.**
- **Format => Borders and Shading...** to get Fig. C.5.
- Select the **Borders** tab, if it's not already selected.
- Under **Settings** click the box to the left of **None** to remove them.
- Click **OK.**

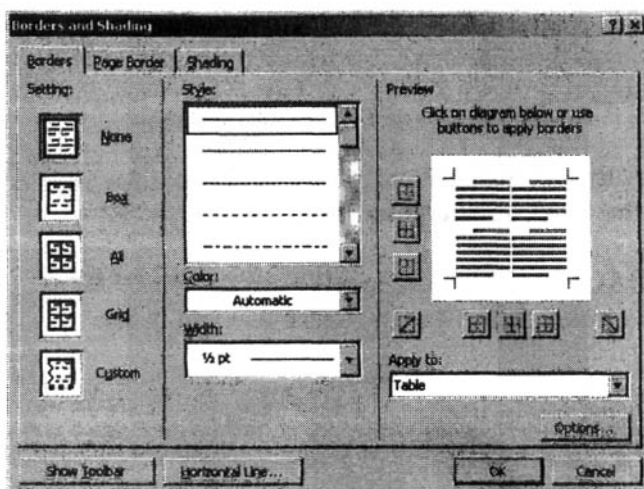


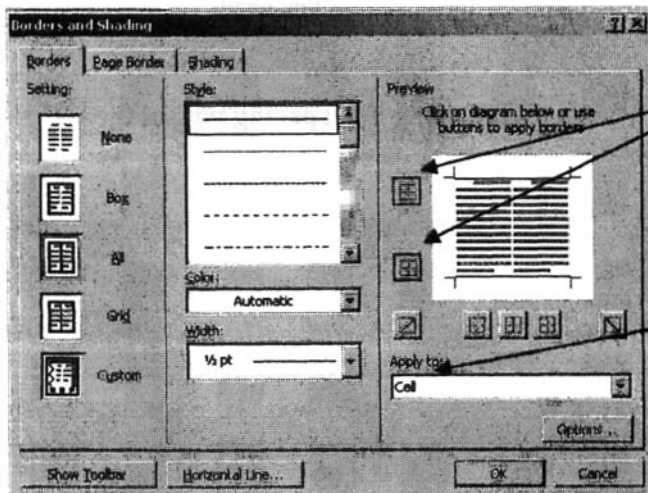
Fig. C.5. Clearing the borders.

To add the correct lines in for an APA table:

- Clicking anywhere in the top row and **Table => Select => Row.**

Appendix C - Making Tables and Figures

- **Format => Borders and Shading...** to get Fig. C.6.
- Make sure the solid line **Style** is selected and the **Width** is **1/2 pt.**
- In the **Preview** picture click the **Upper** and **Lower** bar buttons. This inserts the top two lines in the table.
- Click **OK**.
- Select the last row in your table.
- Click the **Lower** bar button only. This inserts the bottom line in the table.
- Click **OK**.



Upper and Lower bar buttons

Fig. C.6. Setting the horizontal lines.

Note: the **Apply to** is set to **Cell** but since you have selected the entire row it will look like a solid line across the table.

Compare your table to Table C.4.

Table C.4.

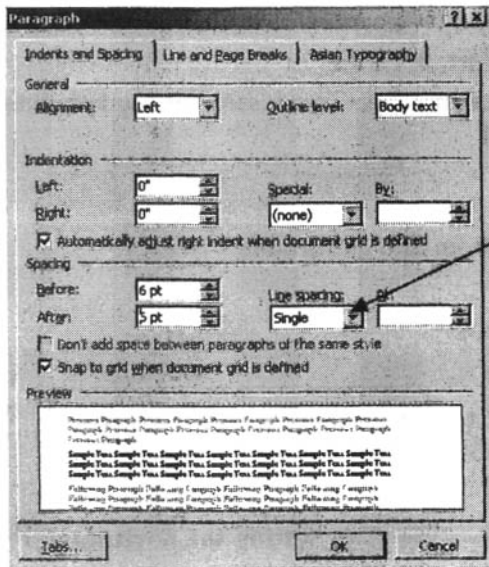
Note: If you can't see the gridlines, turn them on to better see where the rows and cells are. They won't show when printed. Click **Table => Show Gridlines**

The text in the body of an APA table is equal distance between the top and bottom of the cell:

- Select the table by clicking anywhere inside the table, then: **Table => Select => Table**.
- Click **Format => Paragraph...**
- Set **Line spacing** to **Single** (see note on Fig. C.7).
- Set **Spacing** to **Before and After** to **6pt** (see Fig. C.7).
- Click **OK**.

Enter the headings and data into each cell; the SPSS printout will have all of the information to accomplish this. Don't worry about wrapping and aligning at this time. That is easier to do after the data are entered.

Compare your table to Table C.5.



Note: With normal SPSS tables with numerical data and short heads, **Single** and the **6pt** before and after sets the text equal distant in the cell. Wrapped text will come out single spaced which is better presentation for student papers and dissertations. However, if you have text that wraps and the journal requires that all text be double spaced then you will need to **Double** space only those cells that wrap.

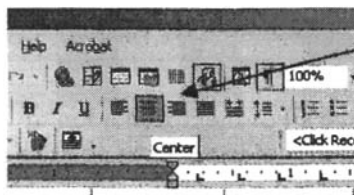
Fig. C.7. Setting line spacing within the cell.

Table C.5.

Source	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Between groups	2	18.14	9.07	4.09	.02
Within groups	70	155.23	2.22		
Total	72	173.37			

In an APA table the **Heads** should be center aligned in the cell and the **Stubs** are left aligned. The numbers in the **Cell** are decimal aligned and centered under the Heads. Notice also that “Between groups” wrapped. Let’s first align the **Heads** and **Stubs**, then fix that wrap and finally align the data off of the **Heads**. To center align the **Heads**:

- Select the **Header Row** of your table by clicking anywhere in the top row and **Table => Select => Row**.
- Click the **Center** align button in the Formatting Toolbar, see Fig. C.8.
- The stub column should already be left aligned; if not, then select the cells and click the **Align Left** button.



Note: If the **Align Buttons** aren’t showing on the **Formatting Toolbar** then you can select the proper alignment from the menu: **Format => Paragraph...**

Fig. C.8. Center aligning the Heads.

Appendix C - Making Tables and Figures

When MS Word creates a table it will generally make all of the columns the same width. To fix the wrap on the “Between groups” cell, that column will need to be widened slightly and then to keep the table within the margins the data columns will need to be decreased slightly. This may be a trail and error process to get the right spacing for your text.

- Click anywhere on the **Stubs** column.
- **Table** => **Table Properties...** to get Fig. C.9.
- Click the **Column** Tab.
- Set the **Preferred width** to 1.4”.
- Click the **Next Column** button and set it to 1.0”.
- Repeat for all of the columns, setting them to 1.0”.
- Click **OK**.

Note: This can also be accomplished by dragging the vertical column separator lines until the “Between groups” is not wrapped and then dragging the other column separator lines so that they are within the margins. However this produces uneven column spaces. We recommend the method outlined.

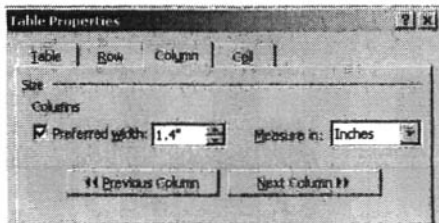


Fig. C.9. Adjusting the column widths.

Compare your table to Table C.6.

Table C.6.

Source	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Between groups	2	18.14	9.07	4.09	.02
Within groups	70	155.23	2.22		
Total	72	173.37			

To set the **Cell** columns so that they are all centered under its **Head**, you will need to set the **Tabs** for each column of data cells to a **Decimal Tab**. We recommend this method of setting all columns the same and then adjusting them separately so they look right, because it requires less individual column adjustment:

- Select just the data cells by clicking in the upper left one, hold the shift key down, and then click in the lower right cell.
- **Format** => **Tabs...** to get Fig. C.10.
- Clear all of the Tabs in the selected cells first by clicking the **Clear All** button.
- Click **Alignment Decimal**.
- Type .35” in the **Tab stop position** box.
- Click the **Set** button.
- Click **OK**.

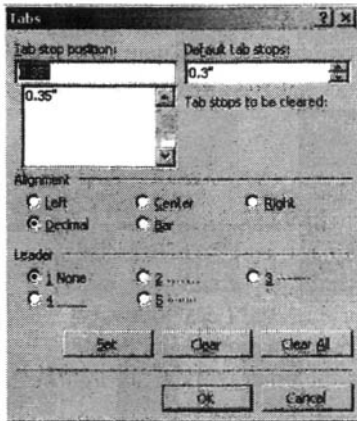


Fig. C.10. Setting the decimal tabs.

Compare your table to Table C.7.

Table C.7.

Source	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Between groups	2	18.14	9.07	4.09	.02
Within groups	70	155.23	2.22		
Total	72	173.37			

The *df* column looks like it could be adjusted slightly to the right and the *p* column slightly to the left. We show you this so that you will know how to get a perfect decimal alignment of the data under the column head text. This may be trial and error depending on your data.

- Select the cells of the *df* column by clicking first on the top data cell, "2," hold the **Shift** key down, and then click on the bottom data cell, "72."
- **Format => Tabs...**
- Clear all of the Tabs in the selected cells first by clicking the **Clear All** button.
- Click **Alignment Decimal**.
- Type .45" in the **Tab stop position** box, to set decimal tap .45" from the left edge of the cell.
- Click the **Set** button.
- Click **OK**.
- Repeat for the *p* column but set it to .25" to set decimal tap .25" from the left edge of the cell.

Compare your finished table to Table C.8.

Table C.8.

Table 2

One-Way Analysis of Variance of Grades in High School by Father's Education

Source	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>p</i>
Between groups	2	18.14	9.07	4.09	.02
Within groups	70	155.23	2.22		
Total	72	173.37			

Adjusting the SPSS Output to Approximate the APA Format

The preceding example shows how the standard SPSS output can be used to create a table in APA format. However, this does require some knowledge of your word processing program's table creation capabilities in order to accomplish this task. It also requires retyping the data into the table. You can adjust SPSS so that the output will approximate the APA format. We would not recommend submitting this to an APA journal, but it may be acceptable for student papers and some graduate program committees.

In SPSS follow these commands BEFORE running your SPSS analysis of your data:

- Click **Edit => Options.**
- Under the **Pivot Tables** tab select **Academic 2.tlo** (see Fig. C.11).
- Press **OK.**

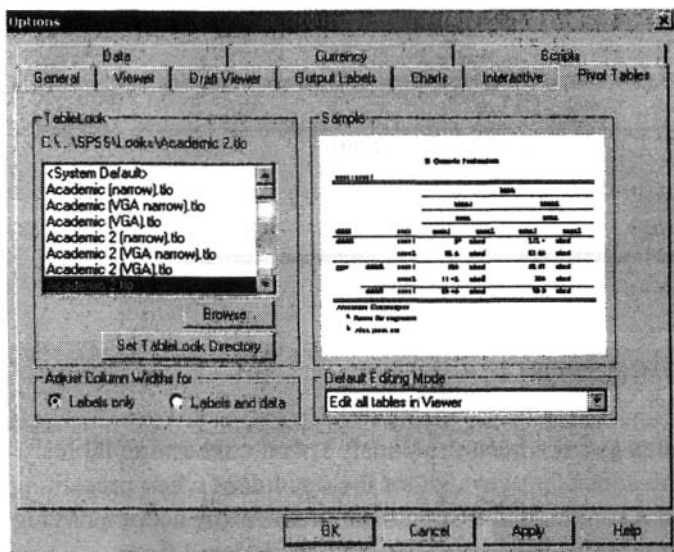


Fig. C.11. Setting SPSS for an approximate APA format output.

- Run the SPSS statistical analysis.

Your outputs will look similar to Table C.9, which approximates an APA table. In order to transfer it to MS Word:

- On the SPSS output, **right click** on the table that you want to transfer.
- Select **Copy objects** from the short menu presented, see Fig. C.12.

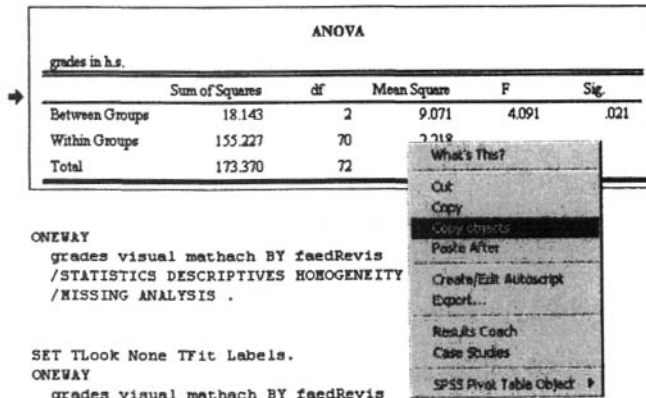


Fig. C.12. Copying tables from SPSS.

- Place the curser in the MS Word file where you want to put the table.
- Select **Paste** in MS Word.

You can then move it and format it like any other image in MS Word, but it can not be edited.

Table C.9. An Example of the SPSS "Academic" Output

Table 2

One-Way Analysis of Variance of Grades in High School by Father's Education

ANOVA					
grades in h.s.					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	18.143	2	9.071	4.091	.021
Within Groups	155.227	70	2.218		
Total	173.370	72			

Using Figures

Generally, the same concepts apply to figures as have been previously stated concerning tables: they should be easy to read and interpret, be consistent throughout the document when presenting the same type of figure, kept on one page if possible, and it should supplement the accompanying text or table. Considering the numerous types of figures, I will not attempt to detail their construction in this document. However, one thing is consistent with all figures. The figure

Appendix C - Making Tables and Figures

number and caption description are located below the figure and the description is detailed, similar to that of the title of a table. See Fig. C.12.

Some cautions in using figures:

- 1) Make it simple. Complex diagrams that require lengthy explanation should be avoided unless it is an integral part of the research.
- 2) Use a minimum number of figures for just your important points. If too many figures are used, your important points may be lost.
- 3) Integrate text and figure. Make sure your figure compliments and enhances the accompanying text.

Correlation of math achievement with

mosaic pattern test

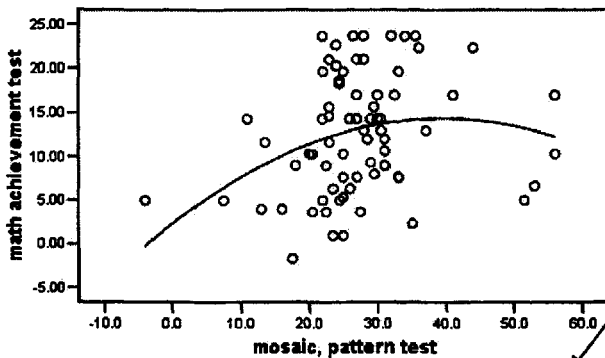


Fig. C.12. An example of using a figure in APA with caption.

Note: The figure number is italicized but the caption itself is not. Also, the caption text is sentence case (only the first word is capitalized).

Fig. 1. Correlation of math achievement with mosaic pattern test.

APPENDIX D

Answers to Odd Interpretation Questions

1.1 What is the difference between the independent variable and the dependent variable?

Independent variables are predictors, antecedents, or *presumed* causes or influences being studied. Differences in the Independent variable are hypothesized to affect, predict, or explain differences in the dependent or outcome variable. So, independent variables are predictor variables; whereas dependent variables are the variables being predicted, or outcome variables.

1.3 What kind of independent variable is necessary to infer cause? Can one always infer cause from this type of independent variable? If so, why? If not, when can one clearly infer cause and when might causal inferences be more questionable? A variable must be an *active* independent variable in order for the possibility to exist of one's inferring that it caused the observed changes in the dependent variable. However, even if the independent variable is active, one can not attribute cause to it in many cases. The strongest inferences about causality can be made when one randomly assigns participants to experimentally manipulated conditions and there are no pre-existing differences between the groups that could explain the results. Causal inferences are much more questionable when manipulations are given to pre-existing groups, especially when there is not pretest of the dependent variable prior to the manipulation, and/or the control group receives no intervention at all, and/or there is no control group.

1.5 Write three research questions and a corresponding hypothesis regarding variables of interest to you, but not in the HSB data set (one associational, one difference, and one descriptive question). Associational research question: What is the relation between guilt and shame in 10-year-old children? Associational hypothesis: Guilt and shame are moderately to highly related in 10 year-old children. Difference question: Are there differences between Asian-Americans and European-Americans in reported self-esteem? Difference hypothesis: Asian-Americans, on the average, report lower self-esteem than do European-Americans. Descriptive question: What is the incidence of themes of violence in popular songs, folk songs, and children's songs? Descriptive hypothesis: There will be more violent themes in popular songs than in folk songs.

1.7 If you have categorical, ordered data (such as low income, middle income, high income), what type of measurement would you have? Why? Categorical, ordered data would typically be considered ordinal data because one can not assume equal intervals between levels of the variable, there are few levels of the variable, and data are unlikely to be normally distributed, but there is a meaningful order to the levels of the variable.

1.9 What percent of the area under the standard normal curve is between the mean and one standard deviation above the mean? Thirty-four percent of the normal distribution is within one standard deviation *above* the mean. Sixty-eight percent is within one standard deviation above or below the mean.

2.1 Using Output 2.1a and 2.1b: a) What is the mean *visualization test* score? 5.24 b) What is the range for *grades in h.s.*? 6 c) What is the minimum score for *mosaic pattern test*? -4 How does this compare to the values for that variable as indicated in chapter 1? It is the

Appendix D – Answers to Odd Interpretation Questions

lowest possible score. **Why would this score be the minimum?** It would be the minimum if at least one person scored this, and this is the lowest score anyone made.

- 2.3 **Using Output 2.4:** a) Can you interpret the means? Explain. Yes, the means indicate the percentage of participants who scored “1” on the measure. b) How many participants are there all together? 75 c) How many have complete data (nothing missing)? 75 d) What percent are *male* (if *male*=0)? 45 e) What percent took *algebra 1*? 79
- 2.5 **In Output 2.8a:** a) Why are matrix scatterplots useful? What assumption(s) are tested by them? They help you check the assumption of linearity and check for possible difficulties with multicollinearity.
- 3.1 a) Is there only one appropriate statistic to use for each research design? No.
b) Explain your answer. There may be more than one appropriate statistical analysis to use with each design. Interval data can always use statistics used with nominal or ordinal data, but you lose some power by doing this.
- 3.3 **Interpret the following related to effect size:**
a) $d = .25$ small c) $R = .53$ large e) $d = 1.15$ very large
b) $r = .35$ medium d) $r = .13$ small f) $\eta = .38$ large
- 3.5. What statistic would you use if you had two independent variables, income group (<\$10,000, \$10,000-\$30,000, >\$30,000) and ethnic group (Hispanic, Caucasian, African-American), and one normally distributed dependent variable (self-efficacy at work). Explain. Factorial ANOVA, because there are two or more between groups independent variables and one normally distributed dependent variable. According to Table 3.3, column 2, first cell, I should use Factorial ANOVA or ANCOVA. In this case, both independent variables are nominal, so I'd use Factorial ANOVA (see p. 49).
- 3.7 What statistic would you use if you had three normally distributed (scale) independent variables and one dichotomous independent variable (weight of participants, age of participants, height of participants and gender) and one dependent variable (positive self-image), which is normally distributed. Explain. I'd use multiple regression, because all predictors are either scale or dichotomous and the dependent variable is normally distributed. I found this information in Table 3.4 (third column).
- 3.9 What statistic would you use if you had one, repeated measures, independent variable with two levels and one nominal dependent variable? McNemar because the independent variable is repeated and the dependent is nominal. I found this in the fourth column of Table 3.1.
- 3.11 What statistic would you use if you had three normally distributed and one dichotomous independent variable, and one dichotomous dependent variable?
I would use logistic regression, according to Table 3.4, third column.
- 4.1 **Using Output 4.1 to 4.3, make a table indicating the mean interitem correlation and the alpha coefficient for each of the scales. Discuss the relationship between mean interitem correlation and alpha, and how this is affected by the number of items.**

SPSS for Intermediate Statistics

Scale	Mean inter-item correlation	Alpha
Motivation	.386	.791
Competence	.488	.796
Pleasure	.373	.688

The alpha is based on the inter-item correlations, but the number of items is important as well. If there are a large number of items, alpha will be higher, and if there are only a few items, then alpha will be lower, even given the same average inter-item correlation. In this table, the fact that both number of items and magnitude of inter-item correlations are important is apparent. Motivation, which has the largest number of items (six), has an alpha of .791, even though the average inter-item correlation is only .386. Even though the average inter-item correlation of Competence is much higher (.488), the alpha is quite similar to that for Motivation because there are only 4 items instead of 6. Pleasure has the lowest alpha because it has a relatively low average inter-item correlation (.373) and a relatively small number of items (4).

- 4.3 For the pleasure scale (Output 4.3), what item has the highest item-total correlation? Comment on how alpha would change if that item were deleted.** Item 14 (.649). The alpha would decline markedly if Item 14 were deleted, because it is the item that is most highly correlated with the other items.
- 4.5 Using Output 4.5: What is the interrater reliability of the ethnicity codes? What does this mean?** The interrater reliability is .858. This is a high kappa, indicating that the school records seem to be reasonably accurate with respect to their information about students' ethnicity, assuming that students accurately report their ethnicity (i.e., the school records are in high agreement with students' reports). Kappa is not perfect, however (1.0 would be perfect), indicating that there are some discrepancies between school records and students' own reports of their ethnicity.
- 5.1 Using Output 5.1: a) Are the factors in Output 5.1 close to the conceptual composites (motivation, pleasure, competence) indicated in Chapter 1 ?** Yes, they are close to the conceptual composites. The first factor seems to be a competence factor, the second factor a motivation factor, and the third a (low) pleasure factor. However, Item01 (I practice math skills until I can do them well) was originally conceptualized as a motivation question, but it had its strongest loading from the first factor (the competence factor), and there was a strong cross-loading for item02 (I feel happy after solving a hard problem) on the competence factor. **b) How might you name the three factors in Output 5.1?** Competence, motivation, and (low) mastery pleasure **c) Why did we use Factor Analysis, rather than Principal Components Analysis for this exercise?** We used Factor Analysis because we had beliefs about underlying constructs that the items represented, and we wished to determine whether these constructs were the best way of understanding the manifest variables (observed questionnaire items). Factor analysis is suited to determining which latent variables seem to explain the observed variables. In contrast, Principal Components Analysis is designed simply to determine which linear combinations of variables best explain the variance and covariation of the variables so that a relatively large set of variables can be summarized by a smaller set of variables.
- 5.3 What does the plot in Output 5.2 tell us about the relation of *mosaic* to the other variables and to component 1?** *Mosaic* seems not to be related highly to the other variables nor to component 1. **How does this plot relate to the rotated component matrix?** The plot

illustrates how the items are located in space in relation to the components in the rotated component matrix.

- 6.1. In Output 6.1: a) What information suggests that we might have a problem of collinearity? High intercorrelations among some predictor variables and some low tolerances ($< 1-R^2$) b) How does multicollinearity affect results? It can make it so that a predictor that has a high zero-order correlation with the dependent variable is found to have little or no relation to the dependent variable when the other predictors are included. This can be misleading, in that it appears that one of the highly correlated predictors is a strong predictor of the dependent variable and the other is not a predictor of the dependent variable. c) What is the adjusted R^2 and what does it mean? The adjusted R^2 indicates the percentage of variance in the dependent variable explained by the independent variables, after taking into account such factors as the number of predictors, the sample size, and the effect size.
- 6.3 In Output 6.3. a) Compare the adjusted R^2 for model 1 and model 2. What does this tell you? It is much larger for Model 2 than for Model 1, indicating that *grades in high school*, *motivation*, and *parent education* explain additional variance, over and above that explained by *gender*. b) Why would one enter *gender* first? One might enter gender first because it was known that there were gender differences in math achievement, and one wanted to determine whether or not the other variables contributed to prediction of math achievement scores, over and above the “effect” of gender.
- 7.1 Using Output 7.1: a) Which variables make significant contributions to predicting who took algebra 2? Parent’s education and visualization b) How accurate is the overall prediction? 77.3% of participants are correctly classified, overall c) How well do the significant variables predict who took algebra 2? 71.4% of those who took algebra 2 were correctly classified by this equation. d) How about the prediction of who *didn’t* take it? 82.5% of those who didn’t take algebra 2 were correctly classified.
- 7.3 In Output 7.3: a) What do the discriminant function coefficients and the structure coefficients tell us about how the predictor variables combine to predict who took algebra 2? The function coefficients tell us how the variables are weighted to create the discriminant function. In this case, *parent’s education* and *visual* are weighted most highly. The structure coefficients indicate the correlation between the variable and the discriminant function. In this case, *parent’s education* and *visual* are correlated most highly; however, *gender* also has a substantial correlation with the discriminant function. b) How accurate is the prediction/classification overall and for who would not take algebra 2? 76% were correctly classified, overall. 80% of those who did not take algebra 2 were correctly classified; whereas 71.4% of those who took algebra 2 were correctly classified. c) How do the results in Output 7.3 compare to those in Output 7.1, in terms of success at classifying and contribution of different variables to the equation? For those who took algebra 2, the discriminant function and the logistic regression yield identical rates of success; however, the rate of success is slightly lower for the discriminative function than the logistic regression for those who did not take algebra 2 (and, therefore, for the overall successful classification rate).
- 7.5 In Output 7.2: why might one want to do a hierarchical logistic regression? One might want to do a hierarchical logistic regression if one wished to see how well one predictor successfully distinguishes groups, over and above the effectiveness of other predictors.

8.1 In Output 8.1: a) Is the interaction significant? Yes b) Examine the profile plot of the cell means that illustrates the interaction. Describe it in words. The profile plot indicates that the “effect” of math grades on math achievement is different for students whose fathers have relatively little education, as compared to those with more education. Specifically, for students whose fathers have only a high school education (or less), there is virtually no difference in math achievement between those who had high and low math grades; whereas for those whose fathers have a bachelor’s degree or more, those with higher math grades obtain higher math achievement scores, and those with lower math grades obtain lower math achievement scores. **c) Is the main effect of *father’s education* significant? Yes. Interpret the eta squared.** The eta squared of .243 ($\eta^2 = .496$) for father’s education indicates that this is, according to Cohen’s criteria, a large effect. This indicates that the “effect” of the level of fathers’ education is larger than average for behavioral science research. However, it is important to realize that this main effect is qualified by the interaction between *father’s education* and *math grades* **d) How about the “effect” of *math grades*? The “effect” of *math grades* also is significant. Eta squared is .139 for this effect ($\eta^2 = .37$), which is also a large effect, again indicating an effect that is larger than average in behavioral research. **e) Why did we put the word effect in quotes?** The word, “effect,” is in quotes because since this is not a true experiment, but rather is a comparative design that relies on attribute independent variables, one can not impute causality to the independent variable. **f) How might focusing on the main effects be misleading?** Focusing on the main effects is misleading because of the significant interaction. In actuality, for students whose fathers have less education, math grades do not seem to “affect” math achievement; whereas students whose fathers are highly educated have higher achievement if they made better math grades. Thus, to say that math grades do or do not “affect” math achievement is only partially true. Similarly, fathers’ education really seems to make a difference only for students with high math grades.**

8.3 In Output 8.3: a) Are the adjusted main effects of *gender* significant? No. b) What are the adjusted *math achievement* means (marginal means) for males and females? They are 12.89 for males and 12.29 for females **c) Is the effect of the covariate (*mathcrs*) significant? Yes d) What do a) and c) tell us about *gender* differences in *math achievement* scores?** Once one takes into account differences between the genders in math courses taken, the differences between genders in math achievement disappear.

9.1 In Output 9.2: a) Explain the results in nontechnical terms Output 9.2a indicates that the ratings that participants made of one or more products were higher than the ratings they made of one or more other products. Output 9.2b indicates that most participants rated product 1 more highly than product 2 and product 3 more highly than product 4, but there was no clear difference in ratings of products 2 versus 3.

9.3 In Output 9.3: a) Is the Mauchly sphericity test significant? Yes. Does this mean that the assumption is or is not violated? It is violated, according to this test. **If it is violated, what can you do?** One can either correct degrees of freedom using epsilon or one can use a MANOVA (the multivariate approach) to examine the within-subjects variable **b) How would you interpret the *F* for *product* (within subjects)?** This is significant, indicating that participants rated different products differently. However, this effect is qualified by a significant interaction between product and gender. **c) Is the interaction between product and gender significant? Yes. How would you describe it in non-technical terms?** Males rated different products differently, in comparison to females, with males rating some higher and some lower than did females. **d) Is there a significant difference between the genders? No. Is a post hoc multiple comparison test needed? Explain.** No post hoc test is needed for gender, both because the effect is not significant and because there are only two groups, so

one can tell from the pattern of means which group is higher. For product, one could do post hoc tests; however, in this case, since products had an order to them, linear, quadratic, and cubic trends were examined rather than paired comparisons being made among means.

- 10.1 In Output 10.1b: a) Are the multivariate tests statistically significant? Yes. b) What does this mean? This means that students whose fathers had different levels of education differed on a linear combination of *grades in high school*, *math achievement*, and *visualization scores*. c) Which individual dependent variables are significant in the ANOVAs? Both *grades in h.s.*, $F(2, 70) = 4.09, p = .021$ and *math achievement*, $F(2, 70) = 7.88, p = .001$ are significant. d) How are the results similar and different from what we would have found if we had done three univariate one-way ANOVAs? Included in the output are the very same 3 univariate one-way ANOVAs that we would have done. However, in addition, we have information about how the *father's education* groups differ on the three dependent variables, taken together. If the multivariate tests had not been significant, we would not have looked at the univariate tests; thus, some protection for Type I error is provided. Moreover, the multivariate test provides information about how each of the dependent variables, over and above the other dependent variables, distinguishes between the *father's education* groups. The **parameter estimates** table provides information about how much each variable was weighted in distinguishing particular *father's education* groups.
- 10.3 In Output 10.3: a) What makes this a “doubly multivariate” design? This is a doubly multivariate design because it involves more than one dependent variable, each of which is measured more than one time. b) What information is provided by the multivariate tests of significance that is not provided by the univariate tests? The multivariate tests indicate how the two dependent variables, taken together, distinguish the intervention and comparison group, the pretest from the posttest, and the interaction between these two variables. Only it indicates how each outcome variable contributes, over and above the other outcome variable, to our understanding of the effects of the intervention. c) State in your own words what the interaction between time and group tells you. This significant interaction indicates that the change from pretest to posttest is different for the intervention group than the comparison group. Examination of the means indicates that this is due to a much greater change from pretest to posttest in Outcome 1 for the intervention group than the comparison group. What implications does this have for understanding the success of the intervention? This suggests that the intervention was successful in changing Outcome 1. If the intervention group and the comparison group had changed to the same degree from pretest to posttest, this would have indicated that some other factor was most likely responsible for the change in Outcome 1 from pretest to posttest. Moreover, if there had been no change from pretest to posttest in either group, then any difference between groups would probably not be due to the intervention. This interaction demonstrates exactly what was predicted, that the intervention affected the intervention group, but not the group that did not get the intervention (the comparison group).

For Further Reading

- American Psychological Association (APA). (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington, DC: Author.
- Cohen, J. (1988). *Statistical power and analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Gliner, J. A., & Morgan, G. A. (2000). *Research methods in applied settings: An integrated approach to design and analysis*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Hair, J. F., Jr., Anderson, R.E., Tatham, R.L., & Black, W.C. (1995). *Multivariate data analysis* (4th ed.). Englewood Cliffs, NJ: Prentice Hall.
- Huck, S. J. (2000). *Reading statistics and research* (3rd ed.). New York: Longman.
- Morgan, G. A., Leech, N. L., Gloeckner, G. W., & Barrett, K. C. (2004). *SPSS for introductory statistics: Use and Interpretation*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Morgan, S. E., Reichart, T., & Harrison T. R. (2002). *From numbers to words: Reporting statistical results for the social sciences*. Boston: Allyn & Bacon.
- Newton R. R., & Rudestam K. E. (1999). *Your statistical consultant: Answers to your data analysis questions*. Thousand Oaks, CA: Sage.
- Nicol, A. A. M., & Pexman, P. M. (1999). *Presenting your findings: A practical guide for creating tables*. Washington, DC: American Psychological Association.
- Nicol, A. A. M., & Pexman, P. M. (2003). *Displaying your findings: A practical guide for creating figures, posters, and presentations*. Washington, DC: American Psychological Association.
- Rudestam, K. E., & Newton, R. R. (2000). *Surviving your dissertation: A comprehensive guide to content and process* (2nd ed.). Newbury Park, CA: Sage.
- Salant, P., & Dillman, D. D. (1994). *How to conduct your own survey*. New York: Wiley.
- SPSS. (2003). *SPSS 12.0: Brief guide*. Chicago: Author.
- Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Thousand Oaks, CA: Sage.
- Vogt, W. P. (1999). *Dictionary of statistics and methodology* (2nd ed.). Newbury Park, CA: Sage.
- Wainer, H. (1992). Understanding graphs and tables. *Educational Researcher*, 21(1), 14–23.
- Wilkinson, L., & The APA Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, 54, 594–604.

Index¹

Active independent variable, see Variables
Adjusted R^2 , 95-96, 103, 133
Alternate forms reliability, see Reliability
Analysis of covariance, see General linear model
ANOVA, 188, 197-198
ANOVA, see General linear model
Approximately normally distributed, 12, 13-14
Associational inferential statistics, 46-47, 53
 Research questions, 47-51, 53
Assumptions, 27-44, also see Assumptions for each statistic
Attribute independent variable, see Variables
Bar charts, see Graphs
Bar charts, 20, 38-39
Basic (or bivariate) statistics, 48-52
 Associational research questions, 49
 Difference research questions, 49
Bartlett's test of sphericity, 77, 82, 84
Between groups designs, 46
Between groups factorial designs, 47
Between subjects effects, 168-169, 173, see also Between groups designs
Binary logistic, 110, 115
Binary logistic regression, 109
Bivariate regression, 49-50, 53
Box plots, 18-20, 31-36, see also Graphs
Box's M, 120
Box's M, 123-124, 147, 165-168, 171, 173
Calculated value, 53
Canonical correlation, 52, 181-187
 Assumptions 182
 Writing Results, see Writing
Canonical discriminant functions, see Discriminate analysis
Case summaries, 190
Categorical variables, 15-16
Cell, see Data entry
Chart editor, 191
Charts, see Graphs
Chi-square, 49-50, 191
Cochran Q test, 50
Codebook, 191, 211-212
Coding, 24-26
Cohen's Kappa, see Reliability
Compare means, 188, see also t test and One-way ANOVA
Complex associational questions
 Difference questions, 49-51
Complex statistics, 48-51
Component Plot, 87-88
Compute variable, 43, 134-136, 203
Confidence intervals, 54-55

¹ Commands used by SPSS are in bold.

- Confirmatory factor analysis, 76
- Continuous variables, 16
- Contrasts, 136-140, 150**
- Copy and paste cells – see Data entry**
 - Output – see Output**
 - Variable – see Variable**
- Copy data properties, 192**
- Correlation, 192-193**
- Correlation matrix, 82
- Count, 192
- Covariate, 2
- Cramer's V, 50, 191
- Create a new file – see Data entry
 - Syntax – see Syntax
- Cronbach's alpha, see Reliability
- Crosstabs. 191**
- Cut and paste cells – see Data entry**
 - Variable – see Variable**
- Cummulative percent, 38
- d*, 55
- Data, see Data entry
- Data entry
 - Cell, 190
 - Copy and paste, 191, 193
 - Cut and paste, 191
 - Data, 193
 - Enter, 193, 195
 - Export, 193
 - Import, 193
 - Open, 193, 195
 - Print, 193
 - Save, 194, 196
 - Split, 196
 - Restructure, 201
- Data reduction, 77, 84, see also Factor analysis
- Data transformation, 42-44, see also Tranform
- Data View, 10, 148**
- Database information display, 194
- Define labels, see Variable label**
- Define variables – see Variables**
- Delete output – see Output
- Dependent variables, 48, also see Variables
- Descriptive research questions – see Research questions
- Descriptives, 29-31, 36-37, 191-192, 194**
- Descriptive statistics, 18, 29-31, 36-37
- Design classification, 46-47
- Determinant, 77, 82, 84
- Dichotomous variables, 13-15, 20, 36-37
- Difference inferential statistics, 46-47, 53
 - Research questions, 46-53
- Discrete missing variables, 15
- Discriminant analysis, 51, 109, 118-127
 - Assumptions, 119
 - Writing Results, see Writing
- Discussion, see Writing
- Dispersion, see Standard deviation and variance

- Display syntax (command log) in the output, see Syntax**
- Dummy coding, 24, 91**
- Edit data, see Data entry**
- Edit output, see Output**
- Effect size, 53-58, 96, 103, 130, 133-134, 143, 150, 164, 168-169, 172, 175**
- Eigenvalues, 82**
- Enter (simultaneous regression), 91**
- Enter (edit) data, see Data entry**
- Epsilon, 152**
- Equivalent forms reliability, see Reliability**
- Eta, 49-50, 53, 132, 167-168, 172**
- Exclude cases listwise, 192-193**
- Exclude cases pairwise, 192-193**
- Exploratory data analysis, 26-27, 52**
- Exploratory factor analysis, 76-84**
 - Assumptions, 76-77
 - Writing Results, see Writing
- Explore, 32-36, 194**
- Export data, see Data entry**
- Export output to MsWord, see Output**
- Extraneous variables, see Variables**
- Factor, 77, 84**
- Factor analysis, see Exploratory factor analysis**
- Factorial ANOVA, see General linear model**
- Figures, 213, 224-225**
- Files, see SPSS data editor and Syntax**
 - Data, 195
 - Merge, 195
 - Output, 195
 - Syntax, 195-196
- File info, see codebook**
- Filter, 190**
- Fisher's exact test, 196**
- Format output, see Output**
- Frequencies, 18-19, 29, 37-38, 196**
- Frequency distributions, 12-13, 20**
- Frequency polygon, 20, 40**
- Friedman test, 50, 147, 154-157**
- General linear model, 52-53**
 - Analysis of covariance (ANCOVA), 51, 141-146
 - Assumptions, 141
 - Writing Results, see Writing
 - Factorial analysis of variance (ANOVA), 49-51, 53, 129-140, 188
 - Assumptions, 129
 - Post-hoc analysis, 134-140
 - Writing Results, see Writing
 - Multivariate, see Multivariate analysis of variance
 - Repeated measures, see Repeated measures ANOVA
- GLM, see General linear model**
- Graphs**
 - Bar charts, 189
 - Boxplots, 189
 - Histogram, 197
 - Interactive charts/graph, 197
 - Line chart, 198
- Greenhouse-Geisser, 152, 159**

- Grouping variable, 3**
- Help menu, 196-197**
- Hierarchical linear modeling (HLM), 52**
- High school and beyond study, 5-6**
- Hide results within an output table, see Output**
- Histograms, 13, 20, 39, 197**
- Homogeneity-of-variance, 28, 119, 121, 124, 132, 138, 143-144, 192**
- HSB, see High school and beyond study**
- HSBdata file, 7-10**
- Import data, see Data entry**
- Independence of observations, 28, 147**
- Independent samples *t* test, 49-50, 53**
- Independent variable, see Variables**
- Inferential statistics**
 - Associational, 5, 46
 - Difference, 5, 46
 - Selection of, 47
- Insert cases, 189-190**
 - Text/title to output, see Output
 - Variable, see Variable
- Interactive chart/graph, see Graphs**
- Internal consistency reliability, see Reliability**
- Interquartile range, 19-20**
- Interrater reliability, see Reliability**
- Interval scale of measurement, 13-14, 16-17**
- Kappa, see Reliability**
- Kendall's tau-b, 49, 197**
- KMO test, 77, 82, 81**
- Kruskal-Wallis test, 50, 197**
- Kurtosis, 21-22**
- Label variables, see Variables**
- Layers, 197-198**
- Levels of measurement, 13**
- Levene's test, 131, 138-140, 166, 172-173**
- Line chart, see Graph**
- Linearity, 28, 197-198**
- Log, see Syntax**
- Logistic regression, 51, 109-114**
 - Assumptions, 109-110
 - Hierarchical, 114-118
 - Writing Results, see Writing
- Loglinear analysis, 49-51**
- Mann-Whitney U, 50, 198**
- MANOVA, see Multivariate analysis of variance**
- Matrix scatterplot, see Scatterplot**
- Mauchly's test of sphericity, 152, 177**
- McNemar test, 50**
- Mean, 198-199**
- Mean, 18-20**
- Measures of central tendency, 18-20**
 - Of variability, 19-20
- Median, 18-20**
- Merge, 195**
- Methods, see Writing**
- Missing values, 199**
- Mixed ANOVAs, see Repeated measures ANOVA**

- Mixed factorial designs, 47, 147
- Mode, 18-20
- Move variable, see Variable
- Multicollinearity, 91-104
- Multinomial logistic regression, 109
- Multiple regression, 51, 53, 198
 - Adjusted R^2 , 95-96, 103
 - Assumptions, 91
 - Block, 105
 - Hierarchical, 92, 104-107
 - Model summary table, 96, 103, 107
 - Simultaneous, 91-104
 - Stepwise, 92
 - Writing Results, see Writing
- Multivariate analysis of variance, 50-51, 162-181
 - Assumptions, 162
 - Mixed analysis of variance, 174-181
 - Assumptions, 175
 - Single factor, 162-169
 - Two factor, 169-174
 - Writing Results, see Writing
- Multivariate analysis of covariance, 51
- Nominal scale of measurement, 13-14, 15, 17, 19-20, 38-39
- Non experimental design, 2
- Nonparametric statistics, 19, 27
 - K independent samples, 50, 197
 - K related samples
 - Two independent samples, 50, 198
 - Two related samples, 205
- Normal, see Scale
- Normal curve, 12, 20-22
- Normality, 28
- Normally distributed, 13, 20-22
- Null hypothesis, 54
- One-way ANOVA, 50, 53
- Open data, see Data entry
 - File, see File
 - Output, see Output
- Ordinal scale of measurement, 13-14, 16, 17, 19-20, 29
- Outliers, 33
- Output, 194-196
 - Copy and paste, 199
 - Create, 195
 - Delete, 200
 - Display syntax, see syntax
 - Edit, 200
 - Export to MsWord, 200
 - Format, 200
 - Hide, 200
 - Insert, 200
 - Open, 195, 201
 - Print, 201
 - Print preview, 201
 - Resize/rescale, 201
 - Save, 196, 201
- Paired samples t test, 49-50

- Parallel forms reliability, see Reliability
- Parametric tests, 19, 27
- Pearson correlation, 49-50, 192, see also **Correlate**
- Phi, 49, 191
- Pillai's trace, 168
- Pivot tables, 201
- Power, 19
- Practical significance, 53-58
- Principal components analysis, 52, 76
 - Assumptions, 76-77
- Print data, see **Data entry**
 - Output, see **Output**
 - Preview, see **Output**
 - Syntax, see **Syntax**
- Profile plots, 175-180
- Randomized experimental designs, 2
- Range, 19-20
- Ranks, 16
- Ratio scale of measurement, 16-17
- Recode into different variable, see **Variable**
- Recode into same variable, see **Variable**
- Regression, see **Multiple regression**
- Reliability 64-71, 192
- Reliability
 - Alternate forms, see **equivalent forms**
 - Assumptions, 63-64
 - Cohen's Kappa, 63, 172-174, 191-192
 - Cronbach's alpha, 52, 63-71, 192
 - Equivalent forms, 63
 - Internal consistency, 63-71
 - Interrater, 63, 72-74
 - Test-retest, 63, 71-72
 - Writing Results, see **Writing**
- Regression
 - Linear regression, 198
 - Multiple regression, see **Multiple regression**
- Repeated measures ANOVA, 50-51, 147-154
 - Assumptions, 147
 - Writing Results, see **Writing**
- Repeated measures designs, 46
- Repeated measures, mixed ANOVA, 157-160
 - Assumptions, 147-148
 - Writing Results, see **Writing**
- Replace missing values, see **Missing values**
- Research problems, 6
- Research questions/hypotheses, 3
 - Basic associational, 4, 6
 - Basic descriptive, 4, 6
 - Basic difference, 3, 4, 6
 - Complex associational, 5, 6
 - Complex descriptive, 5, 6
 - Complex difference, 5, 6
 - Types, 6
- Resize/rescale output, see **Output**
- Restructure data wizard, see **Data entry**
- Results, see **Writing**

- Results coach, 201**
- Rotated component matrix, see Factor analysis
- Run command, see Syntax**
- Saving
 - Data file, see Data entry
 - Output, see Output
 - Syntax/log, see Syntax
- Scale, scale of measurement, 13-14, 16-17, 20, 29, 148
- Scatterplots, 201-202
 - Matrix, 40-42, 97-103
 - Residual, 102-103
- Scree plot, 85-88
- Select cases, 190**
- Selection of inferential statistics, see Inferential statistics
- Sig., see Statistical significance**
- Significance, see Statistical significance
- Single factor designs, 47
- Skewness, 13, 20, 22, 27, 29
- Sort cases, 190**
- Spearman rho, 49, 192-193, 202
- Split file, 190, 196**
- SPSS program
 - Open and start, 206
- SPSS data editor, 194,208**
- Standard deviation, 19-20
- Standardized beta coefficients, 96
- Statistical assumptions, see Assumptions
- Statistical significance, 53-54, 96
- Statistics coach, 202**
- Stem-and-leaf plot, 31, 34-36, 202
- Summarize, 202, see also Descriptive statistics
- Summated scale. 63
- Syntax, 194-196
 - Create, 195, 202
 - Display in output, 200, 209
 - Open, 196
 - Print, 203
 - Run, 202
 - Save, 196, 203
- System missing, 8**
- t* test, 203
- Tables, 203, 213-224**
- Test-retest, see Reliability
- Tolerance, 95-97, 103, 110
- Transform, 43, 192, 199, 203**
- Transform data, see Data transformation
- Two-way ANOVA, see General linear model ANOVA
- Valid *N*, 33
- Values, 8
- Variable label, 3, 209-210**
 - View, 7-8, 148, 208
- Variables, 1, 48, see also Compute
 - Copy and paste, 204
 - Cut and paste, 204
 - Definition, 1
 - Dependent variables, 3, 7, 48, 148

- Extraneous variables, 3, 7
- Independent variables, 1, 7
 - Active, 1, 2
 - Attribute, 2
 - Values, 2
- Information, 204
- Move, 205
- Recode, 205
- Varimax rotation, see Factor analysis
- VIF, 95, 103, 110-114
- Wilcoxon signed-ranks test, 50, 205
- Wilks lambda, 152, 166-168, 183
- Within-subjects designs, see Repeated measures designs
- Within-subjects effects, 152
- Within-subjects factor/variable, 148
- Writing Results
 - Analysis of covariance (ANCOVA), 145-146
 - Canonical correlation, 186-187
 - Discriminant analysis, 126-127
 - Factor analysis, 83
 - Factorial analysis of variance (ANOVA) , 140
 - Friedman, 157
 - Logistic regression, 114
 - Mixed analysis of variance (ANOVA), 160
 - Mixed multivariate analysis of variance (ANOVA), 181
 - Reliability, 67, 72, 74
 - Repeated measures analysis of variance (ANOVA), 153-154
 - Single factor multivariate analysis of variance (MANOVA), 169
 - Simultaneous multiple regression, 103-104
 - Two factor multivariate analysis of variance (MANOVA), 181
- z scores, 22, 205