Quantitative Research in Education
Intermediate & Advanced Methods

DIMITER M. DIMITROV
George Mason University

Whitier Publications, Inc.
New York
This book offers a comprehensive presentation of quantitative research design and statistical methods in the context of education and related fields. The text is intended primarily for use by students who take intermediate and advanced quantitative research courses as a part of their graduate degree program, but it can be a useful resource for researchers in education, counseling, rehabilitation, psychology, sociology, social work, and human development as well.

The main purpose of this book is to provide the readers with an in-depth conceptual and methodological understanding of intermediate and advanced quantitative research methods, as well as the skills necessary to apply such methods using SPSS and to interpret the results. This is achieved by building layers of context-based understanding of research concepts and methods, their statistical translation, methodological principles, computer-based data analysis, presentation of the results in APA style format, and contextual interpretations. The text allows people who experience difficulties with analytic representations of statistical concepts to capitalize on conceptual understanding and still be able to master the research tools necessary for their work on theses, dissertations, and professional research.

While there are many excellent introductory books on research design and statistics in education and the social sciences, most books at the intermediate and advanced levels tend to be either too technical and mathematical or too simplistic. Typically, claiming to have an "applied orientation," such books are dominated by presentations of SPSS dialog boxes and printouts at the expense of theoretical and methodological rigor. To bridge the gap between these extremes, this book attempts to provide a balance between conceptual meaning and its statistical translation by developing understanding and application skills in a spiral exposure to quantitative concepts and methods. For example, the comparison of groups on variables of interest is addressed in a sequence from univariate cases of $t$-tests, nonparametric methods, and analysis of variance (ANOVA) to scenarios illustrating the use of multivariate analysis of variance (MANOVA) and structural equation modeling (SEM). As another example, the concept of validity is addressed in the framework of measurement, research design, and structural equation modeling. Particular attention is devoted to potential problems associated with violation of assumptions, common misconceptions (e.g., conducting MANOVA versus separate ANOVAs), effect sizes, confidence intervals, and sample size. The book is organized in four parts comprising 24 chapters. Each chapter ends with a summary and study questions.

Part I [Measurement in Educational Research] consists of three chapters. Chapter 1 presents variables and measurement scales in the context of education. The focus is on the nature of measurement in education, types of variables, types of scales and their transformations, permissible arithmetic operations with scale values, summation symbols, and basic rules of summation. Chapter 2 introduces the classical model of reliability of scores, types of reliability, and reliability of composite scores. Chapter 3 deals with the concept of validity for measurement instruments (e.g., tests, questionnaires, or inventories) and types of validity (content-related validity, criterion-related validity, and construct-related validity).

Part II [Research Design] consists of two chapters. Chapter 4 deals with research problems, hypotheses, and types of quantitative research: nonexperimental research, experimental research, and threats to internal and external validity. Chapter 5 presents pre-experimental and true experimental research designs that involve quantitative methods of data analysis. The focus is primarily on conceptual understanding and methodological principles underlying the application of such designs in educational research.
Part III [Univariate Statistics in Educational Research] consists of fourteen chapters. The first five of these chapters (6, 7, 8, 9, and 10) cover introductory statistics and prepare the ground for understanding and practical applications of intermediate statistics in educational research. The next six chapters (11 through 16) provide intermediate treatment of correlation, regression, and analysis of variance (ANOVA) including some nonparametric methods. The last three chapters in this section (17, 18, and 19) provide more advanced treatment of multiple regression, analysis of variance, and the relations between them.

Part IV [Multivariate Statistics in Educational Research] consists of five chapters. This part covers the topics of logistic regression, multivariate analysis of variance (MANOVA), exploratory factor analysis, confirmatory factor analysis, and elements of structural equation modeling (SEM). The analytic framework of these topics is simplified and tailored to conceptual understanding, computer-aided applications, and interpretations in the context of educational research.

Supplements
Data sets for computer-based applications in examples using SPSS can be downloaded from the online supplement to this book [http://cehd.gmu.edu/book/dimitrov]. This supplement provides also (a) answers to the study questions for each chapter, (b) addendum to some topics discussed in the book, (c) syntax for confirmatory factor analysis, path analysis, and group comparison on latent variables in the framework of major computer programs — LISREL, AMOS, EQS, and Mplus [used for illustrations in Chapters 23 and 24], and (d) additional references (books, articles, and online products) related to the content of this book.

Acknowledgments
I would like to acknowledge the assistance of my graduate research assistant, Jill Lammert, who provided valuable editorial comments and suggestions throughout the writing of this book. I also appreciate the feedback and encouragement of colleagues from different universities as well as graduate students who used draft chapters of this book in quantitative research courses that they took with me in the Graduate School of Education at George Mason University.

Dimiter M. Dimitrov
Contents

Preface xv

PART I MEASUREMENT IN EDUCATION 1

Chapter 1
VARIABLES AND MEASUREMENT SCALES 3
1.1 Variables in Educational Research 3
  1.1.1 Observable versus Latent Variables 4
  1.1.2 Continuous versus Discrete Variables 4
1.2 Scales of Measurement 5
  1.2.1 What is Measurement? 5
  1.2.2 Nominal Scale 6
  1.2.3 Ordinal Scale 6
  1.2.4 Interval Scale 6
  1.2.5 Ratio Scale 7
  1.2.6 Scale Transformations and Operations 7
  1.2.7 Scaling of Individual Items 9
1.3 Symbols and Rules for Summation of Variables 9
  1.3.1 Symbolic Notations 9
  1.3.2 Summation Operator 10
1.4 Summary 11
1.5 Study Questions 13

Chapter 2
RELIABILITY 15
2.1 What is Reliability? 15
2.2 Classical Concept of Reliability 16
  2.2.1 True Score 16
  2.2.2 Definition of Reliability 16
  2.2.3 Standard Error of Measurement 17
2.3 Types of Reliability 18
  2.3.1 Internal Consistency Reliability 18
  2.3.2 Test-retest Reliability 19
  2.3.3 Alternate Forms Reliability 20
  2.3.4 Criterion-referenced Reliability 20
  2.3.5 Interrater Reliability 22
2.4 Reliability of Composite Scores 23
  2.4.1 Reliability of Sum of Scores 23
  2.4.2 Reliability of Difference of Scores 24
2.5 Reliability Estimation with SPSS 25
   2.5.1 Calculation of Cronbach’s alpha  25
   2.5.2 Calculation of Cohen’s kappa 26
2.6 Summary  27
2.7 Study Questions  28

Chapter 3
VALIDITY 29
3.1 What is Validity? 29
3.2 Aspects of Construct Validity 30
   3.2.1 Content Aspect of Validity 30
   3.2.2 Substantive Aspect of Validity 31
   3.2.3 Structural Aspect of Validity 31
   3.2.4 Generalizability Aspect of Validity 32
   3.2.5 External Aspect of Validity 33
   3.2.6 Consequential Aspect of Validity 34
3.3 Summary  35
3.4 Study Questions  35

PART II RESEARCH DESIGN 37

Chapter 4
QUANTITATIVE RESEARCH 39
4.1 Research Questions and Hypotheses 39
4.2 Types of Quantitative Research 41
   4.2.1 Nonexperimental Research 41
      4.2.1.1 Descriptive research 41
      4.2.1.2 Correlational research 42
      4.2.1.3 Ex post facto research 43
      4.2.1.4 Meta-analysis research 43
   4.2.2 Experimental Research 45
      4.2.2.1 True experimental research 45
      4.2.2.2 Quasi-experimental research 46
      4.2.2.3 Single-case research 46
   4.2.3 Internal and External Validity in Experimental Research 47
      4.2.3.1 Threats to internal validity 47
      4.2.3.2 Threats to external validity 48
4.3 Summary  50
4.4 Study Questions  51
Chapter 5

**BASIC RESEARCH DESIGNS** 53

5.1 Pre-experimental Designs 53
   5.1.1 One Group Posttest-Only Design 53
   5.1.2 One Group Pretest-Posttest Design 54
   5.1.3 Nonrandomized Control Group Posttest-Only Design 54

5.2 True Experimental Designs 55
   5.2.1 Randomized Pretest-Posttest Control Group Design 55
   5.2.2 Randomized Solomon Four-Group Design 56
   5.2.3 Randomized Control-Group Posttest Only Design 57

5.3 Quasi-Experimental Designs 58
   5.3.1 Nonrandomized Pretest-Posttest Control Group Design 58
   5.3.2 One Group Time-Series Design 59
   5.3.3 Control Group Time-Series Design 60

5.4 Summary 61
5.5 Study Questions 63

Part III UNIVARIATE DATA ANALYSIS 65

Chapter 6

**REVIEW OF INTRODUCTORY STATISTICS** 67

6.1 Organizing and Graphing Data 67
   6.1.1 Frequency Table 67
   6.1.2 Basic Distribution Graphs 68

6.2 Describing Distributions 72
   6.2.1 Percentiles 72
   6.2.2 Measures of Central Tendency 74
     6.2.2.1 Mode 74
     6.2.2.2 Median 74
     6.2.2.3 Mean 74
     6.2.2.4 Properties of the mode, median, and mean 75
   6.2.3 Measures of Variation 75
     6.2.3.1 Variance 76
     6.2.3.2 Standard deviation 76
     6.2.3.3 Pooled variance 77
     6.2.3.4 Some basic rules 77
   6.2.4 Standard Scores 78
   6.2.5 Scale Transformation 79

6.3 Summary 80
6.4 Study Questions 82
Chapter 7

BASIC DISTRIBUTIONS 83

7.1 Normal Distribution 83
  7.1.1 What is a Normal Distribution? 83
  7.1.2 Basic Properties of the Normal Distribution 85
  7.1.3 Determining Percentiles and Percentile Ranks 86
  7.1.4 Sampling Distribution of the Mean 87
  7.1.5 Normal Q-Q Plot 89
7.2 Student’s t-Distribution 91
7.3 F-Distribution 92
7.4 Chi-square Distribution 93
7.5 Summary 94
7.6 Study Questions 95

Chapter 8

HYPOTHESIS TESTING 97

8.1 What is Hypothesis Testing? 97
8.2 When To Reject (or Not) the Null Hypothesis? 98
8.3 Testing Hypotheses about the Mean 100
  8.3.1 One-sample Case for the Mean 100
  8.3.2 Two-sample Case for the Mean: Independent Samples 106
  8.3.3 Two-sample Case for the Mean: Dependent Samples 111
8.4 Summary 113
8.5 Study Questions 114

Chapter 9

HYPOTHESIS TESTING FOR PROPORTIONS 117

9.1 One-Sample Case for Proportion 117
9.2 Testing H0: \( P_1 = P_2 \) for Independent Samples 122
9.3 Testing H0: \( P_1 = P_2 \) for Dependent Samples 127
9.4 Summary 131
9.5 Study Questions 132

Chapter 10

CORRELATION AND SIMPLE REGRESSION 135

10.1 Correlation between Two Variables 135
  10.1.1 What is Linear Relationship (Correlation) between Two Variables? 135
  10.1.2 The Pearson Product-Moment Correlation Coefficient 138
10.2 Simple Linear Regression 144
  10.2.1 Correlation, Prediction, and Causation 144
  10.2.2 The Regression Line 144
10.2.3 Interpretation of the Slope 147
10.2.4 Conditional Distributions of Y-scores 149
10.2.5 Assumptions with Simple Linear Regression 152
10.3 Summary 154
10.4 Study Questions 155

Chapter 11
PARTIAL AND PART CORRELATION 159
11.1 Partial Correlation 159
11.2 Part Correlation 163
11.3 Summary 165
11.4 Study Questions 166

Chapter 12
NONPARAMETRIC TESTS 167
12.1 The Man-Whitney U Test 167
12.2 The Wilcoxon Signed-Rank Test for Dependent Samples 170
12.3 Chi-Square Goodness-of-fit Test 172
12.4 Chi-Square Test for Association 177
12.5 Summary 182
12.6 Study Questions 183

Chapter 13
MULTIPLE REGRESSION 185
13.1 The Concept of Multiple Regression 185
13.2 Comparison of Full and Restricted Regression Models 191
13.3 Multicollinearity 194
13.4 Cross-validation 197
13.5 Statistical Power, Effect Size, and Sample Size 197
13.6 Outliers and Influential Data Points 199
13.7 Categorical Predictors in Multiple Regression 201
13.8 Interaction between Predictors in Multiple Regression 203
13.8.1 What is Interaction between Predictors? 203
13.8.2 Testing for Interaction between Predictors 206
13.8.3 Centering Predictors 207
13.9 Selection of Predictors in Multiple Regression 209
13.10 APA Style for Multiple Regression Results 210
13.11 Summary 211
13.12 Study Questions 214
Chapter 14

ONE-FACTOR ANALYSIS OF VARIANCE 217

14.1 The Concept of One-Factor Analysis of Variance 218
14.2 Assumptions in ANOVA 219
14.3 Effects in One-factor ANOVA 220
14.4 Within-groups and Between-groups Variance 221
14.5 Linear Model for One-factor ANOVA 223
14.6 Testing the ANOVA Null Hypothesis 223
14.7 Multiple Comparisons 225
   14.7.1 Post Hoc Comparisons 225
      14.7.1.1 The Tukey method of multiple comparisons 225
      14.7.1.2 Bonferroni method of multiple comparisons 226
   14.7.2 Planned Comparisons 227
      14.7.2.1 Contrasts for planned multiple comparisons 227
      14.7.2.2 Dunnett method of multiple comparisons 228
14.8 Determining Effect Size 230
   14.8.1 Effect Size of Mean Differences 230
   14.8.2 Omnibus Effect Size 230
14.9 Determining the Sample Size 234
14.10 Consequences of Violating the ANOVA Assumptions 234
14.11 Interpretation of SPSS Output for One-factor ANOVA 236
14.12 Summary 238
14.13 Study Questions 239

Chapter 15

TWO- AND THREE-FACTOR ANOVA 241

15.1 Two-factor ANOVA 241
   15.1.1 Null Hypotheses in Two-factor ANOVA 241
   15.1.2 Assumptions in Two-factor ANOVA 242
   15.1.3 Effects in Two-factor ANOVA 243
   15.1.4 Linear Model for the Data in Two-factor ANOVA 245
   15.1.5 Sum of Squares in Two-factor ANOVA 246
   15.1.6 Mean Squares in Two-factor ANOVA 247
   15.1.7 Testing the Null Hypotheses in Two-factor ANOVA 248
   15.1.8 Omnibus Effect Size in Two-factor ANOVA 249
   15.1.9 Types of Interaction in Two-factor ANOVA 251
   15.1.10 Testing for Simple Main Effects 252
   15.1.11 Using SPSS for Two-factor ANOVA 252
15.2 Three-factor ANOVA 257
15.3 Summary 264
15.4 Study Questions 265
Chapter 16

ANALYSIS OF COVARIANCE 267

16.1 The Logic Behind ANCOVA 267
16.1.1 Basic Concepts in ANCOVA 267
16.1.2 Adjusted Group Means in ANCOVA 268
16.1.3 Increased Test Power with ANCOVA 270
16.1.4 Assumptions in ANCOVA 271
16.2 Performing ANCOVA and Interpreting the Results 271
16.3 ANCOVA versus ANOVA on Gain Score 277
16.4 Summary 279
16.5 Study Questions 280

Chapter 17

MULTIPLE REGRESSION AND ANOVA 281

17.1 One-Factor ANOVA via Multiple Regression 281
17.1.1 Contrast Coding for ANOVA with Two Groups 281
17.1.2 Contrast Coding for One-factor ANOVA with Three Groups 282
17.1.3 Orthogonal Contrasts 284
17.2 Two-Factor ANOVA via Multiple Regression 286
17.3 Summary 293
17.4 Study Questions 293

Chapter 18

ANOVA WITH RANDOM FACTORS 295

18.1 ANOVA with One Random Factor 295
18.1.1 Random Effects 295
18.1.2 Assumptions of the Random-factor ANOVA 296
18.1.3 Expected Mean Square in the Random-factor ANOVA 296
18.1.4 The Primary Question in a Random-factor ANOVA 297
18.2 Two-factor Mixed-Effects ANOVA Model 299
18.2.1 The Concept of a Mixed-effects Model 299
18.2.2 Assumptions of the Two-factor Mixed ANOVA Model 300
18.2.3 Expected Mean Square in the Two-factor Mixed ANOVA 301
18.2.4 Effect Size of Mean Differences among Levels of the Fixed Factor 304
18.2.5 Generalizations with the Two-factor Mixed ANOVA 305
18.3 Summary 306
18.4 Study Questions 307

Chapter 19

REPEATED-MEASURES ANOVA 309

19.1 A Simple Repeated-Measures ANOVA 309
## PART IV  MULTIVARIATE DATA ANALYSIS  329

### Chapter 20  
**LOGISTIC REGRESSION**  331

20.1 The Concept of Logistic Regression  331  
20.1.1 Probability, Odds, and Odds Ratio  331  
20.1.2 The Logistic Model  333  
20.1.3 Logit Form of the Logistic Regression Model  334  
20.1.4 Interpretation of the Regression Coefficients  334  
20.2 Tests and Interpretations of Logistic Regression Results  336  
20.2.1 Goodness-of-fit Tests  336  
20.2.2 Hosmer-Lemeshow Goodness-of-fit Test  337  
20.2.3 Test for Significance of Predictor Variables  338  
20.2.4 Effect Size Information with Logistic Regression  339  
20.2.5 Classification Table  341  
20.3 Coding Categorical Predictors  342  
20.4 Using SPSS for Binary Logistic Regression  343  
20.5 Comparison of Full and Restricted Models  347  
20.6 Selection of Predictors in Logistic Regression  348  
20.7 Assumptions in Logistic Regression  349  
20.8 Summary  349  
20.9 Study Questions  351

### Chapter 21  
**MULTIVARIATE ANALYSIS OF VARIANCE**  353

21.1 The Concept of MANOVA  353  
21.2 MANOVA versus Separate ANOVAs  354  
21.3 When to Use Separate ANOVAs?  354  
21.4 When to Use MANOVA?  355  
21.5 Assumptions in MANOVA  355  
21.6 MANOVA with Discriminant Analysis  356  
21.7 MANOVA with Planned Comparisons  361  
21.8 Sample Size in MANOVA  366
EDA in Education. Quantitative research requires statistical modeling and statistical modeling requires EDA. There are studies on student performance (particularly international comparative studies like the Programme for International Student Assessment (PISA)). There are investigations of the effects of social background and school type. Our PhD program in quantitative research methods in education (QRME) develops researchers, scholars, and policy leaders who engage in traditions of inquiry that create knowledge and understanding founded in empirical evidence. The program focuses on: Building a strong understanding of quantitative methods in research. Studying causal and complex relationships within applied educational, social, and institutional settings. Using research to inform policy and practice in applied educational and social settings. Ideal QRME PhD students: Understand the importance of evidence and data in making dec