# CONTENTS

## CHAPTER 1  AN INTRODUCTION TO CONTEMPORARY EDUCATIONAL TESTING AND MEASUREMENT 1

Tests are Only Tools; Their Usefulness Can Vary 2  
Why We Developed This Text: Enhancing Test Usefulness 3  
Technical Adequacy 3  
Test User Competency 4  
Matching the Test’s Intended Purpose 4  
Matching Diverse Test-Takers to the Test 5  
Test Results and Diversity Considerations 6  
Tests are Only Tools: A Video Beats a Photo 7  
The Difference Between Testing/Assessment and The Assessment Process 8  
Tests and Assessments 9  
Assessment Process 9  
Types of Tests/Assessments 10  
Objective, Essay, and Performance/Portfolio Tests/Assessments 11  
Teacher-Made and Standardized Tests 12  
Norm-Referenced Tests (NRTs) and Criterion-Referenced Tests (CRTs) 12  
Curriculum-Based Measurements (CBMs) 13  
Effects on the Classroom Teacher 13  
About the Text 15  
What If You Are “No Good in Math”? 16  
Summary 16  
For Discussion and Practice 17  

## CHAPTER 2  NATIONAL DEVELOPMENTS: IMPACT ON CLASSROOM TESTING AND MEASUREMENT 19  

Education Reform 21  
Regular Education Reform 21  
Special Education Reform 23  
Merging Regular and Special Education Reform: IDEIA and NCLB 24  
Standards-Based Reform 25  
Race to the Top (RTT) 27  
The Future: NCLB, RTT, CCSS, and the CCSS-Aligned Tests 28  
Other Trends: Computer-Adaptive Achievement Testing 29  
Globalization and International Competitiveness 30  
Competency Testing for Teachers 31  
Teacher Evaluation Based on Student Test Scores: Value-Added Models (VAM) 32  
Increased Interest from Professional Groups 32  
A Professional Association–Book Publisher Information Initiative 33  
Summary 33  
For Discussion and Practice 35  

## CHAPTER 3  RESPONSE TO INTERVENTION (RTI) AND THE REGULAR CLASSROOM TEACHER 36  

What Is RTI? 37  
What If You Have Not Heard of RTI Before? 37  
How New Is RTI? 38  
Why Do Regular Education Teachers Need to Know About RTI? 38  
An RTI Scenario 38  
How Important Is RTI to Regular Education Teachers? 40  
How Is RTI Supposed to Help Students and Schools? 41  
RTI Definitions, Components, and Approaches 42  
RTI Definitions 42  
RTI Components 43  
RTI Approaches 48  
How Widely Is RTI Being Implemented? 50  
Some Benefits of RTI 51  
RTI: The Promise and Some Controversies 52  
Technical Issues: Reliability, Validity, and Fairness 52  
Implementation Issues 52  
Summary 53  
For Discussion and Practice 54  

## CHAPTER 4  TESTING AND EDUCATIONAL DECISION MAKING 56  

Testing, Accountability, and The Classroom Teacher 57  
Special Learners, the Regular Curriculum, and Annual Assessments 58  
Types of Educational Decisions 59  
A Pinch of Salt 62  
“Pinching” in the Classroom 63  
What to Measure 64  
How to Measure 65  
Written Tests 65  
Summary 66  
For Discussion and Practice 67  

ix
| CONTENTS |
|-----------------|-----------------|-----------------|-----------------|
| **CHAPTER 5**  NORM-REFERENCED AND CRITERION-REFERENCED TESTS AND CONTENT VALIDITY EVIDENCE 68 | Defining Norm-Referenced (NRT) and Criterion-Referenced (CRT) Tests 69 | Comparing NRTs and CRTs 72 | Differences in The Construction of NRTs and CRTs 73 |
|                              | NRTs, CRTs, and Linguistic and Cultural Diversity 74 | NRTs, CRTs, and Validity Evidence 75 | A Three-Stage Model of Classroom Measurement 76 |
|                              | Are There Different Kinds of Goals and Objectives? 81 | How Can Instructional Objectives Make a Teacher’s Job Easier? 83 | Where Do Goals Come From? 79 |
|                              | For Discussion and Practice 86 | For Discussion and Practice 86 | For Discussion and Practice 86 |
| **CHAPTER 6**  MEASURING LEARNING OUTCOMES 87 | Writing Instructional Objectives 88 | Identifying Learning Outcomes 88 | Identifying Observable and Directly Measurable Learning Outcomes 89 |
|                              | Stating Conditions 90 | Stating Criterion Levels 90 | Keeping It Simple and Straightforward 91 |
|                              | Matching Test Items to Instructional Objectives 92 | Taxonomy of Educational Objectives 94 | Cognitive Domain 94 |
|                              | Affective Domain 97 | The Psychomotor Domain 99 | The Psychomotor Domain 99 |
|                              | The Test Blueprint 100 | Content Outline 102 | Categories 102 |
|                              | Number of Items 102 | Functions 102 | Number of Items 102 |
|                              | Summary 104 | For Discussion and Practice 104 | For Discussion and Practice 104 |
| **CHAPTER 7**  WRITING OBJECTIVE TEST ITEMS 106 | Which Format? 107 | True–False Items 108 | True–False Items 108 |
|                              | Suggestions for Writing True–False Items 110 | Matching Items 111 | Matching Items 111 |
|                              | Faults Inherent in Matching Items 111 | Suggestions for Writing Matching Items 114 | Suggestions for Writing Matching Items 114 |
|                              | Higher-Level Multiple-Choice Questions 114 | Suggestions for Writing Multiple-Choice Items 123 | Higher-Level Multiple-Choice Questions 114 |
|                              | Completion Items 123 | Suggestions for Writing Completion Items 125 | Completion Items 123 |
|                              | Gender and Racial Bias in Test Items 126 | Guidelines for Writing Test Items 127 | Guidelines for Writing Test Items 127 |
|                              | Advantages and Disadvantages of Different Objective Item Formats 128 | True–False Tests 128 | True–False Tests 128 |
|                              | Matching Tests 129 | Multiple-Choice Tests 129 | Matching Tests 129 |
|                              | Completion Tests 129 | Completion Tests 129 | Completion Tests 129 |
|                              | Summary 130 | For Discussion and Practice 130 | For Discussion and Practice 130 |
| **CHAPTER 8**  WRITING ESSAY TEST ITEMS 132 | What Is An Essay Item? 133 | Essay Items Should Measure Complex Cognitive Skills or Processes 134 | Essay Items: Extended or Restricted Response 135 |
|                              | Examples of Restricted Response Essays 136 | Pros and Cons of Essay Items 138 | Advantages of the Essay Item 138 |
|                              | Disadvantages of the Essay Item 138 | Suggestions for Writing Essay Items 139 | Disadvantages of the Essay Item 138 |
|                              | Scoring Essay Questions 140 | Scoring Extended Response and Higher-Level Questions 142 | Scoring Essay Questions 140 |
|                              | General Essay Scoring Suggestions 146 | Assessing Knowledge Organization 147 | General Essay Scoring Suggestions 146 |
|                              | Open-Book Questions and Exams 149 | Some Open-Book Techniques 149 | Open-Book Questions and Exams 149 |
|                              | Guidelines for Planning Essays, Knowledge Organization, and Open-Book Questions and Exams 153 | Summary 154 | Guidelines for Planning Essays, Knowledge Organization, and Open-Book Questions and Exams 153 |
|                              | For Discussion and Practice 155 | For Discussion and Practice 155 | For Discussion and Practice 155 |
|                              | Performance Tests Can Assess Affective and Social Skills 158 | Developing Performance Tests for Your Learners 160 | Developing Performance Tests for Your Learners 160 |
|                              | Step 1: Deciding What to Test 160 | Step 2: Designing the Assessment Context 162 | Step 1: Deciding What to Test 160 |
|                              | Step 3: Specifying the Scoring Rubrics 165 | Step 4: Specifying Testing Constraints 170 | Step 2: Designing the Assessment Context 162 |
|                              | A Final Word 171 | Summary 172 | A Final Word 171 |
|                              | For Discussion and Practice 173 | For Discussion and Practice 173 | For Discussion and Practice 173 |
CHAPTER 10  PORTFOLIO ASSESSMENT  174

Rationale for the Portfolio  175
Ensuring Validity of the Portfolio  176
Developing Portfolio Assessments  176
  Step 1: Deciding on the Purposes for a Portfolio  177
  Step 2: Identifying Cognitive Skills and Dispositions  177
  Step 3: Deciding Who Will Plan the Portfolio  177
  Step 4: Deciding Which Products to Put in the Portfolio and How Many Samples of Each Product  178
  Step 5: Building the Portfolio Rubrics  179
  Step 6: Developing a Procedure to Aggregate All Portfolio Ratings  184
  Step 7: Determining the Logistics  185
Summary  188
For Discussion and Practice  189

CHAPTER 11  ADMINISTERING, ANALYZING, AND IMPROVING THE TEST OR ASSESSMENT  190

Assembling the Test  191
  Packaging the Paper-and-Pencil Test  191
Administering the Test  193
Scoring the Test  195
Analyzing the Test  196
  Quantitative Item Analysis  196
  Qualitative Item Analysis  202
  Item Analysis Modifications for the Criterion-Referenced Test  203
Debriefing  207
  Debriefing Guidelines  208
The Process of Evaluating Classroom Achievement  209
Summary  210
For Discussion and Practice  212

CHAPTER 12  MARKS AND MARKING SYSTEMS  213

What Is the Purpose of a Mark?  213
  Why Be Concerned about Marking?  214
  What Should a Mark Reflect?  214
Marking Systems  215
  Types of Comparisons  216
  Types of Symbols  219
Combining and Weighting the Components of a Mark  221
  Who Is the Better Teacher?  221
Combining Grades into a Single Mark  223
Practical Approaches to Equating Before Weighting in the Busy Classroom  225
  Front-End Equating  226
  Back-End Equating  226
Summary  229
For Discussion and Practice  230

CHAPTER 13  SUMMARYING DATA AND MEASURES OF CENTRAL TENDENCY  231

What Are Statistics?  232
Why Use Statistics?  233
Tabulating Frequency Data  233
  The List  234
  The Simple Frequency Distribution  234
  The Grouped Frequency Distribution  235
Steps in Constructing a Grouped Frequency Distribution  237
Graphing Data  239
  The Bar Graph, or Histogram  240
  The Frequency Polygon  240
  The Smooth Curve  243
Measures of Central Tendency  246
  The Mean  246
  The Median  248
  The Mode  252
  The Measures of Central Tendency in Various Distributions  254
Summary  255
For Discussion and Practice  257

CHAPTER 14  VARIABILITY, THE NORMAL DISTRIBUTION, AND CONVERTED SCORES  258

The Range  259
The Semi-Interquartile Range (SIQR)  260
The Standard Deviation  261
  The Deviation Score Method for Computing the Standard Deviation  264
  The Raw Score Method for Computing the Standard Deviation  265
The Normal Distribution  267
  Properties of the Normal Distribution  268
Converted Scores  270
  z-Scores  273
  T-Scores  277
Summary  278
For Discussion and Practice  278

CHAPTER 15  CORRELATION  280

The Correlation Coefficient  282
  Strength of a Correlation  282
  Direction of a Correlation  282
Scatterplots  283
  Where Does r Come From?  285
Causality  286
Other Interpretive Cautions  287
Summary  289
For Discussion and Practice  290
### CHAPTER 16 VALIDITY EVIDENCE 292

- Why Evaluate Tests? 292
- Types of Validity Evidence 293
  - Content Validity Evidence 293
  - Criterion-Related Validity Evidence 294
  - Construct Validity Evidence 295
- What Have We Been Saying? A Review 296
- Interpreting Validity Coefficients 298
  - Content Validity Evidence 298
  - Concurrent and Predictive Validity Evidence 298
- Summary 302
- For Discussion and Practice 302

### CHAPTER 17 RELIABILITY 304

- Methods of Estimating Score Reliability 305
  - Test–Retest or Stability 305
  - Alternate Forms or Equivalence 306
  - Internal Consistency 306
- Interpreting Reliability Coefficients 309
- Summary 312
- For Discussion and Practice 313

### CHAPTER 18 ACCURACY AND ERROR 315

- Error—What Is It? 316
- The Standard Error of Measurement 317
  - Using the Standard Error of Measurement 318
  - More Applications 321
  - Standard Deviation or Standard Error of Measurement? 323
- Why All the Fuss About Error? 324
  - Error Within Test-Takers 324
  - Error Within the Test 324
  - Error in Test Administration 325
  - Error in Scoring 325
- Sources of Error Influencing Various Reliability Coefficients 325
  - Test–Retest 325
  - Alternate Forms 326
  - Internal Consistency 326
- Band Interpretation 328
  - Steps: Band Interpretation 329
- A Final Word 332
- Summary 332
- For Discussion and Practice 333

### CHAPTER 19 STANDARDIZED TESTS 335

- What Is a Standardized Test? 337
- Do Test Stimuli, Administration, and Scoring Have to Be Standardized? 338
- Standardized Testing: Effects of Accommodations and Alternative Assessments 339
- Uses of Standardized Achievement Tests 340
- Will Performance and Portfolio Assessment Make Standardized Tests Obsolete? 341
- Administering Standardized Tests 341
- Types of Scores Offered for Standardized Achievement Tests 343
  - Grade Equivalents 343
  - Age Equivalents 344
  - Percentile Ranks 345
  - Standard Scores 346
- The Norms Table 347
  - Interpreting a Norms Table 349
- Interpreting Standardized Tests: Test and Student Factors 350
  - Test-Related Factors 351
  - Student-Related Factors 354
- Aptitude–Achievement Discrepancies 358
- Interpreting Standardized Tests: Parent–Teacher Conferences and Educational Decision Making 361
  - An Example: Pressure to Change an Educational Placement 361
- Interpreting Standardized Tests: Score Reports from Publishers 366
  - Report-Based Interpretive Scenarios 371
  - Authors’ Responses 372
- Summary 373
- For Discussion and Practice 375

### CHAPTER 20 TYPES OF STANDARDIZED TESTS 377

- Standardized Achievement Tests 378
  - Achievement Test Batteries, or Survey Batteries 379
  - Single-Subject Achievement Tests 380
  - Diagnostic Achievement Tests 380
- Standardized Academic Aptitude Tests 381
  - The History of Academic Aptitude Testing 381
  - Stability of IQ Scores 382
  - What Do IQ Tests Predict? 383
  - Individually Administered Academic Aptitude Tests 384
  - Group-Administered Academic Aptitude Tests 384