

SIXTH EDITION

**EDUCATIONAL  
TESTING AND  
MEASUREMENT**

**CLASSROOM**

**APPLICATION**

**AND PRACTICE**

Suranaree University of Technology



31051000608451

**Tom Kubiszyn  
Gary Borich**

# CONTENTS

PREFACE

xv

## CHAPTER 1

### **An Introduction to Contemporary Educational Testing and Measurement** 1

Tests Are Only Tools 1

Tests Are Not Infallible 2

Testing: Part of Assessment 2

Testing and Assessment Skills: Vital to Teachers 4

Recent History in Educational Measurement 5

Current Trends in Educational Measurement 6

1997 Amendments to the Individuals With Disabilities Education Act (IDEA-97) 6

Performance and Portfolio Assessment 7

Education Reform and the Global Economy 7

“High-Stakes” Testing 8

Competency Testing for Teachers 12

Increased Interest from Professional Groups 12

Effects on the Classroom Teacher 13

About the Text 15

What If You’re “No Good in Math?” 16

Summary 16

For Discussion 17

## CHAPTER 2

### **The Purpose of Testing** 19

Testing, Accountability, and the Classroom Teacher 20

Types of Educational Decisions 22

“Pinching” in the Classroom 26

What to Measure 27

How to Measure 28

Written Tests 28

Summary 29

For Discussion 31

## CHAPTER 3

### **Norm-Referenced and Criterion-Referenced Tests** 33

Defining Norm-Referenced and Criterion-Referenced Tests 33

Comparing Norm-Referenced and Criterion-Referenced Tests	37
Differences in the Construction of Norm-Referenced and Criterion-Referenced Tests	38
NRTs, CRTs, and Language, Cultural, and Social Sensitivity	40
Summary	41
For Discussion	42

## CHAPTER 4

<b>Instructional Goals and Objectives</b>	<b>45</b>
A Three-Stage Model of Classroom Measurement	46
Why Objectives? Why Not Just Write Test Items?	48
Where Do Goals Come From?	50
Are There Different Kinds of Goals and Objectives?	51
How Can Instructional Objectives Make a Teacher's Job Easier?	53
Summary	56
For Discussion	56

## CHAPTER 5

<b>Measuring Learning Outcomes</b>	<b>59</b>
Writing Instructional Objectives	59
Identifying Learning Outcomes	59
Identifying Observable and Directly Measurable Learning Outcomes	61
Stating Conditions	62
Stating Criterion Levels	62
Keeping It Simple and Straightforward	63
Matching Test Items to Instructional Objectives	64
Taxonomy of Educational Objectives	67
Cognitive Domain	67
Affective Domain	70
The Psychomotor Domain	72
The Test Blueprint	73
Content Outline	76
Categories	76
Number of Items	76
Functions	76
Summary	78
For Practice	79

## CHAPTER 6

<b>Writing Objective Test Items</b>	<b>81</b>
Which Format?	81
True-False Items	83
Suggestions for Writing True-False Items	86
Matching Items	86
Suggestions for Writing Matching Items	89

Multiple-Choice Items	90
Higher Level Multiple-Choice Questions	96
Suggestions for Writing Multiple-Choice Items	100
Completion Items	100
Suggestions for Writing Completion or Supply Items	103
Gender and Racial Bias in Test Items	104
Guidelines for Writing Test Items	105
Summary	108
For Practice	109

## CHAPTER 7

### Writing Essay Test Items 111

What Is an Essay Item?	112
Extended Response Items	113
Restricted Response Items	114
Essay Items and Instructional Objectives	114
Suggestions for Writing and Using Essay Items	115
Why Use Essay Items?	117
Advantages of the Essay Item	117
Disadvantages of the Essay Item	118
<i>When Should Essay Questions Be Used?</i>	118
Scoring Essays	118
Why Are Essay Scores Unreliable?	119
How To Improve Scoring Reliability (and Save Time!)	119
Scoring Criteria for High-Level Essay Items	120
Summary	126
For Practice	127

## CHAPTER 8

### Administering, Analyzing, and Improving the Test 129

Assembling the Test	129
Packaging the Test	130
Reproducing the Test	131
Administering the Test	133
Scoring the Test	135
Analyzing the Test	135
Quantitative Item Analysis	136
Qualitative Item Analysis	142
Item Analysis Modifications for the Criterion-Referenced Test	143
Debriefing	149
Debriefing Guidelines	150
The Process of Evaluating Classroom Achievement	151
Summary	152
For Practice	155

**CHAPTER 9****Performance-Based Assessment****157**

- Performance Tests: Direct Measures of Competence 157
- Performance Tests Can Assess Processes and Products 158
- Performance Tests Can Be Embedded in Lessons 159
- Performance Tests Can Assess Affective and Social Skills 160
- Developing Performance Tests for Your Learners 162
  - Step 1: Deciding What To Test 162
  - Step 2: Designing the Assessment Context 168
  - Step 3: Specifying the Scoring Rubrics 168
  - Step 4: Specifying Testing Constraints 175
- A Final Word 177
- Summary 177
  - For Discussion and Practice 178

**CHAPTER 10****Portfolio Assessment****181**

- Rationale for the Portfolio 182
  - Ensuring Validity of the Portfolio 182
- Summary 199
  - For Practice 200

**CHAPTER 11****Marks and Marking Systems****201**

- What Is the Purpose of a Mark? 201
  - Why Be Concerned About Marking? 201
  - What Should a Mark Reflect? 202
- Marking Systems 204
  - Types of Comparisons 204
  - Which System Should You Choose? 208
  - Types of Symbols 208
- Combining and Weighting the Components of a Mark 210
  - Who Is the Better Teacher? 211
  - Combining Grades from Quizzes, Tests, Papers, Homework, Etc., into a Single Mark 212
- Equating Before Weighting in the Busy Classroom: Combining Performance and Traditional Assessment Scores into a Single Mark 215
  - Front-end Equating 216
  - Back-end Equating 216
- Summary 219
  - For Practice 220

**CHAPTER 12****Summarizing Data and Measures of Central Tendency 223**

What Are Statistics? 223

Why Statistics? 224

Tabulating Frequency Data 225

The List 226

The Simple Frequency Distribution 226

The Grouped Frequency Distribution 227

Steps in Constructing a Grouped Frequency Distribution 227

Graphing Data 232

The Bar Graph, or Histogram 232

The Frequency Polygon 233

The Smooth Curve 235

Measures of Central Tendency 240

The Mean 240

The Median 242

The Mode 246

The Measures of Central Tendency in Various Distributions 248

Summary 250

For Practice 252

**CHAPTER 13****Variability, the Normal Distribution, and Converted Scores 255**

The Range 255

The Semi-Interquartile Range (SIQR) 256

The Standard Deviation 258

The Deviation Score Method for Computing the Standard Deviation 261

The Raw Score Method for Computing the Standard Deviation 262

The Normal Distribution 265

Properties of the Normal Distribution 266

Converted Scores 270

Z-Scores 273

T-Scores 277

Summary 279

For Practice 280

**CHAPTER 14****Correlation 283**

The Correlation Coefficient 284

Strength of a Correlation 285

Direction of a Correlation 285

Scatterplots 286

Where Does  $r$  Come From? 288



Causality	290
Other Interpretive Cautions	291
Summary	294
For Practice	295

## **CHAPTER 15**

### **Validity**

297

Why Evaluate Tests?	297
Types of Validity	297
Content Validity	298
Criterion-Related Validity	298
Construct Validity	300
What Have We Been Saying? A Review	300
Interpreting Validity Coefficients	302
Content Validity	302
Concurrent and Predictive Validity	302
Summary	307
For Practice	308

## **CHAPTER 16**

### **Reliability**

311

Methods of Estimating Reliability	311
Test-Retest	311
Alternate Forms	312
Internal Consistency	312
Interpreting Reliability Coefficients	316
Summary	320
For Practice	321

## **CHAPTER 17**

### **Accuracy and Error**

323

Error—What Is It?	323
The Standard Error of Measurement	325
Using the Standard Error of Measurement	326
More Applications	330
Standard Deviation or Standard Error of Measurement?	332
Why All the Fuss About Error?	333
Error Within Test Takers	333
Error Within the Test	333
Error in Test Administration	334
Error in Scoring	334
Sources of Error Influencing Various Reliability Coefficients	335
Test-Retest	335
Alternate Forms	336

Internal Consistency	336
Band Interpretation	337
Steps: Band Interpretation	339
A Final Word	342
Summary	343
For Practice	344

## CHAPTER 18

<b>Standardized Tests</b>	<b>347</b>
What Is a Standardized Test?	347
Do Test Stimuli, Administration, and Scoring Have To Be Standardized?	349
Standardized Testing: Effects of Accommodations and Alternative Assessments	349
Uses of Standardized Achievement Tests	350
Will Performance and Portfolio Assessment Make Standardized Tests Obsolete?	351
Administering Standardized Tests	352
Types of Scores Offered for Standardized Achievement Tests	354
Grade Equivalents	354
Age Equivalents	355
Percentile Ranks	356
Standard Scores	357
Interpreting Standardized Tests: Test and Student Factors	359
Test-Related Factors	359
Student-Related Factors	366
Aptitude-Achievement Discrepancies	370
Interpreting Standardized Tests: Parent-Teacher Conferences and Educational Decision Making	374
An Example: Pressure To Change an Educational Placement	374
A Second Example: Pressure from the Opposite Direction	379
Interpreting Standardized Tests: Score Reports from Publishers	383
The Press-On Label	386
A Criterion-Referenced Skills Analysis or Mastery Report	387
An Individual Performance Profile	388
Other Publisher Reports and Services	389
Summary	390
For Practice	392

## CHAPTER 19

<b>Types of Standardized Tests</b>	<b>395</b>
Standardized Achievement Tests	395
Achievement Test Batteries, or Survey Batteries	396
Single-Subject Achievement Tests	398
Diagnostic Achievement Tests	398
Standardized Academic Aptitude Tests	400
The History of Academic Aptitude Testing	400
Stability of IQ Scores	401



What Do IQ Tests Predict?	401
Individually Administered Academic Aptitude Tests	403
Group Administered Academic Aptitude Tests	404
Standardized Personality Assessment Instruments	404
What Is Personality?	405
Objective Personality Tests	406
Projective Personality Tests	407
Planning a School- or District-wide Testing Program	407
Step 1: Ask the Right Question	408
Step 2: Decide on the Purposes for Your Testing Program	408
Step 3: Consider the Program's Integration	408
Step 4: Evaluate and Select the Specific Tests	411
Step 5: Develop a Specific and a District-Wide Testing Schedule	412
Summary	413
For Discussion	415

## CHAPTER 20

<b>Testing and Assessing the Special Learner in the Regular Classroom</b>	<b>417</b>
A Brief History of Special Education	420
P.L. 94-142 and the Individuals With Disabilities Education Act (IDEA)	420
Section 504 of the Rehabilitation Act	421
Special Education Service Delivery: An Evolution	422
Evolution: Service Delivery Setting	422
Evolution: Determining Eligibility for Services	423
Evolution: Disability Categories to Developmental Delays	425
IDEA-97 and the Classroom Teacher	427
Testing or Assessment?	427
Child Identification	428
Individual Assessment	430
<i>Individual Educational Plan (IEP) Development</i>	433
Individualized Instruction	436
Reviewing the IEP	436
Manifestation Determinations	437
At the Other End of the Curve: The Gifted Child	438
Defining "Gifted"	438
Assessment and Identification	439
Current Trends in Teaching and Assessing the Gifted and Talented	443
Summary	444
For Discussion	445

## CHAPTER 21

<b>Assessing Special Learners in Regular Education Classrooms</b>	<b>447</b>
IDEA-97: Issues and Questions	448
Assistance for Teachers	448

Can Teachers Assess Special Learners?	449
Should Teachers Assess Special Learners?	449
Assessing Academic Performance and Progress	450
Teacher-Made Tests and Assessments	450
Standardized Tests and Assessments	451
Limitations of Accommodations and Alternative Assessments	451
Assessing Behavioral and Attitudinal Factors	452
Assessment, Not Diagnosis	452
Classroom Diversity, Behavior, and Attitudes	453
Behavior Plans Requirements under IDEA-97	454
Teacher-Made Behavior and Attitude Assessments	454
Distinguishing Behavior from Attitude	454
Assessing Behavior	455
Assessing Attitudes	462
Monitoring Special Learners on Medication	471
Medication Use Is Increasing	471
Side Effects Are Always Present	471
The Teacher's Role in Evaluating Medication and Psychosocial Interventions	471
Commonly Used Scales and Checklists	472
Summary	477
For Discussion	478

## CHAPTER 22

### In the Classroom: A Summary Dialogue

481

Criterion-Referenced Versus Norm-Referenced Tests	485
New Responsibilities for Teachers under IDEA-97	486
Instructional Objectives	486
The Test Blueprint	487
Essay Items and the Essay Scoring Guides	487
Reliability, Validity, and Test Statistics	488
Grades and Marks	489
Some Final Thoughts	490

APPENDIX A: MATH SKILLS REVIEW	493
APPENDIX B: PEARSON PRODUCT-MOMENT CORRELATION	500
APPENDIX C: STATISTICS AND MEASUREMENT TEXTS	502
APPENDIX D: ANSWERS FOR PRACTICE QUESTIONS	503
SUGGESTED READINGS	509
REFERENCES	517
CREDITS	521
INDEX	523