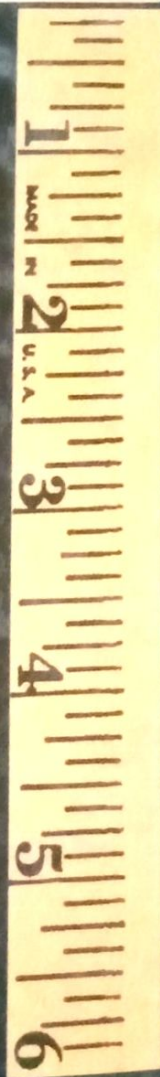
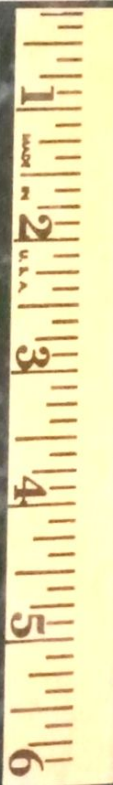


Modern Educational Measurement

W. James Popham

5564



DECIDING ON THE PURPOSE OF AN EDUCATIONAL TEST

1

Perspectives on Purposes- Past and Present 3

- Prominent Applications of Educational Tests, 6*
- The Purposes of Testing, 9*
- Prior Purposes of Testing: A Capsule History of Educational Measurement, 10*
- Related Issues, 18*
- Practice Exercises, 20*
- Answers to Practice Exercises, 21*
- Discussion Questions, 21*
- Suggestions for Additional Reading, 22*
- And For Those who Tire of Reading, 23*

Contents

2

Norm-Referenced and Criterion-Referenced Measurement 24

- The Fundamental Distinction, 25*
- Targets for Educational Tests, 33*
- Matching the Measurement Strategy with the Mission, 34*
- Practice Exercises, 38*
- Answers to Practice Exercises, 39*
- Discussion Questions, 40*
- Suggestions for Additional Reading, 41*
- And For Those Who Tire of Reading, 41*



EVALUATING EDUCATIONAL TESTS

3

What to Look for in an Educational Test 45

- Evaluative Factor One: Description of Measured Behavior, 47*
- Evaluative Factor Two: Items per Measured Behavior, 50*
- Evaluative Factor Three: Scope of Measurement, 56*
- Evaluative Factor Four: Reliability, 58*
- Evaluative Factor Five: Validity, 59*
- Evaluative Factor Six: Comparative Data, 60*
- Other Considerations, 61*
- Organizing the Test Evaluation, 62*
- Practice Exercises, 63*
- Answers to Practice Exercises, 64*
- Discussion Questions, 64*
- Suggestions for Additional Reading, 65*

4

Statistical Concepts Needed to Evaluate Tests 66

- Frequency Distributions, 68*
- Graphic Displays, 71*
- Indicators of Central Tendency, 75*
- Indicators of Variability, 79*
- Indicators of Relationship, 85*
- Practice Exercises, 95*
- Answers to Practice Exercises, 96*
- Discussion Questions, 97*
- Suggestions for Additional Reading, 97*

5

Establishing the Validity of Educational Tests 98

- Content Validity, 101*
- Criterion-Related Validities, 108*
- Construct Validity, 113*
- Practice Exercises, 123*
- Answers to Practice Exercises, 124*
- Discussion Questions, 124*
- Suggestions for Additional Reading, 124*

6

Reliability 126

- Stability, 128*
- Equivalence, 132*
- Equivalence and Stability, 134*
- A Simple-Minded Explanation of Rasch Test Equating, 135*
- Internal Consistency, 141*

The Standard Error of Measurement, 146
Illustrations of Published Reliability Data, 148
Practice Exercises, 152
Answers to Practice Exercises, 154
Discussion Questions, 154
Suggestions for Additional Reading, 154

7

Comparative Data 156

The Normal Curve, 157
Percentiles, 160
Standard Scores, 162
Norms, 172
Practice Exercises, 178
Answers to Practice Exercises, 179
Discussion Questions, 179
Suggestions for Additional Reading, 180

8

Test Bias 181

Culture-Fair Tests, 184
Evaluating for Test Bias, 190
Practice Exercises, 197
Answers to Practice Exercises, 197
Discussion Questions, 198
Suggestions for Additional Reading, 198



CREATING EDUCATIONAL TESTS

9

Specifying What a Test Should Measure 203

A Range of Assessment Options, 205
Test Specifications for Norm-Referenced Tests, 206
Test Specifications for Criterion-Referenced Tests, 211
Two Functions of Test Specifications, 230
Practice Exercises, 232
Answers to Practice Exercises, 233
Discussion Questions, 233
Suggestions for Additional Reading, 233
And For Those who Tire of Reading, 234

10 Selected-Response Tests 235

- Binary-Choice Items, 241*
- Matching Items, 247*
- Multiple-Choice Items, 251*
- Practice Exercises, 262*
- Answers to Practice Exercises, 263*
- Discussion Questions, 264*
- Suggestions for Additional Reading, 264*

11 Constructed-Response Tests 266

- Relative Merits of Selected and Constructed Response Items, 268*
- Short-Answer Items, 271*
- The Essay Item, 274*
- Practice Exercises, 283*
- Answers to Practice Questions, 283*
- Discussion Questions, 284*
- Suggestions for Additional Reading, 284*

12 Improving Test Items 286

- Judgmental Improvement of Test Items, 289*
- Empirical Improvement of Test Items, 293*
- Practice Exercises, 305*
- Answers to Practice Exercises, 306*
- Discussion Questions, 307*
- Suggestions for Additional Reading, 308*

13 Observations and Ratings 309

- Systematic Observations, 311*
- Ratings, 315*
- Practice Exercises, 325*
- Answers to Practice Exercises, 326*
- Discussion Questions, 326*
- Suggestions for Additional Reading, 326*
- And For Those who Tire of Reading, 327*

14**Creating Affective Measures 328**

- Isolating the Attribute to be Measured, 333*
- Assessment Options, 335*
- Individual and Group Affective Assessment, 345*
- Practice Exercises, 349*
- Discussion Questions, 349*
- Suggestions for Additional Reading, 349*
- And For Those who Tire of Reading, 350*

IV**USING EDUCATIONAL TESTS****15****Administering Educational Tests 353**

- Preparing Effective Test Directions, 355*
- Administering the Test, 357*
- Scoring the Tests, 361*
- Correcting for Guessing, 363*
- Additional Considerations, 365*
- Recording Test Results, 365*
- Test Security, 367*
- Practice Exercises, 369*
- Discussion Questions, 369*
- Suggestions for Additional Reading, 370*

16**Setting Performance Standards 371**

- General Considerations, 375*
- Major Factors in Setting Standards, 377*
- Alternative Standard-Setting Procedures, 384*
- Appraising the Alternatives, 392*
- Separate or Aggregate Standards, 394*
- Practice Exercises, 396*
- Answers to Practice Exercises, 398*
- Discussion Questions, 398*
- Suggestions for Additional Reading, 398*

17

Using the Results of Educational Tests 400

Grading, 400

The Design and Improvement of Instruction, 412

Practice Exercises, 422

Answers to Practice Exercises, 425

Discussion Questions, 426

Suggestions for Additional Reading, 426

Index 429