

Contents

1 The Need for Better Classroom Tests, 1

The importance of classroom tests, 3. Can tests provide valid measures of achievement? 5. The functions of classroom tests, 6. Who should prepare classroom tests? 8. Some requirements for effective testing, 12. Common mistakes of teachers in testing, 13. Sources of additional information about problems in the measurement of educational achievement, 16. Summary, 20.

2 What Should Achievement Tests Measure? 22

The dual problems of test construction, 22. The measurability of educational outcomes, 25. The problem of quantification, 28. The uses and limitations of educational objectives in test development, 30. Behavioral goals of education, 35. The cognitive outcomes of education, 38. The problem of rote learning, 45. The problem of meaningless verbalization, 47. The problem of forgetting, 47. Noncognitive outcomes of education, 48. Development of mental abilities as an educational goal, 50. Summary, 53.

3 How to Plan a Classroom Test, 57

Decide when to test, 57. Decide what kind of questions to use, 58. Decide how many questions to include in the test,

3 How to Plan a Classroom Test (Cont.)

65. Decide what emphasis to give to various aspects of achievement, 67. List topics for test items which provide a representative sample of the items of knowledge or of the abilities considered appropriate for inclusion in the test, 72. Decide what level and distribution of difficulties are appropriate for the questions included in the test, 74. Decide whether or not to apply a "correction for guessing," 75. Decide what means and what format to use in presenting the test to the students, 77. Decide whether, and if so how, to determine the passing score on the test, 80. Determine to test important outcomes of instruction, 81. Summary, 82.

4 The Characteristics and Uses of Essay Tests, 84

Some differences between essay tests and objective tests, 84. Some similarities of essay and objective tests, 102. When to use essay or objective tests, 109. Suggestions for preparing essay tests, 110. Suggestions for grading essay tests, 116. Summary, 121.

5 How to Use True-False Tests, 124

Criticisms of true-false tests items, 124. Advantages of true-false tests items, 126. Equivalence of true-false and multiple-choice item forms, 127. Confidence weighting of responses to true-false test items, 130. Confidence weighting and personality, 133. Empirical studies of confidence weighting, 134. The problem of ambiguity, 135. The problem of significance, 137. Suggestions for writing true-false test items, 137. Illustrations of good and poor true-false items, 138. Techniques for developing true-false test items, 143. Desirable balance between true and false statements, 146. Summary, 146.

6 How to Write Multiple-Choice Test Items, 149

The use of multiple-choice items in classroom testing, 149. Suggestions for preparing good multiple-choice test items, 151. Illustrations of desirable and undesirable characteristics of multiple-choice test items, 171. Illustrative reading interpretation test items, 189. Summary, 198.

7 How to Administer and Score an Achievement Test, 201

Test presentation, 202. Preparing the students, 203. Developing skill in test taking, 203. The problem of test-wiseness, 205. The problem of test anxiety, 206. Administering the test, 207. The problem of cheating, 208. Honoring systems, 211. Loss of test security, 212. Procedures for test scoring, 213. Scoring machines, 216. Correction for guessing, 223. Differential scoring weights, 233. Summary, 236.

8 Describing Test Scores Statistically, 240

Frequency Distributions, 240. The median and the mean, 242. The variance and the standard deviation, 244. The normal distribution, 248. Percentile ranks, 251. Interpretation of percentile ranks, 255. Stanine standard scores, 259. Estimating standard deviations from percentiles and stanines, 363. Coefficients of correlations, 265. Interpreting coefficients of correlation, 272. Other references, 276. Summary, 277.

9 How to Judge the Quality of a Classroom Test, 281

Relevance, 284. Balance, 287. Efficiency, 295. Objectivity, 296. Specificity, 297. Difficulty, 300. Discrimination, 300. Reliability, 302. Fairness, 304. Speededness, 305. Concluding statement, 306. Summary, 306.

10 How to Estimate, Interpret, and Improve Test Reliability, 308

The importance of reliability, 308. The concept of test reliability, 310. An operational definition of test reliability, 311. Methods of obtaining independent scores on equivalent tests, 312. Derivation of a general Kuder-Richardson formula, 320. Derivation of Kuder-Richardson Formula 20, 326. Derivation of the Spearman-Brown formula, 327. Computation of Kuder-Richardson reliability coefficients, 328. The interpretation of reliability coefficients, 330. Improving test reliability, 336. Summary, 343.

11 How to Improve Test Quality Through Item Analysis, 346

The value of item analysis data, 346. The process of item analysis, 347. An example of item analysis, 347. Selection of the criterion groups, 348. Special criterion groups, 350.

11 How to Improve Test Quality Through Item Analysis (Cont.)

Counting the responses, 351. The index of discrimination, 352. The criterion for item analysis, 356. Must all items discriminate? 358. The index of item difficulty, 359. Distribution of item difficulty values, 360. Item selection, 364. Item revision, 367. Summary, 372.

12 The Validity of Classroom Tests, 376

The concept of validity, 376. Types of validity, 379. Derived validity, 382. Operational definitions, 383. The role of judgment, 385. Tasks and traits, 386. Validity and reliability, 388. Validity and relevance, 390. Characteristics of validity, 392. Summary, 392.

13 Marks and Marking Systems, 396

The problem of marking, 396. The need for marks, 399. Shortcomings of marks, 401. Institutional marking systems, 404. Absolute versus relative marking, 406. Measurement or evaluation, 414. Achievement or attitude and effort, 416. Status or growth, 417. Single or multiple marks, 420. Few marks or many? 421. Letters or numbers, 424. Quality control in a marking system, 424. A method of marking, 426. The basis for marks, 435. Illustration of mark determination, 437. Summary, 439.

Appendix Glossary of Terms Used in Educational Measurement, 444**Index, 471**