

Contents

List of Figures	vii
List of Tables	xi
Foreword by Logan Wilson	xiii
Editor's Preface	xv
1. Educational Measurement for the Seventies	3
by Robert L. Thorndike	

PART ONE—Test Design, Construction, Administration, and Processing

2. Defining and Assessing Educational Objectives	17
by David R. Krathwohl and David A. Payne	
3. Planning the Objective Test	46
by Sherman N. Tinkelman	
4. Writing the Test Item	81
by Alexander G. Wesman	
5. Gathering, Analyzing, and Using Data on Test Items	130
by Sten Henryssen	
6. Reproducing the Test	160
by Robert L. Thorndike	
7. Test Administration	188
by William V. Clemans	
8. Automation of Test Scoring, Reporting, and Analysis	202
by Frank B. Baker	

PART TWO—Special Types of Tests

9. Performance and Product Evaluation	237
by Robert Fitzpatrick and Edward J. Morrison	
10. Essay Examinations	271
by William E. Coffman	
11. Prediction Instruments for Educational Outcomes	303
by Paul A. Schwarz	

PART THREE—Measurement Theory

12. The Nature of Measurement	335
by Lyle V. Jones	
13. Reliability	356
by Julian C. Stanley	
14. Test Validation	443
by Lee J. Cronbach	

- 15. Scales, Norms, and Equivalent Scores 508
by William H. Angoff
- 16. Techniques for Considering Multiple Measurements 601
by William W. Cooley

PART FOUR—Application of Tests to Educational Problems

- 17. Measurement in Learning and Instruction 625
by Robert Glaser and Anthony J. Nitko
- 18. Use of Measurement in Student Planning and Guidance 671
by Junius A. Davis
- 19. Use of Measurement in Selection and Placement 680
by John R. Hills
- 20. The Evaluation of Educational Programs 733
by Alexander W. Astin and Robert J. Panos
- Index 753

List of Figures

3.1	A Classification of Educational Objectives	51
3.2	Blueprint of a College Proficiency Examination in History of American Education	53
3.3	Two-Axis Chart of Specifications for a Final Examination in Natural Science	54
5.1	Normal Distribution of Subjects on the Item Continuum	139
5.2	Normal Distribution Illustrating Relation between Item Proportion and Difficulty Level for Item 15 in the Tryout Test	139
5.3	Relation between Item Difficulty and Item Proportion, When Normality Is Assumed	140
5.4	Assumed Normal Bivariate Distribution between Item and Criterion	141
5.5	Difficulty and Discrimination Indices for 16 Items in a Tryout Version of a Vocabulary Subtest	144
5.6	Item-Test Regression for Item 15 in Tables 5.3 and 5.4	146
5.7	Item-Test Regression When the Best Subjects Tend to Choose a Wrong Answer	147
5.8	Several Hypothetical Item Characteristic Curves of the Normal Ogive Forms	147
5.9	Plot of Validity and Reliability Indices for Item Selection	155
6.1	Facsimile of Test Item Illustrating Photo-Offset Reproduction of a Photograph with Copy Set by Linotype	162
6.2	Illustration of Copy Marked for Typesetting	164
6.3	Reading Passage Printed with Lines Too Long for Easy Legibility	166
6.4	Use of Very Short Lines Breaks Up the Text Excessively	167
6.5	Space Wasted on Lines Containing Choices	167
6.6	Double-Column Page Permits More Compact Arrangement of Same Material	167
6.7	Machine-Scorable Test Permitting Student to Mark Answer Directly under His Choice in Test Booklet	170
6.8	Item Is Difficult to Read in Strung-Out Form	171
6.9	Types of Items for Which Horizontal (Strung-Out) Arrangement Is Suitable	171
6.10	Illustration of Various Ways of Arranging Multiple-Choice Items	172
6.11	Diagram Interrupts Verbal Sequence from Stem to Choices	173
6.12	Stem and Choices Kept Together	173
6.13	Compact Arrangement, Especially Useful When Economy of Space is Important	173
6.14	Item Elements Differentiated by Use of Varying Amounts of Space between Lines	173
6.15	Illustrations of Ways of Setting Off Choice Numbers	174
6.16	Answers Arranged in Order of Magnitude	174
6.17	A Better Arrangement, Avoiding Confusion	175

6.18	Illustration of Ways of Arranging True-False Items	175
6.19	Illustration of Page Arrangement with Reference Material	176
6.20	Illustration of Page Arrangement with Reference Material	177
6.21	Illustration of Page Arrangement with Reference Material and Grouped Items	178
6.22	Illustration of Arrangement of Grouped Items	180
6.23	Illustration of Use of Guidelines	182
6.24	Illustration of Different Sizes of Type	185
6.25	Illustration of Different Styles of Type Faces	185
6.26	Illustration of Text Set Solid	185
6.27	Illustration of Text Set with 2-Point Leading	185
7.1	The Grid for George Henry	195
8.1	A General-Purpose Answer Sheet for a Mark-Scan Optical Scanner	204
8.2	Optical Read-Station Using the Transmitted-Light Principle	206
8.3	Optical Read-Station Using the Reflected-Light Principle	207
8.4	A High-Capacity Test-Processing System at the Measurement Research Center	209
8.5	Flow Chart for Test-Processing System	213
8.6	An Alphanumeric Display Console	215
8.7	Student Report Produced by a Commercial Test-Scoring Service	217
8.8	Class Roster and Test Scores for a Teacher	218
8.9	Results Produced by a Test Analysis Program for a Digital Computer	219
8.10	Administrator's Summary Report as Produced by an X-Y Plotter	220-221
8.11	Test Report as Might Be Produced by Computer Programs	229
9.1	Type of Score Sheet Utilized in Leaders Reaction Test	243
9.2	Police Lieutenant In-Basket Test-Scoring Form	244
9.3	Forms Used in the ASCOT Game	246
9.4	Sample ASCOT Game Outcomes	247
9.5	Extract from the Instructions to Students in a Science Fair	249
9.6	Example of Tab Items in Television Repair	250
9.7	Summary Scoring Chart for the Seashore-Bennett Test	251
9.8	A Checklist for Evaluation of Vehicle Operators	253
9.9	Instructions for a Performance Test in Metalworking	254
9.10	Instructions for a Test of Physical Fitness	255
9.11	Task Analysis Worksheet for Taking Oral Temperature	256
9.12	A Sample Task in a Pilot Flight Test	266
12.1	Distribution of Estimates of the Gravitational Constant Using Steel Bar No. 1, August 1958	347
12.2	Distribution of Estimates of the Gravitational Constant Using Steel Bar No. 2, December 1959	348
13.1	Regression of X on T	433
13.2	Regression of T on X	434
14.1	A Well-Developed Theoretical Network	477
14.2	A Portion of the Interpretative Network for n Achievement	478
14.3	Components of Criterion Variance as a Function of Number of Observations k	497
14.4	Possible Predictor-Criterion Relations When There Are Two Treatments	500

14.5	Artificial Data Contrasting Analyses under the Discriminant-Function and Regression Models	502
15.1	Ogive for Form 4a of the STEP Mathematics Test	518
15.2	Relation between the Raw Score Scales of Five Test Forms and Scores on the Reference Scale	566
15.3	Ogives for Two Forms of the STEP Social Studies Test	572
15.4	Equivalent Raw Scores on Two Forms of the STEP Social Studies Test	573
16.1	Overlap among Top 10 Percent on Two Traits Correlated .80 ($N=1000$)	602
16.2	Johnny's Profile on Five Different Traits	603
16.3	Comparison of Johnny's Profile with That for Chemists	603
16.4	Comparison of Profiles in a Two-Dimensional Trait Space	604
16.5	Distributions for Groups A and B on Trait X	608
16.6	Distributions for Groups A and B	609
16.7	Bivariate Normal Distributions for Populations A and B	609
16.8	Ellipses Representing Three Centours	610
16.9	Regions for Two Populations in One-Dimensional Space	610
16.10	Rotation of Axes through Angle θ	612
16.11	Projection from Test Space (X_1, X_2) to Discriminant Space (F_1)	615
16.12	Chemists and Chemistry Teachers in the One-Dimensional Discriminant Space	615
16.13	The 36 Centroids in the Ability Space	617
17.1	Curriculum Hierarchy on the Addition of Integers	634
17.2	Curriculum Hierarchy for Counting a Collection of Movable Objects	635
17.3	Curriculum Hierarchy for Placing an Object in a Two-Dimensional Matrix	636
17.4	Two Possible Hierarchies of Sequence of Instruction	638
17.5	Illustration of Alternative Instructional Sequences	640
17.6	Hierarchies of Objectives for an Arithmetic Unit in Addition and Subtraction	642
17.7	Flow Diagram for an Instructional System	648
17.8	Instructional Process Flowchart for the IPI Procedure	649
17.9	Examples of Item Forms from the Subtraction Universe	657
17.10	Illustration of Hively's Task Format and Task Generation Rules	658
17.11	An Example of a Verbal Replacement Set for a Variable Element in an Item Form	659
17.12	Basic Operational Elements in Development and Evaluation of a System for IPI	665
19.1	Regressions of English Literature Grades on Placement Test Scores Where Placement Is Effective	714
19.2	Regressions of English Literature Grades on Placement Test Scores Where Placement Is Ineffective	714
19.3	Regressions of Mathematics Criterion Scores on Numerical Operations Placement Test Standard Scores for Two Treatment Groups in Mathematics	719
19.4	Regressions of Payoff from Mathematics Criterion Scores on Numerical Operations Placement Test Scores for Three Treatments	720
20.1	Components of an Educational Program	737

List of Tables

2.1	Ways of Conceptualizing Educational Objectives	32
3.1	Topics for Mechanical Comprehension Test with Their Assigned Weights	57
5.1	Distribution of Criterion Scores for Students Choosing Different Answers on Pretryout Item 16	136
5.2	Distribution of Criterion Scores for Students Choosing Different Answers on Tryout Item 15	138
5.3	Distribution of Criterion Scores for the Students Choosing Right and Wrong Answers on Item 15	140
5.4	Number of Examinees from the Upper 27 Percent and Lower 27 Percent Groups Choosing Different Answers on Item 15	145
5.5	Percentage of Examinees from the Upper 27 Percent and Lower 27 Percent Groups Choosing Different Answers on Item 15	145
5.6	Proportion of Right Answers to Item 15 for Each Criterion Score	146
10.1	Intercorrelations of Objective Score, Operational Essay Scores, and Experimental Essay Scores	281
10.2	Frequency Distributions of Scores on Twelve Questions	290
10.3	Frequency Distributions of Essay Subscore for One Question of the 1966 Advanced Placement Examination in English	294
10.4	Estimates of Components of Variance and Reliability Coefficients for Scores Based on the Sum of Two Ratings for Questions in Advanced Placement Examinations, 1967	295
10.5	Computer Simulation of Human Judgments for Five Essay Traits	296
13.1	Possible Sources of Variance of Scores on a Particular Test	364
13.2	Reliability of a Difference Score	387
13.3	Schema for a Persons-by-Items Matrix of Item Scores	419
13.4	Test Consisting of $mL=I$ Stratified-Random Items Administered to n Males and n Females by the t^{th} Tester; X_{split} Notation	431
14.1	Summary of Types of Validation	446
14.2	Hypothetical Correlations of Clarity and Warmth Scores Obtained in Three Different Ways	468
14.3	Number of Categories for Which a Hypothesized Difference in Correlations Appears in Hoepfner-Guilford Data for Divergent Tests	472
14.4	Subjects, Classified on Proj and TAQ, Compared to Median for All Subjects on Experimental Measures	479
14.5	A Specimen Expectancy Table	493
15.1	Normalization of Scores on Form 4a of the STEP Mathematics Test	517
15.2	Efficiency of Two-Stage Sampling Procedures for Sixth-Grade Data and for Tenth-Grade Data	555

15.3	Scaled Score Values for Five Forms of a 60-Item Test	570
15.4	Distributions of Raw Scores on Two Forms of the STEP Social Studies Test	574
15.5	Equipercntile Points on Two Forms of the STEP Social Studies Test	575
15.6	Equivalent Raw Scores on Two Forms of the STEP Social Studies Test	575
16.1	Roster of Scores for 10 Chemists on Five Variables	604
16.2	Some Illustrative Basic Matrix Operations	605
16.3	Variance-Covariance Matrix	606
16.4	Correlation Matrix	606
16.5	Scores, Chi Squares, and Centours for the 10 Chemists	607
16.6	Correlation Matrix for Factor Analysis Example	613
16.7	Results of Factor Analysis	614
16.8	Roster of Scores for 10 Chemistry Teachers on Five Variables	615
16.9	Weights and Correlations for the Discriminant Function	616
16.10	Six-Category Classification Scheme	616
16.11	Career Group Self-Predictions (Grade 9 Males)	617
16.12	Multiple Correlation Results for Chemists	618
16.13	Canonical Correlation Results	620
17.1	Objectives for Computer-Assisted Branched Testing for Addition-Subtraction	643
19.1	Bivariate Grouped Frequency Distribution of College Board English Achievement Test Score vs. Grades in Regular Freshman English	711