

Longmans' Education Series

THE USE AND INTERPRETATION OF EDUCATIONAL TESTS

A HANDBOOK FOR THE CLASSROOM TEACHER

BY

HARRY A. GREENE, PH.D.

ASSOCIATE PROFESSOR OF EDUCATION, AND
DIRECTOR OF BUREAU OF EDUCATIONAL RESEARCH AND SERVICE
UNIVERSITY OF IOWA

AND

ALBERT N. JORGENSEN, PH.D.

DIVISION OF EXPERIMENTAL RESEARCH AND
DIRECTOR OF EDUCATIONAL ADMINISTRATION
MICHIGAN STATE NORMAL COLLEGE

LONGMANS, GREEN AND CO.
NEW YORK · LONDON · TORONTO

1929

**THE USE AND
INTERPRETATION OF
EDUCATIONAL TESTS**

PREFACE

THIS book is designed to place in the hands of the classroom teacher and the student of education an introductory and elementary discussion of the essential principles of measurement in education. The material and the plan of organization are the results of a number of years of experience in teaching courses in educational measurements and in practical work in this field with the classroom teacher. It is not a part of its purpose to compete with or to repeat the numerous more technical discussions of test construction and interpretation now available. These are for the advanced statistician and research student. The present treatment is planned to provide a systematic handbook for any person who feels the need of a well-organized treatment of the fundamental ideas and tools of measurement in education. It is written from the practical point of view of the classroom teacher, and in language sufficiently non-technical to appeal to the beginning student in this field which to some seems highly technical. It is hoped that because of this direct method of attack, the practical point of view taken, and the concrete illustrations, it may appeal to the group for which it is intended. Only through the active enlistment of the interest of the classroom teacher and of the student who himself is to serve in that capacity will the fullest possibilities of measurement in education be realized.

The authors believe that the chief purpose served by standard tests is the improvement of classroom instruction through the more exact location, diagnosis, and remedy of individual pupil difficulties in learning. They believe, too, that the responsibility for the use of these instruments for the location,

diagnosis, and remedy of pupil difficulties must be more and more assumed by the classroom teacher. This does not mean any less responsibility on the part of the supervisor. It means a clearer recognition of it on the part of the classroom teacher, for after all the use and interpretation of educational tests is a coöperative project. It is all directed to the same end, the goal of all good supervision, the improvement of the conditions under which teachers teach and children learn.

To accomplish these purposes, much more emphasis must be placed upon the training of the prospective teacher, as well as of the more experienced teacher, in the use of measuring devices and in the development and application of remedial materials following diagnosis. This book is an attempt to place in the hands of the student, the classroom teacher, and the supervisor some basic suggestions of ways in which educational measuring instruments may be of practical use in the teaching of children and some general rules for guidance in selecting, using, and interpreting educational tests as valuable aids in accomplishing this task. If, through the reading of this brief treatment of the field as limited above and the working of the practical problems in measurement given in the *Appendix*, the teacher or student shall be prompted to a more intensive examination of the possibilities of standard tests as teaching and supervisory instruments, the purpose of this book will be fulfilled.

The indebtedness of the authors to the many experienced teachers and supervisors who have contributed indirectly as well as directly to the formulation of much of the material incorporated herein is gratefully acknowledged. Special thanks are due to Miss Mabel Snedaker, Extension Supervisor, University of Iowa, and to Miss Maud McBroom, Supervisor of the University Elementary School, as well as to the teachers of this school for the suggestion of a number of remedial devices in the field of reading and language. The authors' special thanks are also due Dr. G. M. Ruch, Professor of Education, University of California, who read much of the

manuscript and made many helpful suggestions as to its organization. We are also grateful to the authors and publishers who gave permission for the reproduction of many selected portions of their work and publications. Specific acknowledgments of these are made throughout the text.

We wish also to acknowledge our very deep indebtedness to Professor Frederic B. Knight for his careful study of this manuscript and for his many valuable suggestions as to organization and content, and to Miss Neva Steele, Miss Esther Flynn, and Mr. Paul K. Frazer for their assistance in the final preparation of the manuscript.

CONTENTS

CHAPTER	PAGE
I. INTRODUCTION.....	1
I. BACKGROUND OF THE GROWTH OF THE MEASUREMENT IDEA.	
1. Educational Measurements Not New Idea.	
2. Real Pioneers in Measurements. 3. Factors in Growth of Measurement Idea. 4. Rapid Growth of Interest in Tests. 5. Purpose of this Book.	
II. THE MEANING OF EDUCATIONAL TESTS.....	10
I. WHAT TESTS ARE.	
6. Measurement in Education. 7. Characteristics of Educational Tests. 8. What Tests Do. 9. What Tests Do Not Do.	
II. KIND OF TESTS.	
10. Tests and Scales. 11. Mental Tests. 12. Individual Mental Examinations. 13. Group Tests of Mental Ability. 14. Special Aptitude and Prognostic Tests. 15. Tests of Achievement. 16. Quality and Product Scales. 17. Survey vs. Diagnostic Tests.	
III. SUMMARY.	
III. THE USES OF EDUCATIONAL TESTS.....	47
18. Tests in Relation to Classroom Instruction.	
I. USES OF ACHIEVEMENT TESTS.	
19. For Purposes of Comparison. 20. For Measuring Efficiency of Instruction. 21. For Purposes of Class Diagnosis. 22. For Purposes of Individual Pupil Diagnosis. 23. For Pupil Gradation and Guidance.	
II. USES OF MENTAL TESTS.	
24. A Warning. 25. For Class Diagnosis. 26. For Individual Diagnosis. 27. Inter-	

CHAPTER	PAGE
dependence of Achievement and Mental Test Results.	
III. SUMMARY.	
IV. USING STANDARD TESTS IN THE CLASSROOM.....	67
28. Introduction.	
I. GIVING TESTS.	
29. The Teacher's Responsibility.	30. Scoring and Interpreting Tests.
II. WHEN TO GIVE TESTS.	
31. Periodical vs. Continuous Use of Tests.	
III. THE BASIS FOR THE TESTING PROGRAM.	
32. A Clear-cut Teaching Problem as Basis.	
33. Steps in a Testing Program.	34. Illustrations of Problems.
IV. THE CRITERIA FOR SELECTING TESTS.	
35. Validity.	36. Reliability.
37. Objectivity.	
38. Educational Significance.	39. Administrability.
40. Economy.	41. Adequacy of Test in Measuring Specific Functions.
V. GENERAL RULES FOR A TESTING PROGRAM.	
42. Preparation for Administering Tests.	43. Timing.
44. Scoring.	
VI. SUMMARY.	
V. CONSTRUCTING AND USING INFORMAL OBJECTIVE EXAMINATIONS IN THE CLASSROOM.....	89
I. CHARACTERISTICS OF INFORMAL OBJECTIVE EXAMINATIONS.	
45. Introduction.	46. The New Examination.
47. Informal vs. Standardized Tests.	
II. ADVANTAGES AND DISADVANTAGES OF THE INFORMAL OBJECTIVE EXAMINATION.	
48. Advantages of the Informal Objective Examination.	49. Disadvantages of the Objective Examination.
III. PRINCIPLES OF CONSTRUCTING AND USING OBJECTIVE EXERCISES.	
50. Need for Specific Directions.	51. Number of Exercises to Use.
52. Selection of the Subject-matter Items to be Tested.	53. Clear State-

CONTENTS

xiii

CHAPTER

PAGE

ment of Exercises. 54. Difficulty of the Exercises. 55. Arrangement of the Exercises. 56. Arrangement of Answers for Convenience in Scoring. 57. Use of Various Types of Exercises.

IV. PRACTICAL USE OF RESULTS.

58. Use of Objective Test Scores as Basis for Class Grades.

V. CLASSIFICATION AND ILLUSTRATIONS OF TYPES OF OBJECTIVE EXERCISES.

59. Classification of Objective Tests. 60. Samples of Types of Objective Test Exercises. 61. Chance Factor in Objective Tests.

VI. SUMMARY.

VI. SUMMARIZING THE RESULTS OF TESTING..... 116

62. Introduction.

I. CLASSIFICATION AND TABULATION OF TEST SCORES.

63. Need for a Method of Grouping Data. 64. Steps Involved in Preparing a Frequency Table. 65. Determining the Range of Scores. 66. Determining the Number and Size of Class-Intervals. 67. Determining the Limits of the Class-Intervals. 68. Tabulating the Scores. 69. Summary of Steps in Making a Frequency Table.

II. MEASURES OF CENTRAL TENDENCY.

70. Computing the Arithmetic Mean from Ungrouped Data. 71. Computing the Arithmetic Mean—Data grouped in Frequency Distribution. 72. Summary of Steps in Finding Arithmetic Mean. 73. Use of Arithmetic Mean. 74. The Median as a Measure of Central Tendency. 75. Computing a Counting Median or Mid-Measure. 76. Computing the Median from Frequency Distribution. 77. Use of the Median.

III. MEASURES OF VARIABILITY.

78. Need for Measures of Variability. 79. Range as an Expression of Variability. 80. "Q" or Semi-inter-quartile range as a Measure of Variability. 81. Computing the Quartile Deviation, "Q". 82. Further Practice in Find-

ing "Q." 83. Standard Deviation (Sigma) as a Measure of Variability. 84. Computing the Standard Deviation. 85. Computing the Standard Deviation from Data in Frequency Distribution. 86. Additional Practice in Finding Sigma.

IV. THE RELATIONSHIP OF TEST SCORES.

87. Need for Measures of Relationship. 88. Meaning of Correlation. 89. Tabulating Test Scores in Double Entry or Correlation Tables. 90. Steps in the Computation of the Pearson Product-Moment Coefficient of Correlation. 91. Illustration of the Computation of the Pearson Correlation Coefficient. 92. Meaning of the Correlation Coefficient. 93. Additional Experience with Correlations.

V. ASSIGNMENT OF RELATIVE AND ABSOLUTE RANKS.

94. Need for Method of Ranking. 95. Assignment of Relative Ranks. 96. Assignment of Absolute Ranks. 97. Computing Percentiles (Deciles) in a Frequency Distribution. 98. A Short Method of Computing Percentile Scores. 99. Computation of the Percentile Score Corresponding to a Given Test Score in a Frequency Table.

VI. SUMMARY.

VII. INTERPRETING THE RESULTS OF TESTING..... 176

I. TEST NORMS AND STANDARDS.

100. Norms as Distinctive Features of Standardized Tests. 101. Norms vs. Standards. 102. Uses of Test Norms and Standards. 103. Kinds of Norms. 104. Derivation of Norms. 105. Meaning of Test Norms. 106. Grade Norms. 107. Age Norms. 108. Interpretation of Grade and Age Norms.

II. TYPES OF DERIVED SCORES.

109. Function of Derived Scores. 110. Grade Norms as Derived Scores—Type 1. 111. Age Scores as Derived Scores—Type 2. 112. Percentile Ranks as Derived Scores—Type 3. 113. Quotients as Derived Scores—Type 4.

III. SUMMARY.

CONTENTS

xv

CHAPTER

PAGE

VIII. THE PLACE OF TESTS IN DIAGNOSIS AND REMEDIAL TEACHING..... 207

I. THE FUNCTION OF EXACT DIAGNOSIS.

114. Importance of the Diagnostic Interpretation of Test Results. 115. Exact Diagnosis the Basis for Remedial Work. 116. Diagnosis as the Basis for Preventive Work. 117. General Practice Exercises vs. Remedial Drill.

II. CRITERIA FOR CONSTRUCTION OF REMEDIAL MATERIAL.

118. Analysis must Precede Diagnosis. 119. Diagnosis must be Specific. 120. Each Identified Skill must be Paralleled by a Valid Drill Unit. 121. The Complete Area of the Skill Must be Covered. 122. The Underlying Skills Must be Synthesized in Remedial Units. 123. Time and Achievement Standards Should Accompany Remedial Drill Material.

III. SUMMARY.

IX. DIAGNOSTIC AND REMEDIAL TECHNIQUES; ARITHMETIC..... 231

124. Introduction. 125. Scope of Remedial Teaching in Arithmetic.

I. IDENTIFICATION OF ARITHMETICAL SKILLS.

126. Specific Skills in Arithmetic as Basis for Test and Remedial Units.

II. REMEDIAL INSTRUCTION IN ARITHMETIC.

127. Materials for Remedial Work in Arithmetic. 128. Using the Remedial Materials in Classroom.

III. SUMMARY.

X. DIAGNOSTIC AND REMEDIAL TECHNIQUES; LANGUAGE..... 251

129. Introduction. 130. A Language Consciousness.

I. IDENTIFICATION OF SPECIFIC LANGUAGE ABILITIES.

131. Analysis of Language Skills. 132. Oral Language Skills. 133. Written Language Skills.

II. DIAGNOSIS OF LANGUAGE DISABILITIES.

134. Diagnosis of Oral Language Disabilities.
 135. Diagnosis of Difficulties in Written Language. 136. Diagnostic Exercises in Written Language.

III. REMEDIAL INSTRUCTION IN LANGUAGE.

137. Language Diagnosis as a Basis for Remedial Instruction. 138. Remedial Exercises in Language; Spelling. 139. Remedial Exercises in Punctuation. 140. Remedial Exercises in Form and Appearance. 141. Remedial Exercises in Correct Language Usage. 142. Remedial Exercises in Sentence Structure. 143. Remedial Drill in Choice of Words. 144. Drill on Language Organization.

IV. SUMMARY.

XI. DIAGNOSTIC AND REMEDIAL TECHNIQUES; READING, AND OTHER SUBJECTS. 285

145. Educational and Social Importance of Reading Ability.

I. IDENTIFICATION OF MAJOR READING ABILITIES.

146. Major Objectives in Reading Instruction.
 147. Typical Defects in Reading. 148. Oral vs. Silent Reading.

II. DIAGNOSIS AND REMEDY IN ORAL READING.

149. Diagnosis of Oral Reading Difficulties.
 150. Remedial Drill for Oral Reading difficulties.

III. DIAGNOSIS AND REMEDY OF SILENT READING DIFFICULTIES.

151. Diagnosis of Work-Study Type of Reading; Elementary School Level. 152. Diagnosis of Work-Study Type of Reading; High School and College Level. 153. Remedial Drills for Work-Study Types of Reading Difficulties. 154. Other Remedial Materials in Reading.

IV. DIAGNOSIS AND REMEDY IN OTHER SUBJECTS.

155. Diagnostic and Remedial Treatment of Other School Subjects.

V. SUMMARY.

CONTENTS

xvii

CHAPTER	PAGE
XII. TESTS AND THE CLASSROOM TEACHER.....	327
I. SUMMARY.	
156. Recognition of the Need for Tests in the Classroom. 157. Meaning of Educational Tests. 158. Uses of Educational Tests. 159. Using Tests in the Classroom. 160. Summarizing and Interpreting the Results of Testing. 161. After Testing What? 162. Conclusion,	
APPENDIX.....	335
A. Distributing Agencies for Tests.....	335
B. Classified List of Tests and Publishers.....	337
C. Suggested List of Remedial and Drill Materials....	353
D. Illustrative Problems in Educational Measurements.	356
E. Individual Pupil Test Record Form.....	375
F. Glossary of Terms.....	376
INDEX.....	385

LIST OF TABLES

TABLE NO.	PAGE
1 Distribution of Mental Ability of a 9th Grade Class.	51
2 Marks Assigned by Ten Teachers to an Essay Type Examination in Civics, and a True-False Examination over the Same Material.	80
3 Theoretical Basis for Distribution of Objective Test Grades.	101
4 Suggested Point Values Corresponding to Class Grades.	102
5 Scores on a History Test Taken by an 8th Grade Class.	118
6 Arithmetic Scores of Thirty-Seven 7th Grade Pupils.	119
7 Frequency Distribution Based on Data in Table 6.	125
8 Relation of Range and Size of Class-Interval.	126
9 Distribution of Scores and Calculation of Arithmetic Mean.	133
10 Arithmetic Test Scores Arranged in Descending Order.	134
11 Arithmetic Scores in Frequency Distribution.	136
12 Distribution of Scores on a First Grade Reading Vocabulary Test.	138
13 Frequency Distribution.	143
14 Standard Deviation from Ungrouped Data.	147
15 Standard Deviation from Frequency Distribution.	148
16 Scores of a Tenth Grade Class on a Plane Geometry Test.	151
17 Pairs of Test Scores.	154
18 Illustrating the Computation of XY Products.	158
19 Correlation Table.	159
20 Relative Ranks.	167
21 Table for Assignment of Rank Positions.	168
22 Computing Percentiles (Deciles).	170

LIST OF TABLES

TABLE NO.		PAGE
23	Interpretation of Percentiles.....	171
24	Short Method of Computing Percentiles.....	173
25	Norms and Standards of Accuracy.....	178
26	Distribution of Scores for Computation of Norms..	181
27	Percentile Norms.....	182
28	Woody-McCall Arithmetic Test Results.....	186
29	Grade Increments.....	186
30	Semester Improvement, Monroe Silent Reading Tests.....	188
31	Distribution of Language Test Scores by Ages.....	189
32	Age-Equivalents of Composite Scores on Stanford Achievement Examination.....	190
33	Pupil Record from Compass Diagnostic Test Number 7.....	192
34	Grade Norms for Compass Diagnostic Test Number 7.....	192
35	Age-Equivalents of Total Scores; Compass Diag- nostic Test Number 7.....	193
36	Comparison of Units of Measurement.....	194
37	Grade Levels for Thorndike-McCall Silent Reading Scales.....	196
38	Thorndike-McCall T-Scores and Reading Ages.....	199
39	Data for Reading Ages.....	200
40	Data for Reading Quotients.....	202
41	Data for Accomplishment Quotients.....	204
42	Test Record of a Pupil.....	209
43	Distribution of Practice; Remedial Unit Number 27.	217
44	Analysis of Problem Solving.....	223
45	Distribution of Practice; Remedial Unit Number 31.	227
46	Scope of Remedial Drill Units in Whole Numbers..	243
47	Distribution of Practice in an Addition Test Pre- pared by a Teacher.....	248

LIST OF FIGURES

	PAGE
1 Graphic Record Card.....	57
2 Illustrating the Principle of Sampling.....	77
3 Mid-points and Limits of Class-Interval.....	124
4 Illustrating the Principle of Moments of Forces...	129
5 Illustrating the Calculation of the Mean.....	131
6 Illustrating Need for Measures of Dispersion.....	139
7 Comparison of Sigma and Q.....	145
8 Graphic Presentation of the Median, 25th, and 75th Percentiles.....	184
9 Comparing Achievement with Grade Norms.....	186
10 Illustration of Need for Age Norms.....	190

LIST OF PROBLEMS

APPENDIX D

I. TABULATION OF TEST SCORES

	PAGE
1 Geography Test Scores of 21 Fifth Grade Pupils.....	357
2 Reading Test Scores of 41 Fourth B Grade Pupils....	358
3 Language Test Scores of 31 Sixth Grade Pupils.....	358

II. FINDING THE ARITHMETIC MEAN

4 Computing the Arithmetic Mean from Frequency Distribution.....	361
5 Computing the Arithmetic Mean from Frequency Distribution.....	362
6 Computing the Arithmetic Mean.....	363
7 Computing the Arithmetic Mean.....	363

III. FINDING THE MEDIAN

8 Finding the Mid-Measure.....	363
9 Finding the Mid-Measure.....	363
10 Finding the Mid-Measure.....	364
11 Tabulating, Finding Median, and Arithmetic Mean..	368
12 Tabulating, Finding Median, and Arithmetic Mean..	369

IV. EXPRESSING THE VARIABILITY OF TEST SCORES

13 Finding the Semi-Inter-Quartile Range.....	370
---	-----

V. RELATIVE AND ABSOLUTE RANKS

14 Assignment of Relative Ranks.....	371
15 Computing Percentile Ranks.....	371

VI. USING TABLES OF NORMS

16 Finding Age-Equivalents from Test Scores.....	372
17 Comparing Test Scores with Grade Norms.....	373
18 Finding Quotients from Ages Given.....	374

THE USE AND INTERPRETATION OF EDUCATIONAL TESTS

CHAPTER I

INTRODUCTION

The purpose of this chapter is the discussion of the following points which seem to underlie the development of the idea of measurement in education:

- a. The idea of measurement in education is not necessarily a recent movement, but was in the minds of a number of pioneers many years prior to its general public acceptance.
- b. The real pioneers in the measurement movement.
- c. The effect of studies of the inaccuracy of teachers' marks on the measurement idea.
- d. The effect of the "school survey" idea on the development of tests in school subjects.
- e. The stimulation given to the measurement movement by the organization of research bureaus and kindred agencies.
- f. The effect of detailed analysis of skills underlying achievement in certain fields.
- g. The development of detailed diagnostic tests based upon the analysis of underlying skills.
- h. The minimal principles needed in the use and interpretation of tests.

I. BACKGROUND OF THE GROWTH OF THE MEASUREMENT IDEA

1. Educational Measurements not new idea. Standardized tests and the types of information resulting from their use in