

**MEASUREMENT AND
ADJUSTMENT SERIES**

EDITED BY LEWIS M. TERMAN

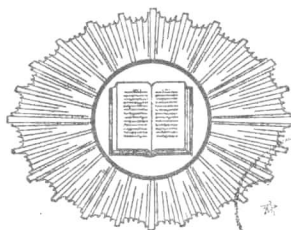
**TESTS
AND MEASUREMENTS
IN HIGH SCHOOL
INSTRUCTION**

By G. M. RUCH

Professor of Education, University of California

AND GEORGE D. STODDARD

Assistant Professor of Psychology and
Education, State University of Iowa



Yonkers-on-Hudson, New York
WORLD BOOK COMPANY
2126 Prairie Avenue, Chicago

PREFACE

THIS volume has been prepared primarily as a summary of the available tests and test methods in high school instruction. Since no separate treatment of measurement in secondary education has appeared to date, it was thought wise to include certain historical facts in Part I. In response to the increasingly critical attitude toward test selection, Chapter IV was prepared.

Part II presents the more important published measurements which are useful to the high school teacher. A great many tests have not been mentioned chiefly because of the lack of adequate data for their evaluation. Such tests are not necessarily undesirable, but no secure basis for their recommendation could be found in the literature. Part II presents numerous reliability coefficients, most of which were calculated by the authors or their students especially for this volume. Certain apologies might be offered for the scantiness of numbers of cases employed at times in the calculation of reliabilities, and for the many omissions in these data. However, it should be recognized that the burden of such determinations does not in fairness rest with the present authors, but rather with the authors of the tests themselves. If the reliabilities given are at times in serious error due to the limited numbers of cases used, or for other reasons, their presentation will at any rate stimulate their refutation and in time lead to the publication of more accurate results. Incidentally, it should be pointed out that very few makers of tests in the high school field have felt the necessity for ascertaining many pertinent statistical considerations which would be of the greatest value in the careful selection of tests.

In preparing Part II, the authors have tried to avoid being unduly swayed by such facts as great popular recogni-

tion of tests merely because of early historical origin, except to the extent that this is a legitimate criterion of merit. Where experimental findings justify the use of a less well-known test, such a test has been given as great recognition as others of more general reputation, or even greater.

Part III is to be justified upon the basis of the rapidly growing interest in informal objective examination methods. No treatment of test methods for classroom use can in the future ignore these unstandardized tests and examinations, which appear to be supplanting many of the traditional examination practices. Here again, especially in Chapter XVI, the authors were forced to take some position on a number of moot points. In the main, it is thought that the weight of the more extensive investigations indicates the fairness of the tentative recommendations advanced.

Part IV was placed last in the book, although logically it might have followed Part I. This was done for the sake of the reader who is not interested in statistical aspects of tests and test construction. The treatment of Part IV, moreover, is from the point of view of the test builder rather than that of the classroom teacher. A moment's thought will show that there is really no alternative. Part IV is to be defended upon two bases: (a) the fact that many high school teachers desire some insight into the more critical phases of test construction and selection; and (b) the fact that this book will doubtless find its way into the college classroom as a text in theoretical as well as applied aspects of measurement. There are so very few brief discussions of the general technique of test construction that it was felt that a few dozen pages could well be devoted to matters which obviously facilitate critical selection, use, and interpretation of measurements.

The authors' thanks are due a large number of graduate students who gave, scored, and tabulated test returns, and

who at times purchased test materials out of their own funds. Mr. Richard R. Foster calculated a considerable number of correlation coefficients for Part II. Grateful acknowledgment is made to Dean Paul C. Packer, of the College of Education, State University of Iowa, for funds used for the purchase and scoring of tests and for incidental clerical and stenographic assistance. Mrs. George D. Stoddard gave valuable aid in the preparation of certain sections of the book. Indirectly, the authors have drawn freely from the published work of many writers; particularly, from the statistical treatises of Dr. Truman Lee Kelley. Comment should be made upon the patience of the editor, Dr. Lewis M. Terman, and of the publishers in awaiting the completion of a task which has required the better part of four years.

G. M. RUCH

G. D. STODDARD

EDITOR'S INTRODUCTION

IN the history of tests and measurements nothing stands out as more striking or significant than the recent increase of interest in this movement among teachers in secondary schools. The measurement of both intelligence and achievement began in the grades below the high school and for several years was rarely applied above the eighth grade. It is possible that the somewhat tardy introduction of measurement methods at the higher levels may have been due in part to a more marked tendency to conservatism on the part of high school teachers, who, at least in many parts of the country, have frequently had less pedagogical training than teachers in the elementary grades. Perhaps a more potent cause lies in the apparently greater difficulty of devising satisfactory measuring instruments for use with high school students. We are finding, however, that the supposed difficulties were more apparent than real. Experience shows that it is about as easy to measure either the natural abilities or the subject-matter achievement of high school pupils as it is to apply such measurements in the elementary grades. In some respects, indeed, it is easier. With older children a standard test procedure can be more rigidly followed and the factors of attention and effort can be more readily controlled. Moreover, there is probably no greater inherent difficulty in measuring achievement in algebra, geometry, physics, ancient history, or a foreign language than there is in measuring achievement in reading, arithmetic, or spelling.

Not only are measurement methods as applicable in the high school as elsewhere; in no part of the school system is more to be gained from their use. It is in the junior and senior high school that pupil guidance, both educational and vocational, assumes such outstanding importance. The student at this age stands at the threshold of life, subject to

conflicting desires and uncertain as to his abilities and as to the direction he should take. He is facing decisions which will affect his entire life: whether to leave school and go to work, whether to complete the high school course, whether to prepare for college, whether to fit himself to enter one or another of the professions. Without objective information of the kind which is obtainable only from standard tests, the guidance of such a student can rest upon little more than guesswork. One is tempted to put it more strongly and to say that *educational guidance without educational testing is professional quackery*, as much so as in the case of the physician who refuses to employ the approved laboratory techniques in the diagnosis and treatment of certain diseases.

That the truth of the above statements is coming to be generally recognized, among high school teachers as well as among university professors of secondary education, is indicated by the multiplication of high school achievement tests and by their widespread use. For most of the high school subjects fairly satisfactory tests have already been devised, and considerable progress has been made in the testing of special aptitudes and interests at this level.

One obstacle to the use of achievement tests in the high school has been the lack of a suitable textbook. The principal or teacher who would acquaint himself with the recent advances in this field has found it necessary to search through several educational magazines and innumerable monographs. In summarizing and interpreting the scattered contributions which have appeared to date in this field, the authors of the present volume have rendered a service to secondary education which the teaching fraternity will not be slow to appreciate.

The treatment of the subject throughout is direct, intelligible, and helpful. On the one hand it avoids over-emphasis of the theoretical problems of measurement, and on the

other hand it avoids too great preoccupation with the minutiae of procedure, scoring, and computations. The extensive experience of the authors (the senior author was formerly principal of the University High School at Eugene, Oregon) has enabled them to write a book ideally adapted to its purpose — one that is informative, balanced, and helpful. Both the uses and the limitations of tests are clearly and convincingly set forth. Criteria are given for the selection of tests suitable for a particular purpose. All the important intelligence and achievement tests intended for use in the high school are described and evaluated. Finally, in order that those who make use of testing methods may have more than a superficial knowledge of the instruments they employ, four chapters are added on the principles of test construction. In the opinion of the editor, no other chapters in the book are more worthy of careful study than these last four.

It is safe to predict that *Tests and Measurements in High School Instruction* will find a wide field of usefulness both as a classroom text and for reading-circle purposes. It is a book which the progressive high school teacher cannot afford to ignore.

LEWIS M. TERMAN

CONTENTS

	PAGE
EDITOR'S INTRODUCTION	xvii
PART ONE: STATUS, USES, LIMITATIONS, AND SELECTION OF TESTS IN SECONDARY SCHOOL INSTRUCTION	
CHAPTER	
I. THE PRESENT STATUS OF MEASUREMENT IN SECOND- ARY SCHOOLS	1
<p>Historical — The first educational tests — The rise of intelligence tests — Other early educational tests — Educational testing at the present time — Limitations in the past: 1. Inertia; 2. College domination and teacher training; 3. Lack of sharp definition of content of high school subjects; 4. Qualitative <i>vs.</i> quantitative values in instruction</p>	
II. USES AND LIMITATIONS OF TESTS IN THE HIGH SCHOOL	8
The major uses of standardized measures	
I. TESTS IN THE ADMINISTRATION AND SUPERVISION OF INSTRUCTION	8
<p>Importance of tests in school supervision — Classification of supervisory uses of tests: 1. The measurement of pupil progress; 2. The diagnosis of teaching efficiency; 3. Setting up standards of performance; 4. The objectification of records of performance</p>	
II. THE DIAGNOSIS OF SPECIAL DIFFICULTIES	18
The requirements for diagnostic tests	
III. USES AND LIMITATIONS OF TESTS IN THE HIGH SCHOOL (<i>Continued</i>)	28
III. TESTS IN GRADING, PROMOTIONS, AND SECTIONING	28
<p>The general problems of classification — Tests in determining promotions — Sectioning of high school classes — The techniques of classification in the high school — Intelligence tests for classification — Educational tests for classification — Prognosis and aptitude tests — Educational tests <i>vs.</i> mental tests for classification</p>	

CHAPTER	PAGE
IV. USES OF TESTS FOR RESEARCH PURPOSES .	42
Tests in school investigations	
V. TESTS IN THE MOTIVATION OF LEARNING .	43
Motivation as a neglected aspect of testing	
IV. CRITERIA FOR THE SELECTION OF EDUCATIONAL TESTS	45
General outline	
I. VALIDATION	48
Curricular validation	
II. RELIABILITY	51
Definition of reliability—The importance of reliability— What is a satisfactory degree of reliability?	
III. EASE OF ADMINISTRATION	56
Adequacy of instructions	
IV. OBJECTIVITY OF SCORING	58
Objectivity an essential in good tests	
V. INTERPRETATION OF RESULTS: NORMS	60
Accuracy of norms—Kinds of norms provided	
VI. DIAGNOSTIC VALUE OF THE TEST	64
The evaluation of the diagnostic functions of a test	
VII. NUMBER AND EQUIVALENCE OF DUPLICATE FORMS	65
The need for duplicate forms—Equivalence of forms	
VIII. TIME REQUIREMENTS, COST, MECHANICAL FEATURES, ETC.	66
PART TWO: DESCRIPTIONS OF HIGH SCHOOL TESTS BY SUBJECTS	
V. MATHEMATICS	71
Introduction	

Contents

ix

CHAPTER	PAGE
I. ALGEBRA	72
<i>Holtz First-Year Algebra Scales — Douglass Standard Diagnostic Tests for Elementary Algebra — Illinois Standardized Algebra Tests — Thurstone Vocational Guidance Tests: Algebra — Iowa Placement Examinations: Mathematics Aptitude, MA-1, Revised; Mathematics Training, MT-1, Revised</i>	
Remedial procedures in algebra	
II. GEOMETRY	84
<i>Minnick Geometry Tests — Schorling-Sanford Achievement Test in Plane Geometry — Columbia Research Bureau Plane Geometry Test — Geometry Test of Thurstone Vocational Guidance Tests</i>	
Remedial procedures in geometry	
III. OTHER MATHEMATICS TESTS	91
<i>Rogers Test of Mathematical Ability — Kelley Mathematical Values Test</i>	
Test materials — References	
VI. ENGLISH (LANGUAGE, GRAMMAR, SPELLING, READING, AND COMPOSITION)	97
I. LANGUAGE AND GRAMMAR TESTS	97
Introduction	
<i>Kirby Grammar Test — Wilson Language Error Test — Cross English Test — Pressey Diagnostic Tests in English Composition — Wakefield Diagnostic English Test — Tressler Minimum Essentials Test — Iowa Placement Examinations: English Aptitude, EA-1, Revised; English Training, ET-1, Revised</i>	
Remedial procedures in language and grammar	
II. SPELLING TESTS	108
Introduction	
<i>Sixteen Spelling Scales Standardized in Sentences for Secondary Schools (Seven S Spelling Scales) — Monroe's Timed Sentence Spelling Test, III</i>	
Remedial procedures in spelling	

CHAPTER	PAGE
III. READING TESTS	112
Introduction	
<i>Haggerty Reading Examination, Sigma 3 — Thorndike-McCall Reading Scale — Van Wagenen Reading Scales — English Literature — Monroe Standardized Silent Reading Tests, III — Pressey Technical Vocabularies of the Public School Subjects — Holley Sentence Vocabulary Scale</i>	
Remedial procedures in reading	
IV. ENGLISH COMPOSITION SCALES	123
Introduction — What the composition scales measure	
<i>Hillegas Scale — Hudelson English Composition Scale — Hudelson Maximal Composition Ability Scales — Lewis Composition Scales — Van Wagenen English Composition Scales</i>	
Reliability of composition scales — Remedial procedures in English composition	
V. MISCELLANEOUS TESTS IN ENGLISH	130
Test materials — References	
VII. SCIENCE	136
Introduction	
I. GENERAL SCIENCE	137
<i>Ruch-Popenoe General Science Test — Dvorak General Science Scales — Van Wagenen Reading Scales: General Science</i>	
II. BIOLOGY	141
<i>Michigan Botany Test — Ruch-Cossmann Biology Test — Information Exercises in Biology (Cooprider)</i>	
Remedial procedures in general science and biology	
III. CHEMISTRY	146
Introduction	
<i>Powers General Chemistry Test — Iowa Placement Examinations: Chemistry Aptitude, CA-1, Revised; Chemistry Training, CT-1, Revised</i>	
Other tests in chemistry	

Contents

xi

CHAPTER	PAGE
IV. PHYSICS	152
<i>Iowa Physics Tests — Hughes Physics Scales — Iowa Placement Examinations: Physics Aptitude, PA-1; Physics Training, PT-1 — Thurstone Vocational Guidance Tests: Physics</i>	
Remedial procedures in chemistry and physics	
Test materials — References	
VIII. FOREIGN LANGUAGE	160
I. FRENCH AND SPANISH	160
Introduction	
<i>Henmon French Test — Handschin Modern Language Tests — Columbia Research Bureau French Test — Iowa Placement Examinations: French Training, FT-1, Revised; Spanish Training, ST-1 — Wilkins Prognosis Test in Modern Languages — Iowa Placement Examinations: Foreign Language Aptitude, FA-1, Revised</i>	
Remedial procedures in French and Spanish	
II. LATIN	166
Introduction	
<i>Henmon Latin Tests — Ullman-Kirby Latin Comprehension Test — White Latin Test — Tests in Latin Vocabulary, Latin Derivatives, Latin Verb Forms, and Latin Syntax — Godsey Latin Composition Test — Orleans-Solomon Latin Prognosis Test</i>	
Comparative data on Latin tests	
Remedial procedures in Latin	
Test materials — References	
IX. SOCIAL STUDIES	177
Introduction	
<i>Barr Diagnostic Tests in American History — Pressey-Richards American History Test — Gregory Tests in American History, Test III — Van Wagenen Reading Scale in History — Kepner Background Tests in Social Sciences — Brown-Woody Civics Test</i>	
Test materials — References	

CHAPTER	PAGE
X. VOCATIONAL SUBJECTS	186
I. MECHANICAL AND TRADE TESTS	186
Introduction	
<i>Stenquist Mechanical Aptitude Tests — Thurstone Vocational Guidance Tests — Trade Tests</i>	
II. COMMERCIAL TESTS	191
Introduction	
<i>Blackstone Stenographic Proficiency Tests, Typewriting</i>	
III. MUSIC TESTS	192
<i>Seashore Measures of Musical Talent — Kwalwasser-Ruch Test of Musical Accomplishment</i>	
Remedial procedures in music	
Test materials — References	
XI. SURVEY TESTS	200
Introduction	
<i>Iowa Comprehension Test — Iowa High School Content Examination — Iowa Placement Examinations — Columbia Research Bureau Tests</i>	
Remedial procedures in connection with survey examinations	
Test materials — References	
XII. GENERAL INTELLIGENCE TESTS	212
Introduction	
Individual intelligence tests — Comparative data on group intelligence tests in the high school range — Reliability of intelligence tests — Intelligence test intercorrelations — Prediction of high school marks — Utilization of results of intelligence testing — Limitations of intelligence testing	
Test materials — References	
XIII. JUNIOR HIGH SCHOOL TESTS	227
Introduction — The scope of measurement in the junior high school	
Description of tests for junior high school grades	

Contents

xiii

CHAPTER

PAGE

I. MATHEMATICS TESTS 228

Buckingham Scale for Problems in Arithmetic, Division III — Compass Diagnostic Tests, Form A — Courtis Standard Research Tests, Series B — Monroe General Survey Scales in Arithmetic — Monroe Standardized Reasoning Tests in Arithmetic, Test III — Otis Arithmetic Reasoning Test — Spencer Diagnostic Tests in Arithmetic, III — Stanford Arithmetic Test — Stone Reasoning Test — Woody Arithmetic Scales — Woody-McCall Mixed Fundamentals, Forms I, II, III, and IV — Wisconsin Inventory Test in Arithmetic

II. ENGLISH TESTS 231

Briggs English Form Test — Charters Diagnostic Language and Grammar Tests — Franseen Diagnostic Tests in Language — Ayres Spelling Scale — Iowa Dictation Exercise and Spelling Test — Stanford Dictation Exercise — Chapman-Cook Speed of Reading Test — Gray Standardized Oral Reading Check Tests — Monroe Standardized Silent Reading Tests, Revised — Stanford Reading Examination — Stone Narrative Reading Tests — Thorndike Test of Word Knowledge — Willing Scale for Measuring Written Composition

III. GEOGRAPHY TESTS 235

Buckingham-Stevenson Place Geography Tests — Courtis Supervisory Geography Test — Gregory-Spencer Geography Tests — Hahn-Lackey Geography Scale — Posey-Van Wageningen Geography Scales

IV. VOCATIONAL SUBJECTS 238

Ayres Scale for Measuring Handwriting, Gettysburg Edition — Courtis Standard Practice Test in Handwriting — Freeman Chart for Diagnosing Faults in Handwriting — Thorndike Handwriting Scale. For General Merit of Children's Handwriting — Gates-Strang Health Knowledge Test — Short Scales for Measuring Habits of Good Citizenship — Home Economics Information Tests — Murdoch Sewing Scale — Murdoch Analytic Sewing Scale for Separate Stitches — King-Clark Foods Test

V. SURVEY TESTS 239

Illinois Examination — Lippincott-Chapman Classroom Products Survey Test — Otis Classification Test — Stanford Achievement Test

CHAPTER	PAGE
VI. INTELLIGENCE TESTS	241
<i>Illinois General Intelligence Scale — National Intelligence Test</i>	
Test materials — References	

PART THREE: INFORMAL OBJECTIVE EXAMINATION METHODS

XIV. THE RÔLE OF INFORMAL OBJECTIVE EXAMINATIONS IN HIGH SCHOOL INSTRUCTION	251
--	-----

Introduction — The relation between standardized and unstandardized tests — Limitations of the traditional examination — The subjectivity of the traditional examination — Unreliability of the traditional examination due to limited sampling

XV. TYPES AND CHARACTERISTICS OF OBJECTIVE EXAMINATIONS	266
---	-----

Classification

Recall or Completion Tests — Multiple-Response Tests — The True-False Tests — Matching Exercises — Best Answer or Judgment Tests — Identification Exercises — Rearrangement Tests

Advantages and limitations of objective tests in general

Illustrations of the more common types of objective examinations

XVI. CRITICAL CONSIDERATIONS IN OBJECTIVE EXAMINATION METHODS	282
---	-----

Introduction — Corrections for chance or guessing — Instructions to guess or not to guess — Experimental data on guessing and corrections for chance — Conclusions from the preceding investigation — Results from similar studies — Time allowances for objective tests — Summary and conclusions

Selected references for Part III

PART FOUR: THE CONSTRUCTION OF EDUCATIONAL AND MENTAL TESTS

CHAPTER	PAGE
XVII. THE CONSTRUCTION OF EDUCATIONAL AND MENTAL TESTS	301
I. VALIDATION	301
Introduction	
Setting up a criterion of validity	
The meaning of validity — The criterion or criteria of validity	
Original selection of items — Methods: 1. Textbook Analysis; 2. Analysis of Courses of Study; 3. Analysis of Examination Questions; 4. Pooled Judgments; 5. Rating Scales; 6. Correlations with School Marks; 7. Rise in Percentage of Successes; 8. Correlations against Validated Measures; 9. The Method of Widely Spaced Groups; 10. Validation by the Principle of Social Utility; 11. Psychological and Logical Analysis; 12. Validation by Correlation with Other Measures	
XVIII. THE CONSTRUCTION OF EDUCATIONAL AND MENTAL TESTS (<i>Continued</i>)	329
The experimental try-out of items	
Computation of item difficulties — Scaling or weighting of items	
II. BREAKING THE TEST ITEMS INTO EQUIVALENT FORMS	338
The second try-out — Determination of final time limits	
XIX. THE CONSTRUCTION OF EDUCATIONAL AND MENTAL TESTS (<i>Continued</i>)	343
III. THE DERIVATION OF NORMS	343
The final try-out of the test — Types of norms: 1. Grade norms; 2. Age norms; 3. Percentile norms; 4. <i>T</i> -scores	

TESTS AND MEASUREMENTS IN HIGH SCHOOL INSTRUCTION

CHAPTER ONE

THE PRESENT STATUS OF MEASUREMENT IN SECONDARY SCHOOLS

Historical. The introduction of educational and mental testing into the high school has not shown the same rapid growth which characterizes the applications of standardized measures in elementary education. A great many factors have been responsible for this situation. The retarding influences, for the most part, appear to be transitory in character, and present indications point toward an increasingly greater application and usefulness of test methods in secondary education in the future.

The technique of measurement, like that of curriculum construction, supervision of instruction, or the psychologizing of teaching methods, presents many unique problems in secondary schools which are not paralleled exactly in elementary school practice. Since the elementary school has in the past occupied the more strategic position in educational theory and administration, it is not surprising that the idea of the measurement of classroom products has had its most thorough try-out in the lower schools. Whether the elementary school will continue to lead in the test movement remains to be seen, but there is no doubt at present about the fact that test methods are gaining a secure foothold in the high school, college, and university. Viewed broadly over a period of years, the slower progress of measurement in the high school may prove to be a blessing in disguise, since it will enable the high school teacher or principal and the test builder to avoid many of the mistakes and pitfalls of the early years of experimentation with objective measures in the elementary