MEASUREMENT AND ADJUSTMENT SERIES

EDITED BY LEWIS M. TERMAN

MEASUREMENT IN HIGHER EDUCATION

BY BEN D. WOOD

Assistant Professor of Collegiate Educational Research in Columbia College: Assistant to the Dean of Columbia College

Yonkers-on-Hudson, New York

WORLD BOOK COMPANY

1923

WORLD BOOK COMPANY

THE HOUSE OF APPLIED KNOWLEDGE

Established 1905 by Caspar W. Hodgson
YONKERS-ON-HUDSON, NEW YORK
2126 PRAIRIE AVENUE, CHICAGO

Scientific method in education involves the careful measurement of each child's ability to learn and of the amount that he has learned. It also involves adjustment of organization, subject matter, and methods of instruction to the varying needs and abilities of pupils. This book is the first of a series that will set forth the value, technique, and applications of educational measurement and adjustment. It will present this subject from the point of view of higher education

MAS : WMHE–1

TO

D. G. W. AND A. W. W.

EDITOR'S NOTE

At the request of the author the usual brief introductory statement by the editor has in this volume been replaced by a more extended discussion which appears as Chapter I. In acceding to the author's wishes in this regard the editor has been influenced by the conviction that the significance of the contribution which this book makes to higher education is so great as to justify its introduction to the American Educational Public in a more attention-compelling manner than is possible in the traditional type of editorial statement.

Lewis M. Terman.

ACKNOWLEDGMENTS

THE writer's unmeasured thanks are due, in the first place, to Professor E. L. Thorndike of Teachers College, whose generous counsel has been mainly responsible for whatever scientific merit this volume may possess.

In the second place, acknowledgment is due to Columbia College. This book would have been quite impossible without the active coöperation of the administration of Columbia College in providing an ideal laboratory and in appropriating generous funds for the prosecution of the research on which it is based.

Dean Herbert E. Hawkes of Columbia College, Professor J. J. Coss, Executive Officer of the Staff in Contemporary Civilization, and Professor A. L. Jones, Director of Columbia University Admissions, have had so large and indispensable a part in the production of the book that it is only by their magnanimity that it is issued under the sole authorship of the writer.

The writer wishes to express his indebtedness to many others, too numerous to mention specifically, who have helped in carrying out this research; but special thanks are due to the Staff in Contemporary Civilization, the members of which are largely responsible for the experiment with the new type content examination described in Chapter VIII: Professors Carman, Coss, Evans, Kendrick, Moon, and Weld, Drs. Edman, Graper, Munro, Schneider, and Messrs. Burtt, Dewey, Friess, Gazley, Leigh, Randall, and Tugwell. Thanks are also due to Registrar Grant, to Assistant Registrar Fox, and to Assistant Director of University Admissions Chadwick. To Professor T. L. Kelley, formerly of Teachers College, now

of Stanford University, the writer is grateful for much helpful advice and inspiration given so disinterestedly during the last half-dozen years. Acknowledgment is made of the careful work of Miss G. A. Taylor and of Mrs. Anna W. Scully on the statistical tables and on the manuscript.

It is a matter not only of pleasure but also of pride to the author to make grateful acknowledgment to Raymond M. Weaver, Assistant Professor of English in Columbia University, for his friendship and encouragement during the writing of this book, and for his careful reading of the proofs in their entirety.

Finally, the writer wishes to express his deep appreciation of Professor Terman's generosity in consenting to write an introductory chapter for the book, in addition to the usual Editor's Introduction.

B. D. W.

COLUMBIA UNIVERSITY

CONTENTS

MEASUREMENT IN HIGHER EDUCATION

CHAPTER ONE

INTRODUCTION

By the Editor, Lewis M. Terman

MENTAL and educational tests have come into very general use in elementary and secondary schools throughout the country. There are a great many standard tests of various types intended for use in Grades 1 to 12, and the number of pupils given one or more of them in a year must run into the millions. After a considerable period of question and doubt, it is now everywhere admitted that the results of standard tests in our lower schools have proved their value in the classification of pupils according to ability, in educational guidance, and in measuring the results of our educational efforts.

Only recently, however, have the methods of tests and measurements been adopted for practical use in the institutions of higher learning, the colleges and universities, which were responsible in the first place for launching them upon the educational world. It must be admitted that American education is most conservative at the top, at least in respect to attitude toward innovations based upon psychological or pedagogical experiment. There are doubtless good and sufficient reasons why this is so. The average college teacher has had little professional training along psychological lines, and he is extremely preoccupied with his special subject of instruction. He is likely to feel a just and conscious pride in his immunity to the ephemeral fads which sweep periodically

over the educational horizon and leave so little but disappointment behind them. At the same time, once an educational reform has been fully demonstrated to rest upon a scientific basis there is no need to fear as to its ultimate reception in the average university. There may be a period of watchful waiting, but when the innovation has been accepted its influence is likely to be thoroughgoing and widespread.

Experiments with psychological tests of college students have been reported from time to time since the early work of Cattell back in the "nineties." For a long time, however, the work remained purely in the experimental stage. It is largely due to psychology's dramatic success in the Great War that we had such rapid progress in this field in recent years. Since 1918 scores of minor and several very important researches along this line have been published. Of such studies by far the most thoroughgoing and conclusive is that described in this volume.

The value and interest of this account, which the author has so ably and convincingly presented, go far beyond the facts and circumstances of the particular research concerned. Instead of a dry description of an experiment the author has given us a vital treatment of the entire measurement problem as it presents itself to our institutions of higher learning. His book is therefore an ideal introduction to the subject and ought to be read by every teacher in every university, college and normal school in the United States. Indeed, the issues which the author treats are so fundamental, and his methods of dealing with them have such wide applicability, that his book should prove no less interesting or useful to high school than to college teachers. In no essential respect do the problems of mental and educa-

tional measurement in the high school differ from those in the college. In both cases the canons of scientific procedure are exactly the same; the tactical difficulties to be overcome in instituting a measurement program are the same; the practical uses to be made of results are the same. There is hardly a fact or conclusion in the entire volume that is not as germane to the high school as to the college situation.

The editor believes that the publication of this volume will exert a profound influence upon higher education in the United States, an influence which may prove as real as that of Flexner's report on the more limited field of medical education. The problems with which he deals are problems of fundamental importance in the conduct of any institution of higher learning. His data have been painstakingly gathered; they have been treated by the most approved scientific methods, and the conclusions rest upon a solid foundation of fact.

That the author's data are based upon the *Thorndike* intelligence test and upon certain *Columbia* tests of academic achievement is of course irrelevant to the main purpose of the book. Far from setting up the tests which he used as being in any sense ideal models, the author is aware of their imperfections and has shown the way to improvements. The point is that even when the tests herein described have all been superseded by better ones, as in time they of course will be, the value of this book will not be appreciably impaired.

The reader untrained in statistical method may at first be somewhat mystified by the frequent use of mathematical terms and by the occasional reference to statistical formulæ and procedures. He can be assured, however, that the essential principles involved are really very simple, at least as far as their practical bearings

are concerned. A careful reading of the text need not leave any intelligent person, however innocent of statistical procedure he may be, greatly in the dark with reference to the real significance of facts presented. From the language of statistics there is no escape if we wish to go beyond the limits of personal opinion and individual bias. Worth-while evaluations in higher education will continue to be as rare as they now unhappily are until the rank and file of college and university teachers become able to think in more exact quantitative terms than they are yet accustomed to. Reliability coefficients of .80 to .95 for the new type content examination, as compared with .40 to .60 for the old style essay examination; a correlation of .98 between two scorings of the same new type examination papers, as compared with .60 or .70 for separate gradings of old style essay examinations; correlations of .70 between college success and Thorndike intelligence scores, as compared with .45 between college success and four years of high school marks; a correlation of .46 between scores on a 30-minute section of the Thorndike test and two years of college English marks, as compared with .36 between first semester English marks and second semester English marks;—these are facts of a kind whose significance does not need to be supported by argument.

It may be well to point out that the value of this book does not hinge on the assumption that intelligence can be measured with perfect accuracy, or that the psychology of intelligence is thoroughly understood. As far as the main purposes of the book are concerned the author could without loss have eliminated the term "intelligence" altogether, as carrying implications likely to confuse the issue. By keeping the term he has merely avoided circumlocutions and unnecessary word coinage.

Those who prefer to do so may think of intelligence tests simply as a more or less satisfactory method of arriving at a prediction of an individual's probable success in certain kinds of intellectual work. The value of the prediction may be judged wholly in terms of correlation coefficients or other quantitative, objective evidence. Personal opinion as to what intelligence is or is not may be left out of the question altogether.

The statistical treatment of mental test data has answered once for all the question regarding their validity. Time and again it has been shown that the scores on a single intelligence examination enable us to predict college success as accurately as we can predict it from four years of high school marks. The great value of the intelligence test lies in the fact that it furnishes data not duplicated from any other source. It gives a new line on the student. More than any other kind of information it tells us what grade of work we have a *right* to expect. It gives a favorable starting point for investigating the causes of failures. It enables us to discriminate between the intelligent student whose failure is avoidable and the student whose inferior native ability renders him comparatively non-educable.

The financial aspects of this problem are by no means unworthy of consideration. One university of about 2,500 students and 1½ million annual income spends more than a quarter of a million annually on students who are on probation or who will later be disqualified for poor scholarship. Actual test shows that about 80 per cent of the failing students in this university have a grade of intelligence which renders their failure in such a university practically inevitable.

However, there is no intention to suggest that only the most brilliant young men and women should attend

college. There are colleges and colleges, also departments and departments within the same college. If we had an intelligence index representing the average mental ability in each college, the indices for a hundred representative colleges would be found to cover such an enormous range that the average as well as the brilliant youth could readily select a college that would correspond to his own grade of ability. The purpose of intelligence tests is not to deprive any one of any educational opportunity from which he is fitted by ability to derive normal profit, but rather to enable us to select the type of curriculum from which a given individual can profit, whether he be bright or dull.

It can hardly be denied, however, that the value of a university's product is determined as much by the original quality of the raw material with which it works as it is by salary budget, instructional method or curriculum. It is true that the relatively dull student may and perhaps usually does obtain some benefit from a college education. But a university may "educate" any number of such students without making any appreciable contribution to the advancement of higher learning. Had space permitted, the author of this book could have quoted extensive data from other investigations besides his own showing how greatly the institutions of higher learning in this country differ from one another with respect to the average intellectual quality of their students. Two universities may be regarded by the general public as in all respects equal when an amount of ability below the minimum requisite for bare survival in one would considerably exceed the average in the other. Comparatively speaking, some student bodies are made up largely of intellectual cream while others have drawn almost entirely from the lower levels. Such differences may have been

suspected in some quarters, but only recently has it been possible to measure them.

Just as colleges differ with respect to the intellectual ability of their students, so do the coördinate departments and classes differ within a given college. Generally speaking, students who major in Latin, philosophy, law, mathematics, physics, or chemistry rate higher in intelligence than students electing dentistry, agriculture, household science, art, or music. Of two professors teaching the same subjects in a given department, one may tend to attract bright students while the other attracts mainly the mediocre or inferior.

At Ohio State University the median army Alpha intelligence score of students in the Liberal Arts department was 147; of students in agriculture, 133; of dental students, 115. At the University of Illinois the median Alpha score of law students was 163; of students in agriculture, 139; of music students, 121. At Miami University the median intelligence score necessary to earn the grade A was 162 in philosophy classes, 155 in chemistry, 145 in history, 132 in drawing or music, and 126 in home economics.

For lack of the kinds of facts which tests give us regarding the raw material in our institutions of higher education it has been impossible to evaluate the institutions themselves, their departments, their professors, their curricula, or their methods. If classically trained students were found to make better records in medical school than medical students who had been trained in zoölogy, the fact was explained in terms of the superior mental discipline afforded by the classics. We now know that in some schools the intellectual differences alone are sufficient to account for such a finding.

There are but two essential factors going to make a

university, or for that matter any other kind of school; namely, raw material and educational processes. Each of these factors enters into every sample of the school's product. As long as the raw material is an unknown quantity the merits of the processes also remain unknown and unknowable. Wherever schools are conducted without the use of modern methods of testing the quality of their material, their work is done largely in the dark.

The possible uses of psychological tests as a method of determining fitness for college work in general or fitness to pursue particular courses are especially interesting. It is not yet clear how far it will be found safe to go in omitting various content requirements in the case of students of very superior native ability. Except in certain types of courses it is probable that scholastic success depends far more upon native ability than upon the extent and kind of previous training. A Stanford student whose pre-university training was limited to the completion of the fifth elementary grade has been able, nevertheless, to carry his university work through two years with a record considerably above average. Other cases of this kind have recently been reported from various universities. Certainly a college is justified in permitting the exceptionally able candidate who is short in some of the usual academic requirements to enter by the test route. Properly safeguarded, the plan involves no risk whatever of lowering academic standards. Instead, it puts the emphasis on ability, where it belongs. The candidate who can earn an exceptionally high test score in spite of inadequate training is the best possible bet as regards scholastic promise. This plan of conceding to intellect has the additional advantage that it tends to attract the best and discourage the weakest from entering.