HOW TO MEASURE IN EDUCATION

WILLIAM A. McCALL, PH.D.
ASSOCIATE PROFESSOR OF EDUCATION
AT TEACHERS COLLEGE, COLUMBIA UNIVERSITY

New York
THE MACMILLAN COMPANY
1923

PRINTED IN THE UNITED STATES OF AMERICA

Set up and electrotyped. Published, January, 1922.

PREFACE

The *science* of measurement is in its infancy and the *art* of measurement is younger still. Yet both have developed at such a phenomenal rate in the last few years as to make this movement for the mental measurement of children the most dramatic tendency in modern education. Educators returning from a few years' sojourn in foreign missionary fields testify that no recent change in educational practices compares with that effected by the growth of scientific educational and intelligence testing. During the war, and since, it was my privilege to confer with educational or military commissions sent here by several foreign governments to study our schools. These representatives testified that the extensive use of scientific mental measurement was one of the most distinctive features of American education. This book describes rather fully the meaning and methods of this movement.

The art of measurement is younger than the science of measurement because the abler workers have, of necessity, devoted their energies almost exclusively to the origination of foundational techniques. But the whole movement was so promising in the way of concrete assistance in meeting educational problems that practical educators have irresistibly demanded that the science of measurement be turned into the art of measurement almost overnight. Hence the last two or three years has witnessed a feverish effort to meet these demands. The result has been numerous mistakes and remarkable successes.

This book has a fourfold aim. First, it aims to present these successes and warn against a repetition of the mistakes.

Secondly, it aims to help forward the movement for mak-

ing teaching a genuine profession. Effective teaching requires a skill and a command of refined procedure in excess of that needed by the physician or surgeon. The medical profession is a genuine profession just because its members are experts in refined procedure. This book is one of the two which I have planned to show that tangible, learnable, refined techniques are possible in teaching and that even a "born" teacher will not, in a few years, be able to compete with a professionally *trained* teacher.

In the third place this book aims to meet the needs of the educators who are interested in both the How and the Why. Extremely elementary books in this field have done an admirable service in diffusing an interest in mental measurement. The more intelligent teachers are now asking, however, for a book which not only brings the art of measurement up to date, but which, in addition, goes sufficiently into the science of it that they may be able to use tests less blindly than heretofore.

The final aim of this book is to bring together in one convenient volume most of the techniques needed by those engaged in mental measurement. At present the worker in this field must go to one book to learn how to construct a mental test, to another book to learn how to give and use the results of the test, to another book to learn how to apply statistical methods, and to still another book to discover methods for graphic and tabular presentation. The expert will and should continue to consult these special treatises. Others do not like to give the necessary time. This book then, is really several small books in one. Furthermore, it has been so arranged that the reader can confine himself to the earlier and less technical portions or he can omit this and study the more technical chapters placed toward the end.

As I think through the contents of this book and reflect upon the influences which have made it, I am keenly conscious that it is both autobiography and biography. It is autobiography because several years of my own thought and

labor and much experimentation in the world's most stimulating educational laboratory, Teachers College and New York City, have gone into it.

This book is a biography because from cover to cover I see the tangible and relatively intangible evidences of my many teachers. My pupils have been my teachers. My contemporaries in mental measurement have been my teachers. In so far as my mention can accomplish it, I wish, however, to honor particularly five individuals. The first is, and has been, an elementary school teacher in the mountains of Kentucky and Tennessee. The second is President Emeritus of Cumberland College, Williamsburg, Ky. The third is President of Lincoln Memorial University, Cumberland Gap, Tennessee. The influence of the fourth is responsible for my interest in realms which transcend those represented in this book. The ideas of the fifth are so completely a part of me that it might be said that he wrote the book through an imperfect medium. They are Jesse Worley, E. E. Wood, George A. Hubbell, Frank M. McMurry, and Edward L. Thorndike.

Wm. A. McCall

CONTENTS

PART ONE

How to Use Measurement

PART TWO

How to Construct and Standardize Tests

PART THREE

Tabular, Graphic, and Statistical Methods

LIST OF TABLES

LIST OF DIAGRAMS

PART ONE

HOW TO USE MEASUREMENT

HOW TO MEASURE IN EDUCATION

CHAPTER I

PLACE OF MEASUREMENT IN EDUCATION

THESIS 1. "WHATEVER EXISTS AT ALL, EXISTS IN SOME AMOUNT" [1]

It is possible to become so immersed in the details of the measurement of pupil achievement as to lose sight of its fundamental significance. It is this absence of perspective which is responsible for two educational afflictions—the lop-sided enthusiast for the scientific measurement of education, and the equally unbalanced opponent of the movement. Educational measurement has a sort of philosophy. I have attempted to condense the main elements of this philosophy into a series of theses which may help those who have not had much time to think along these lines to appreciate the true place which measurement should have in education. The first of these theses is stated above.

Since all sane persons accept this thesis it needs no qualification, but a qualified thesis will suffice for our purpose, namely, whatever change the teacher makes in a pupil must be a change in an amount of something. We teachers will scarcely insist that our effort makes no change in amount. Even though such were the result of our effort it would not so much disprove the thesis but rather prove our own inefficiency.

There is an ever-dwindling group who strenuously oppose

[1] E. L. Thorndike, *The Seventeenth Year Book of the National Society for the Study of Education*, Part II, p. 16; Public School Publishing Co., Bloomington, Ill.

the practical implications of the above thesis. They claim to be interested in the emancipation of education from the quantitative idea. Their effort is directed toward the qualitative in education. According to them there is in every person a non-quantative quality—a

> " . . . something far more deeply interfused,
> Whose dwelling is the light of setting suns."

Did they truly "see into the life of things" they would realize that there is never a quantity which does not measure some quality, and never an existing quality that is non-quantitative. Even our halos vary in diameter.

THESIS 2. ANYTHING THAT EXISTS IN AMOUNT CAN BE MEASURED

At least half a dozen scales now exist by which it would have been possible to measure the quality of the Handwriting on the Wall. Faust said:

> "What she reveals not to thy mental sight
> Thou wilt not wrest from her with levers and with screws."

But science has enormously increased the subtlety of levers and screws, and our mental sight is obtuse compared to some of our present-day mental tests. Jesus tacitly accepted and practiced mental measurement when He estimated the quantity of faith on a mustard-seed scale.

It is possible to measure, at least crudely, an individual's love of a sunset or appreciation of opera. Theoretically the thesis is sound but whether practically we shall ever possess sufficient ingenuity to discover all the things that exist in amount and then measure them with any great accuracy, is a question. All that is necessary to accept for the present is that all the abilities and virtues for which education is consciously striving can be measured and be measured better than they ever have been. The measurement of initiative, judgment of relative values, leadership, appreciation of good literature and the like is entirely possible. We already have a scientific scale for the measure-

ment of poetic appreciation. The measurements may not be as exact as we might wish, but they would have value.

THESIS 3. MEASUREMENT IN EDUCATION IS IN GENERAL THE SAME AS MEASUREMENT IN THE PHYSICAL SCIENCES.

The two types of measurement are fundamentally alike because both measure physical manifestation. Neither adding ability, nor good intentions can be measured by plunging a thermometer into a pupil's spiritual medium, but they can be by measuring his behavior and judging his inner condition therefrom. Unless the witness is a habitual liar, psychologists can, with considerable success, determine by means of a breathing curve, when a witness is not telling the truth.

In a still invisible future it may be possible to secure a "movie" of a pupil's mental machinery when in operation and thus secure the desired information but for the present it is necessary to measure the product produced and, if desired, infer the inner condition of the pupil.

Measurement must frequently meet the objection of being too materialistic. Listen to Gilder in "The Poet's Protest."

> "O man with your rule and measure,
> Your tests and analyses!
> You may take your empty pleasure,
> May kill the pine, if you please,
> You may count the rings and the seasons,
> May hold the sap to the sun,
> You may guess at the ways and the reasons
> Till your little day is done."

To parody Wagner in "The Better Way," one would think that it was the purpose to measure human worth by the ell, the value of a life by the number of its years, the painter's canvas by the yard, or the work of the poet by the pound or bushel. A student writes: "Measurement should not be applied where spiritual factors and ideal values are involved." Those educators who protest most violently against any such measurement of the pupil are daily probing

his mental activity by methods which are comparable to the surgical operations of bygone ages. They find themselves in a position of disapproving the lover who estimates his lady's affection by the radius of the pupil of her eye under standardized lighting, and of approving the scientific father who soothes the mother for his punishment of their infant by saying: "I am not slapping an innocent soul but spanking a physiological reaction."

THESIS 4. ALL MEASUREMENTS IN THE PHYSICAL SCIENCES ARE NOT PERFECT

Physical measurements are, in general, more exact than educational measurements but education has no monopoly upon imperfect tests. There are tests which are now the rule in physical sciences for which an expert in educational measurements would blush. The general superiority of physical measurements is not due to the fact that they are radically different in kind. Physical measurements are subject to practically all the errors which trouble educational measurement. It is not that they do not exist in the former, but that they usually exist in such small amounts that the average person fails to see them. They are large enough to be the despair of experts in the various sciences. Thorndike [2] has given us an excellent statement of this point:

"Nobody need be disturbed at these unfavorable contrasts between measurements of educational products and measurements of mass, density, velocity, temperature, quantity of electricity, and the like. The zero of temperature was located only a few years ago, and the equality of the units of the temperature-scale rests upon rather intricate and subtle presuppositions. At least, I venture to assert that not one in four of, say, the judges of the Supreme Court, bishops of our churches, and governors of our states could tell clearly and adequately what these presuppositions are. Our measurements of educational products would not at

[2] Op. cit. p. 18.

present be entirely safe grounds on which to extol or condemn a system of teaching reading or arithmetic, but many of them are far superior to measurements whereby our courts of law decide that one trade-mark is an infringement on another."

But the imperfections of educational measurements are, in general, far more glaring than the majority of those made in physics, chemistry and like sciences. Some may have gotten the impression from certain emotional and quixotic radicals that standard tests are perfect instruments. This is far from the truth. They have numerous and decided limitations. The recent somewhat unbalanced, but doubtless necessary, propaganda is justified not because of the perfection of the tests recommended, but the much greater imperfection of the tests or lack of tests now in use.

A common criticism of educational measurement is that the tests measure a narrow, limited segment of a pupil's totality. Physical measurements tend to be more handicapped in this respect than educational measurements. Most of their measurements, such as measurements of length, width, weight, and temperature are exceedingly narrow abstractions and they are exceedingly useful too. A totality test for a pupil would certainly be useful but if we possessed one we would proceed immediately to construct tests for the detailed measurement of pupil abilities. Scales for the measurement of composition are useful, but scales for the measurement of the elements which go to make up composition are also useful. Teachers not only teach children "all over"; they teach them in detail. If tests are to aid instruction effectively, there is as much need for them to measure in detail as in totality.

THESIS 5. MEASUREMENT IS INDISPENSABLE TO THE GROWTH OF SCIENTIFIC EDUCATION

Exact measurement has made possible the rapid progress in the natural sciences. It has been stated that the amount