HOW TO EXPERIMENT IN EDUCATION

BY

WILLIAM A. McCALL, Ph.D.

ASSOCIATE PROFESSOR OF EDUCATION, TEACHERS COLLEGE, COLUMBIA UNIVERSITY, NEW YORK CITY

New York

THE MACMILLAN COMPANY

1923

EDITOR'S INTRODUCTORY NOTE

Professor McCall has written this book primarily for the purpose of presenting the methodology of educational experimentation in a practical form for the use of teachers and students of education who wish to engage in experimental work, or who desire to understand the great amount of experimental literature which is appearing in magazine and book form. This is the first book on educational experimentation to be published at home or abroad. There are philosophical treatises on scientific methodology, such as Pearson's "Grammar of Science," and a few scattered suggestions on the method of experimental education in books on scientific education; but there has been no adequate treatment of experimental work in the educational field. This fact led the present writer, when he became editor of the Experimental Education Series, to ask Dr. McCall to prepare this volume. Dr. McCall has conducted courses in Teachers College in the field of experimental education, and he has for a number of years been accumulating concrete data to illustrate the experimental method of procedure. Probably no one is as well equipped as he is to prepare a book for the guidance of all who desire either to understand or to undertake experimental work in education.

With the aid to be gained from this book, intelligent teachers can engage profitably in research work in education even if they are not technically trained in experimental methods. The subject is one of permanent worth; and students of education or teachers who wish to gain an intelligent appreciation of and to keep in touch with American educational progress must be familiar with, and, to some extent at least, must be master of the methodology of educational experimentation. A large proportion of popular educational doctrines has been derived without due regard to the requirements for securing valid conclusions; and it may be safely predicted that superintendents, principals, and teachers, as well as students of education, who read Professor McCall's book understandingly will exercise greater care than they have done heretofore in promulgating educational principles based upon data that have not been secured in an accurate manner or treated according to a technique designed to control or eliminate disturbing or irrelevant factors.

"How to Experiment in Education" is not as technical as it might appear to be at first glance. The formulæ and diagrams as well as the discussion can be easily understood by any reader, even though untrained in experimental methods, if he will begin at the beginning of the work and go through it systematically and leisurely. Concrete examples of experimental problems that have been or that might be successfully studied are described by Professor McCall frequently and clearly enough to illustrate every method of procedure discussed and every diagram presented. Technical terms are sparingly used, and the meaning of those that are employed can be easily gained from the context in which they appear.

M. V. O'Shea.

The University of Wisconsin.

PREFACE

My initiation into educational research, like most initiations, was a rather tragic one with happy consequences. My professors plunged me into practical research situations when my training in experimentation was exceedingly lopsided. They trusted to my genius to supply the missing half of research methodology. The memory of this mistaken trust constitutes the pleasant after effects.

The cause of my tragedy and of others like mine was due to the fact that, heretofore, chief attention has been directed toward statistical refinements, rather than refinements of pre-statistical procedure. There are excellent books and courses of instruction dealing with the statistical manipulation of experimental data, but there is little help to be found on the methods of securing adequate and proper data to which to apply statistical procedure. Training is given and books exist only for the last step of a several-step process. As a result, the final step often becomes little more than statistical doctoring for the ills in the data.

This book, together with its predecessor, "How to Measure in Education," but particularly this book, represents an attempt to assemble or originate a fairly complete methodology of research from the selection of the problem to the conclusion of the research. Material has been drawn from numerous sources, but the largest single source is that unannounced richest course of instruction taken by me at Teachers College, namely, the frequent privilege of out-of-course association with Professor E. L. Thorndike.

The encouragement and support given my work by my departmental Superiors, Professors M. B. Hillegas and Frank M. McMurry, and by Dean James E. Russell have been a continuous surprise because they have exceeded every expectation. Such encouragement has made it a pleasure to shorten vacations and to lengthen the working day so as to finish this book before departing for a year of service with the Chinese National Association for the Promotion of Education.

It is fortunate for the future reader that I am in China while this book is being edited and published. As a result, Dr. M. V. O'Shea has given an unusual amount of time to its editing, and in this he has had the technical assistance of Dr. John G. Fowlkes. Miss Harriet Barthelmess, who has a thorough knowledge of the methodology of experimentation, and my wife, Alma McCall, have volunteered to read the proof. I wish to make grateful acknowledgment of their kindness.

William A. McCall.

Teachers College
Columbia University

CONTENTS

HOW TO EXPERIMENT IN EDUCATION

CHAPTER I

SELECTION AND FORMULATION OF EXPERIMENTAL PROBLEM

I. Value and Prevalence of Experimentation in Education

Prevalence of Experimentation.—Except for sporadic exceptions and for continuous overlapping, the method for the determination of truth has passed through three major stages. The first stage is that of *authority*. When any question arose as to the truth or falsity of any fact or principle, it was referred by consent or force to the oracle, chief, king, church, state, or other temporarily ascendant individual or group. In the year 1922 the legislature of a certain state decided by vote whether the principle of evolution is true or false. In this same year there were further occasional evidences that vital educational matters were still being decided on the basis of authority and authority alone.

The second stage is that of *speculation*. This represents a genuine advance. When this stage was reached, questions were no longer matters merely to be settled; they were matters to be freely discussed. Broadly speaking, America and American education have now advanced well into this stage.

The third stage is that of *hypothesis* and *experimentation*. This stage is not something perceived only in visions. We have seen enough of it to know its aspect and to appraise its promise. Since earliest times a tiny stream of scientific research has trickled through the ages, now above ground, now below, now a dashing stream, now a desert rill, but always flowing forward toward the future, and, in late years, increasing greatly in volume. Today, educational experimentation is accepted but not achieved.

These three, authority, speculation, and experimentation, have been described as stages, and in a sense they are. But, in a truer sense, they supplement each other. Speculation, unless it becomes an end in itself, is a fruitful source of hypotheses or problems for research. Authority, when founded upon tested knowledge rather than upon pure opinion, has an essential function in the scheme of life and education.

Everywhere there are evidences of an increasing tendency to evaluate educational procedures experimentally. Though measurement alone is not research, the marvelous spread of the movement for scientific measurement of educational products is a symptom of a new attitude which is favorable for research. The establishment of numerous city and state bureaus of research is another evidence. Numerous experimental schools have arisen for the purpose of research, pseudo-research, or propaganda. Most of the departments of the better teachers colleges have become saturated with the new point of view. Scientific organizations, research committees, an institute of educational research, and large educational foundations are lending such impetus as make experimental education the most important current movement in education.

But even with all its growth we have barely entered the stage of experimentation. Most educational theory still needs testing. Adequate testing of theory requires a rigid scientific procedure. The technique of experimentation is possessed today, with a few exceptions, mainly by a small group of educational psychologists. Experimental education cannot hope to cope with its great task or develop much faster so long as superintendents, principals, and supervisors, not to mention teachers, are not equipped to solve their own problems for themselves. It is but a question of time until educational leaders will be required to have a command of research technique. Then the third stage has a chance to arrive.

Value of Experimentation. — Experimentation has proved its worth by hastening the day when the test of truth will be verification and conformity to our experience rather than revelation and miraculous departure from our experience. Science asks us to believe in such unthinkable things as the reality of ether, the absence of weight and friction for celestial bodies, the existence of the atom, that food makes thought, and the like. But these matters are in conformity with logic or experimental evidence. As Burroughs states, the helium atom has been proved to be an objective entity as truly as that the sun is in heaven.

The practice of experimentation in a school or school system pays in terms of an altered attitude on the part of the entire staff, willingness to consider new proposals, and an alertness for new methods and devices. Experimentation ploughs up the mental field. Teachers join their pupils in becoming question askers. It is the absence of just such stirrings of the mental soil, which, in all probability, is responsible for the supposed fact that teachers fail to improve after a few years of experience.

Experimentation pays in terms of cash. Three years ago an experiment was conducted in a school of five hundred pupils. The purpose of the experiment was to evaluate a group of teaching methods. A careful account was kept of the increased ability secured. Careful estimates were made of its financial value. A record was kept of expenditures. The value of the increased abilities secured was estimated to be worth $10,000. This estimate was based upon the total cost in previous years of producing each unit of ability. The cost of test material used, and of the special supervision required, amounted to $540. The net annual saving, not counting future compounding of the abilities, was $9,460.

Recently an experiment has been conducted by Dransfield, principal of a school in West New York, New Jersey, and by Barton, superintendent of schools at Sapulpa, Oklahoma. The purpose of these experiments was to evaluate the plan for the teaching of reading described in "How to Measure in Education." The total points of A. Q. growth in reading in the control school were 60. The points of growth in the experimental school were 143. Even without taking into account the improvement in history, geography, arithmetic, etc., resulting from increased reading ability, or the cumulative value to the pupils in future years, and even without considering that the teachers have learned a new process to use with other pupils, still the difference between the two groups is worth thousands of dollars. Consider the value to education of this and similar experiments, when their influence shall have spread to the millions of pupils in American schools.

The foregoing experiments have been described to show that it is not unreasonable to claim that a widespread use of scientific research could so increase the efficiency of instruction as to save a year of instruction. The value of such an achievement in financial terms is shown by the following approximate figures:

Population of the United States	103,600,000
Saving to each person through research	1 yr.
Total saving	103,600,000 yrs.
Value of a year	$1,000
Saving for U. S.	$103,600,000,000
Population engaged in World War	1,300,000,000
Saving for World War Powers	$1,346,800,000,000
Saving for 100 generations	$134,680,000,000,000

$134,680,000,000,000 = 260 times U. S. Wealth = 790 times cost of World War = 395 times cost of all wars in recorded history.

Experimentation will pay the nation, the school system, and the individual school. The time has now arrived when it also pays the individuals who engage in it. If the financial reward is not large, the esteem of the profession is.

There is no denying the fact that those educators who today are constructively studying educational problems by scientific methods have achieved, or are destined to achieve, positions of recognized leadership in education. They become the final arbiters for most educational questions, for the peculiar function of experimentation in education is to be a court of last resort.

Methodology of Research.—Scientific educational research may be grouped conveniently into three major divisions,—descriptive investigations, experimental investigations, and causal investigations. The purpose of descriptive investigations is to describe a situation as accurately and objectively and quantitatively as possible. They involve the collection of data, and the quantitative description of the data by the following means: some mass measure, such as a frequency distribution, frequency surface, order distribution, or rank distribution; or some point measure, such as a mode, mean, median, midscore, or percentile; or some variability measure, such as a quartile deviation, median deviation, mean deviation, or standard deviation; or some relationship measure, such as a scatter diagram, contingency table, or coefficient of correlation; or some reliability measure, such as a standard deviation of the measure, or probable error of the measure; or some other of the standard statistical techniques, such as are described in Rugg's "Application of Statistical Methods to Education," or Thorndike's "Mental and Social Measurements."

The purpose of experimental investigations is to evaluate the methods, materials, and aims of education. It is to determine the absolute or relative effects upon some subject or subjects or pupils of one or more experimental factors.

The purpose of causal investigations is to start with some observed effect and locate the cause or causes; to determine whether hypothetical causes are really causes; or to determine just how much each of several causes contributes to produce the effect.

McCall's "How to Measure in Education" has for its purpose not only to tell how to use practically and construct scientifically mental and educational tests, but also to present the measurement, tabular, graphic, and statistical techniques required for the conduct of descriptive investigations. This book is a sort of companion volume for "How to Measure in Education," and has for its purpose to complete the presentation of the methodology of research. The first book covers descriptive investigations. This book presents the techniques for experimental and causal investigations.

II. Selection of Experimental Problem

Planning an Experiment.—An experimenter ought to think through his experiment from the conception of the problem to the formulation of the conclusions and beyond. If he has six months to devote to an experiment he can, with advantage, spend five months in planning the experiment and one month in conducting it. Ideally an experimenter should not start his experiment until he has gone through, mentally at least, every step even down to the smallest statistical detail. Those who do not possess a vivid imagination can advantageously carry a miniature experiment with hypothetical data through the various tabulation and statistical stages.

The importance of adequate planning cannot easily be exaggerated. There is little justification for the contention that a well-prepared plan is an inflexible plan. A plan can be thorough and yet plastic enough to be altered to meet unexpected emergencies. In fact original adequacy of plan is probably correlated positively with a healthful plasticity.

Whenever the experimenter can afford the time, an actual-trial experiment is superior to a mental-trial experiment. Even the keenest vision of the most experienced experimenter cannot always foresee every difficulty which will arise. Hence the theoretically best procedure is to follow the mental-trial experiment with the actual-trial experiment, to modify and perfect the plan in the light of the actual trial, and, finally, to conduct the real experiment.

How to Find Experimental Problems.—The best way to find genuine experimental problems is to become a scholar in one or more specialties as early as possible. Thorndike has done a great service for the cause of original research by showing, in a convincing way, that the original mind is the informed mind. The idea that much knowledge hampers a man's originality has taken deep root in the popular fancy, as a result of its self-deceptive search for some crumb of comfort for stupidity. The essence of originality is high native intelligence plus adequate knowledge. Spencer describes knowledge as a sphere of light floating in an abyss of darkness. As a rule, only those who live their mental life on or in this sphere conceive fruitful problems.

A second way to discover fruitful problems is to read, listen, and work critically and reflectively. It is well to form the habit of reacting upon every situation with a question mark, and to consider every untested theory as an hypothesis. Between the lines of every worthwhile book are enough problems and enough rich materials to make the finder and utilizer famous.

A third method of discovering fruitful problems is to consider every obstacle an opportunity for the exercise of ingenuity instead of an insuperable barrier. A king once placed a purse full of gold in the middle of a public road. On the purse he placed a large stone. A soldier with his head in the air and whistling a tune chanced that way. He roundly cursed those who drove over that road for not removing the stone and hence for the injury to his pride and person. A wagoner, with the expenditure of much emotion and considerable skill, maneuvered his wagon past the obstacle. Since no one who passed that way had formed the mental habit of considering every obstacle an opportunity, the reward beneath the obstacle went by default to the king.

A fourth method of finding problems is to start a research and watch problems bud out of it. The very process of research stirs up a hornet's nest of insistent problems. Spencer expressed a profound truth when he said that if we enlarge ever so little the sphere of light we increase infinitely its points of contact with the darkness.

A fifth method of finding problems is not to lose those already found. Almost everyone has probably been given for a moment—probably some odd and unexpected moment—some rare insight. These flashes come, linger for a moment, go, and are forgotten beyond recall. Twiss attributed his rise to a university position to one fact. He bought a steel filing case and recorded and filed original ideas and problems before they were forgotten. So vital for professional growth is this matter of finding and recording problems, that the worth of an educator can probably be measured by asking him to list in ten minutes as many as he can of worth-while educational problems.

What Experimental Problem to Select.—It goes without saying, and yet it needs to be said, that experimenters should select problems whose solution is not already known. One of the abler men in educational measurement reported, at a recent gathering of scientific workers, the results of a painstaking and exceptionally original research. Unfortunately the same problem had already been solved and the results published. Thorndike tells of a student who submitted to him the results of a research which the candidate hoped would be acceptable for a Ph.D. thesis. In submitting the manuscript the candidate wrote that he knew the research was original for he had been careful to avoid reading anything whatever about the subject.

As a rule, an experimenter should select and work upon problems in his own specialty. It will be shown later that successful experimentation requires such a detailed knowledge of the factors operating in a particular situation, and of the influence of these factors, as only a trained and experienced individual possesses. Recently, some students of experimentation, who were reasonably expert in education only, attempted to plan an experiment in chemistry. The undertaking was soon abandoned. No one seemed to know the influence of temperature upon certain chemical reactions. This necessity of intimate knowledge probably explains why over 99 per cent of all discoveries are made by experts in the field of discovery. During the World War, the War Department established a clearing house for popular inventions. A few valuable suggestions were received, but in the main the bulk of all research had to be done by a mere handful of experts.

An experimenter should select the relatively more vital problems. There are many problems which are worth solving but not relatively worth solving. The number of those willing or competent to undertake research is too small and their time too valuable to expend effort on problems not of vital consequence.

An experimenter should select a problem whose solution is feasible, and should set up hypotheses capable of proof. However vital the hypothesis, if it is not susceptible of proof it should be discarded, for the present at least. Unfortunately, the solution of many experimental problems of great worth is often not feasible, because needed tests have not been constructed, or because appropriate subjects are not available, or because the experimenter cannot sufficiently control the situation in which the proposed experiment is to be conducted, or for some other reason. Thus, the excellence of an experimental problem depends upon several factors, and hence it should be selected in the light of these factors. A more comprehensive list of these conditioning factors will be given later.

III. Formulation of Experimental Problem

Types of Formulation.—There are three types of individuals engaged in educational research, and the types are clearly indicated by the way they formulate their problems.

The first type of experimenter "flutters in all directions and flies in none!" He formulates problems so that their scope is scarcely less wide than the universe. Such broad formulations offer little practical aid in planning the details of an experiment. Gazing at the stars, this experimenter steps into every snare at his feet. Just as a teacher cannot teach arithmetic in general, or spelling in general, but, instead, must teach particular examples or particular words, so an experimenter is likely to think and act very irrelevantly if he is guided by a broad formulation only.

Recently an experimenter came for consultation about a problem which he had formulated thus: What is the effect of various factors upon learning? After a little urging he departed and returned later with this formulation: What are the effects of distribution of time upon learning? He was commended for the improvement made. At a later stage the problem had become: Will a typical fourth-grade class in silent reading, spending three thirty-minute periods per week, accomplish more or less than an equivalent class spending five periods of eighteen minutes each per week? Even this is too broad for a final working formulation.

The second type may be called the *pot-hole* type. Near the Cumberland Falls, the Cumberland River has a stone bed pitted with pot-holes. These holes were made by small hard pebbles which lodged in originally slight concavities and which, due to the action of the water, have ground round and round, thereby making the pebbles smaller and the hole wider and deeper. There are indefatigable individuals engaged in educational research whose experimental problems are admirably specific. They are as narrow as the pebbles in the pot-hole. And, like the pebbles, their problems become narrower and narrower as their research proceeds. Such experimenters are experimental drudges. They do much excellent work, but each research is isolated from every other. There is an absence of general plan. There is no mental reaching for the larger implications. They are as lop-sided as the first type.

The third type of experimenter is the truly admirable one.

He is the scholarly type. He perceives the larger meanings of each minute investigation. This glorifies the drudgery inherent in all careful research. The scholarly experimenter first formulates a broad problem. This gives the larger goal and permits perspective. He then breaks up the broad problem into very narrow, specific problems. These are the working units. As the results from the specific investigations come in, he fits the bits together into a beautiful mosaic. The solution of any one specific problem may be of no practical value. It merely contributes to the solution of the larger problem which alone has genuine practical significance. Hence, it is desirable that there be a hierarchy of formulations from very broad to very specific.

A working formulation of an experimental problem should clearly describe: (1) the experimental factor or factors whose effect or effects are being studied, (2) the experimental subjects or individuals or pupils to whom the experimental factor or factors are to be applied, and who are expected to register the effect or effects, (3) the nature of the effects expected and to be measured. In sum, a working formulation requires that the experimenter must have analyzed his problem in rough outline at least.

Why and When to Survey Bibliography on a Problem.—The time to make a survey of the bibliography on an experimental problem is the opposite of the time when the survey is all too frequently made. Often an investigator has completed his experiment and has prepared his manuscript for publication before he hurriedly collects a list of references. The prime function of a bibliographical survey is not to provide a dignified list of references to append to an article, but to serve as a practical guide to the formulation of the subordinate problems, and to the general planning of the investigation. Hence, the survey of the bibliography should immediately follow the formulation of the experimental problem or problems.

If there were no other reason, self-respect as a scholar should be adequate motivation for surveying a bibliography.