

# EXPERIMENTATION and MEASUREMENT in RELIGIOUS EDUCATION

---

GOODWIN B. WATSON, Ph.D.

Assistant Professor of Education, Teachers College, Columbia University  
Director of Research, Home Division, National Council  
of the Y M C A

ASSOCIATION PRESS

NEW YORK: 347 MADISON AVENUE

1927

COPYRIGHT, 1927, BY  
THE GENERAL BOARD OF  
THE YOUNG MEN'S CHRISTIAN ASSOCIATION

---

PRINTED IN UNITED STATES OF AMERICA

## INTRODUCTION

This presentation is intended for students in undergraduate and graduate courses in religious education who are anxious to secure a more adequate understanding of the experimental viewpoint and of the necessary techniques. It is hoped that it may prove useful, as well, to enterprising pastors, directors of religious education, secretaries of the Y M C A or Y W C A, and similar professional workers in the field of religious education. Officials charged with the supervision of large areas, writers of curricula, and persons responsible for administrative policy may find suggestions which will help them to establish a new method of work.

The needs of so varied a group demand a text which shall contain some explanations unnecessary for advanced students, and some techniques too complicated for the service of the average experimenter. The author hopes that readers will excuse his errors of judgment in selection and organization with the thought that somewhere in so diverse a population there may be those whose needs will be served.

Little is offered which is original, save the applications to the complex problems of character development and religious education. Experimental forms, methods of scale construction and test making, and statistical methods may all be found discussed more fully in other volumes. It has been the attempt of the author to bring together in one volume for convenient study and reference by the experimenter in religious education the most essential materials, previously scattered through periodicals and dozens of books.

The reader inexperienced in the forms and formulae here presented may find it advantageous to read through the entire book first, passing over the points which are not clear. Later he may wish to read it again more carefully, working forward

and backward, using index and glossary, and working out the formulae and statistical processes. The first survey will perhaps open up the field of possibilities in experimentation and measurement, and lead to awareness of techniques which need to be acquired. The book is so written, however, as not only to serve this need but to form a handbook of reference for the more experienced student.

Fortunately, this book will often be used by those who are actually carrying forward experiments. It can do them little harm. Having once tasted the thrill of the experimental establishment of previously undiscovered truth, no presentation of schemes and formulae, however inept, will long deter them from further indulgence. The sharp barb of curiosity, the patient eagerness of search, the culminating mastery of accomplishment speak with a potency which words will not convey.

It is the author's privilege to acknowledge indebtedness to Professor George Albert Coe for stimulating viewpoints which are implicit in the philosophy of religion here assumed. To Thorndike, Garrett, Otis, and particularly to William A. McCall the author is anxious to give credit for all the best statistical formulae, tables, and explanations. To his associate, Ralph B. Spence, and to his wife, Gladys H. Watson, are due thanks for reading the manuscript and proof, and for numerous useful suggestions.

GOODWIN B. WATSON

New York City,  
*July 1, 1926*

## TABLE OF CONTENTS

LIST OF TABLES . . . . .	viii
INTRODUCTION . . . . .	xi
CHAPTER	PAGE
I. THE SCIENTIFIC APPROACH TO RELIGIOUS EDUCATION . . . . .	1
1. Magic <i>versus</i> Scientific Method . . . . .	1
2. Magic and Method in Religious Work . . . . .	2
3. Problems Not Answered by Experiment . . . . .	3
4. Experiments Needed Today . . . . .	4
5. Pure Research and Practical Research . . . . .	5
6. Sources of Problems for Practical Research . . . . .	6
II. THE CHOICE OF EXPERIMENTAL METHOD . . . . .	10
1. Principles Governing Choice of Problem . . . . .	10
2. Schemes for Organizing Experiments . . . . .	13
<i>a.</i> Single-Group Methods	
<i>b.</i> Equated-Group Methods	
<i>c.</i> Rotation Methods	
<i>d.</i> Survey Methods	
3. General Advices . . . . .	30
<i>a.</i> Miniature Experiments Advisable	
<i>b.</i> The Diary	
<i>c.</i> Experiments with Individuals	
III. METHODS OF MEASUREMENT . . . . .	34
1. Observation as Measurement . . . . .	34
<i>a.</i> Whatever Exists Can Be Measured	
<i>b.</i> Methods of Refining Crude Observation	
<i>c.</i> Value of Ratings	
<i>d.</i> Reliability of Measures	
<i>e.</i> Validity of Measures	
2. Objective Records as Measures . . . . .	53
<i>a.</i> Methods of Collecting Objective Data	
3. Paper and Pencil Tests . . . . .	57
<i>a.</i> Tests <i>versus</i> Questionnaires	
<i>b.</i> Ethical Problems in Testing	
4. Conduct Tests . . . . .	59
5. Interview Methods . . . . .	61
6. Case Studies . . . . .	62
IV. TESTS NOW AVAILABLE . . . . .	67
1. The Growing Field . . . . .	67
2. Intelligence and Vocational Tests . . . . .	68

CHAPTER	PAGE
3. Description of Tests Now Available . . . . .	70
Biblical Knowledge Test (Whitley)	
Bogardus Social Distance Test	
Brief Test in Religious Education	
Brotmarkle Comparison Test	
Church School Examination Alpha	
Colgate Emotional Hygiene Test	
Downey Will-temperament Test	
Emotional History Record	
Fernald Achievement Capacity Test	
Fernald Ethical Discrimination Test	36,035
Fernald Ethical Perception Test	27,105
Freyd's Occupational Interest Blank	
Giles, Sunday School Examination A	
Hart Personnel Assayer	
Hart Test of Social Attitudes and Interests	
Interest Analysis	
Kent Rosanoff Free Association Test	
Koh's Ethical Discrimination Test	
Laycock's Test of Biblical Information	
Lundholm, Emotional Cross-out Test	
Miner's Analysis of Work Interest Blank	
Multiple-choice Test of Religious Ideas	
Pressey X-O Tests for Investigating the Emotions	
Racial Attitudes Test	
Social Relations Test	
Survey of Public Opinion on Some Religious and Economic Issues (Watson Prejudice Test)	
Union Test of Ethical Judgment	
Union Test of Religious Ideas	
Upton Chassell Citizenship Scale	
Woodworth Matthews Questionnaire	
4. Valuable Tests Not Now Available . . . . .	97
<i>a.</i> Case's A True-false Test in Religious Education	
<i>b.</i> Chapman's A Test of Motives	
<i>c.</i> Character Education Inquiry Tests	
<i>d.</i> Chassell Parable Interpretation Test	
<i>e.</i> Fahs Biblical Test	
<i>f.</i> Fernald, Meritorious Acts	
<i>g.</i> Hartshorne, Tracy, Unfinished Stories	
<i>h.</i> Self-ordinary Ideal Rating Scale	
<i>i.</i> Laslett, Controlled Association Test	
<i>j.</i> Orr, Good Manners Test	
<i>k.</i> Porter Advanced Bible Test	
<i>l.</i> Porter, Student Opinion on War	
<i>m.</i> Raubenheimer Character Preference Test	
<i>n.</i> Schwesinger Socio-ethical Vocabulary Test	

# TABLE OF CONTENTS

v

## CHAPTER

## PAGE

<i>o.</i> Travis, Test of Personality Traits	
<i>p.</i> Van Wagenen Character Judgment Scale in American History	
<i>q.</i> Y M C A Religious Education and Character Growth Tests	
5. Life Situation, Conduct, and Behavior Tests . . . . .	106
<i>a.</i> Aggressiveness	
<i>b.</i> Caution	
<i>c.</i> Cheerfulness	
<i>d.</i> Civic Duty	
<i>e.</i> Concentration	
<i>f.</i> Confidence	
<i>g.</i> Conformity	
<i>h.</i> Courtesy	
<i>i.</i> Decision	
<i>j.</i> Delinquency	
<i>k.</i> Emotionality	
<i>l.</i> Group Loyalty	
<i>m.</i> Helpful Behavior	
<i>n.</i> Home Environment	
<i>o.</i> Honesty	
<i>p.</i> Honest Confession	
<i>q.</i> Humor	
<i>r.</i> Imagination	
<i>s.</i> Interest	
<i>t.</i> Negativism	
<i>u.</i> Persistence	
<i>v.</i> Recklessness	
<i>w.</i> Sociability	
<i>x.</i> Social Perception	
<i>y.</i> Studiousness	
<i>z.</i> Suggestibility	
6. Use of Standardized Tests . . . . .	122
<i>a.</i> Discussion of Test Results	
<i>b.</i> Cooperation in Standardization	
V. THE CONSTRUCTION OF TESTS . . . . .	126
1. Definition of What Is to Be Measured . . . . .	127
2. Selection of Content . . . . .	127
3. Form for Test Elements . . . . .	128
<i>a.</i> New Type and Essay Type	
<i>b.</i> True-false	
<i>c.</i> Two-answer	
<i>d.</i> Completion	
<i>e.</i> Word-phrase Answer	
<i>f.</i> Multiple Choice	
<i>g.</i> Degrees	
<i>h.</i> Ranking	

CHAPTER	PAGE
<i>i. Matching and Pairing</i>	
4. Other Significant Classifications . . . . .	145
5. Development of Directions . . . . .	148
6. Methods of Determining Scoring Scheme . . . . .	150
<i>a. Scoring When Best Answers are Known</i>	
<i>b. Scoring on Basis of General Criterion</i>	
<i>c. Technique of Criterion Score Process</i>	
<i>d. Weighting Test Elements</i>	
7. Standardization of Tests . . . . .	159
<i>a. Objectivity</i>	
<i>b. Reliability</i>	
<i>c. Validity</i>	
<i>d. Comparability</i>	
8. Criteria for a Good Test . . . . .	166
VI. STATISTICAL METHODS . . . . .	171
1. The Service of Statistics . . . . .	171
2. Distribution of Scores . . . . .	176
<i>a. Order</i>	
<i>b. Rank</i>	
<i>c. Frequency</i>	
<i>d. Graph</i>	
3. Measures of Central Tendency . . . . .	183
<i>a. Mean</i>	
<i>b. Median</i>	
<i>c. Mode</i>	
4. Other Significant Points . . . . .	187
<i>a. High Quartile</i>	
<i>b. Low Quartile</i>	
<i>c. Deciles</i>	
<i>d. Percentiles</i>	
<i>e. Marks</i>	
5. Measures of Variability . . . . .	191
<i>a. Range</i>	
<i>b. Average Deviation</i>	
<i>c. Standard Deviation</i>	
<i>d. Q</i>	
6. Measures of Relationship . . . . .	197
<i>a. Graphic Measures</i>	
<i>b. Correlation</i>	
<i>c. Regression</i>	
<i>d. Critical Score</i>	
<i>e. Partial Correlation</i>	
<i>f. Partial Regression</i>	
<i>g. Multiple Correlation</i>	
7. Practical Problem in Five Variables . . . . .	224
8. Measures of Reliability . . . . .	237
<i>a. S. D. and P. E.</i>	



# TABLE OF CONTENTS

vii

## CHAPTER

## PAGE

	<i>b.</i> Reliability of Prediction from a Regression Equation	
	<i>c.</i> Reliability of a Difference	
	<i>d.</i> Limitations of Reliability	
VII.	PRESENTATION OF THE RESULTS OF EXPERIMENTATION . . . . .	245
	1. Choice of Audience . . . . .	245
	2. Written Reports . . . . .	247
	3. Use for Publicity . . . . .	248
	4. Education of Those who Participate . . . . .	251
	5. Twenty-five Pitfalls . . . . .	252
	6. Religion and Scientific Research . . . . .	254
APPENDIX	. . . . .	256
	Suggested Problems for Experimental Investigation . . . . .	256
	The Construction of a Sample Scale . . . . .	270
	Statistical Tables . . . . .	273
	Glossary . . . . .	279
	Index . . . . .	292

## LIST OF TABLES

TABLE NO.	PAGE
I. How to Pair Subjects for Equated Groups . . . . .	21
II. Suggested Form for Graphic Rating Scale for Initiative . . . .	41
III. Form for the Weighting of Test Elements in Accord with Criterion Scores . . . . .	157
IV. Order Distributions: Showing Also Mean, Median, Mode $Q_1$ , $Q_3$ , $Q$ , and Range . . . . .	177
V. Rank Distributions . . . . .	179
VI. Frequency Distributions: Showing Also Median, $Q_1$ , $Q_3$ , and $Q$ .	180
VII. Graph Distributions . . . . .	182
VIII. A Curve of Normal Distribution . . . . .	183
IX. Average Deviation and Standard Deviation . . . . .	192
X. Use of Standard Deviation in Comparing Scores . . . . .	195
XI. Graph of Relation between Ethical and Religious Scores . . .	198
XII. Rank Method Correlation between Intelligence and Ethical Score . . . . .	200
XIII. Rank Method Correlation between Intelligence and Religious Score . . . . .	201
XIV. Rank Method Correlation between Ethical Score and Religious Score . . . . .	202
XV. Correlation between Intelligence and Ethical Score . . . .	204
XVI. Correlation between Intelligence and Religious Score . . . .	205
XVII. Correlation between Ethical Score and Religious Score . . .	206
XVIII. Example of Correction for Attenuation . . . . .	210
XIX. Prediction of the Most Probable Ethical Score from a Given Intelligence or Religious Score . . . . .	211
XX. Partial Correlation between Ethical Score and Religious Score When Intelligence is Held Constant . . . . .	214
XXI. Summary of Data for Hypothetical Problem in Administration of a Board of Religious Education . . . . .	226
XXII. Intercorrelations of Measures Obtained by Board . . . .	226
XXIII. The Reliability of Certain Measures Previously Computed . .	239
XXIV. The Reliability of the Predictions in Table XIX . . . . .	241
XXV. How the Board of Religious Education Should Spend Each Dollar . . . . .	250
XXVI. Form of Recording the Number of Judges Preferring One Speci- men to Another, for Ten Samples . . . . .	271

# LIST OF TABLES

TABLE NO.		PAGE
XXVII.	Conversion of Per cent of Judges Believing a Given Sample Superior to a Second Sample, into the Distance in S. D. Units between First and Second Samples . . . . .	272
XXVIII.	Squares and Square Roots of Numbers to 100 . . . . .	273
XXIX.	Value of $r$ Corresponding to Each Value of $\rho$ . . . . .	274
XXX.	Predictive Indices for Certain Correlation Coefficients . . . . .	275
XXXI.	Interpretation of Correlation in Terms of Displacement . . . . .	276
XXXII.	Per cent of Independent Causal Factors Common to Two Measures Yielding Certain Correlations . . . . .	277
XXXIII.	Per cent of Total Number of Cases (Area of Normal Curve) Falling above Given Scores in S. D. Units . . . . .	278

## CHAPTER I

### THE SCIENTIFIC APPROACH TO RELIGIOUS EDUCATION

#### 1. MAGIC VERSUS SCIENTIFIC METHOD

Human progress has involved as one of its essential features, the substitution of scientist for soothsayer. In the primitive village sickness was treated with magic ointment and supposedly potent incantations. Messages were sent to the absent by means of favorable spirits, invoked by suitable shibboleths. Ceremonies and spells were accepted without blasphemous questions. An impertinent inquirer possessing the temerity to ask, "Why?" "How do you know?" and, "Will it always work that way?" might expect uncomfortable consequences after the devout had recovered from their astonishment.

In modern civilization sickness is carefully studied, the symptoms diagnosed, and remedies prescribed. Messages are sent over wires, or indeed without wires at all. The scientific student of medicine or radio transmission is not dismayed by those who would ask "Why?" or "How do you know?" or "Will it always work that way?" Underneath all his activity is an assumption that the most carefully controlled experiments with most accurate measurements will yield truth which is wholly reliable. In the control of the physical universe magic has been widely replaced by method. Rivers are spanned, skyscrapers arise, light and power and thought are transmitted with incredible swiftness. Somewhat more slowly the processes of education and of economic life have emerged from the domain of blind, self-satisfied prejudice into the keen cool atmosphere of the scientific search for laws which will always hold true. Equations are being written which will express the future increase in a city's population, the probable consequences of a bankruptcy, the total energy available in a

## 2 EXPERIMENTATION AND MEASUREMENT

human being with a certain diet, the rate at which he can learn to read and write, or the rate at which his mind will change under the application of given stimuli. Achievement has been the reward of open-mindedness, humility, and patience. The will of God in the life of the universe has been found written clearly for those willing to strip aside the placards of tradition and to search reverently for the relationships which constitute reality.

### 2. MAGIC AND METHOD IN RELIGIOUS WORK

Religion is still apt to be frightened by the imperious scientific search for truth. Suppose some impertinent inquirer to have escaped after his experience in questioning the primitive practices of the medicine men. Let him approach a certain Sunday school and ask, "Why do you sing hymns? What evidence have you that those prayers and sermons will make these varied children into what you would like to have them become? Why do you study such complicated material as the Scriptures? What are the actual consequences of your methods? Do they always work? Why not give up the whole business? Would not society be better off?"

Sometimes the answers would be emphatic opinions of workers who have grown up under such methods and believe in them; frequently the answers would point to the fact that the practice is very old or very widespread. Such answers would differ little from those made by the tribesmen. Indeed, it is not at all impossible that the inquirer would find, after the officials recovered from their astonishment, that he was no more welcome in their ceremonies than he had been among the savages.

Today this stronghold of the soothsayer is crumbling before the attack of a scientific spirit in a new generation. Pointed queries as to *whys* and *wherefores* are met with more investigation and experiment, and with less of that emotional reaction which indicates a lack of intellectual defenses. An ever-increasing number of workers in religious education are more anxious to find truth than to confirm their previous opinions or blindly to preserve the traditions of the past. Frequently, of course, they find that race experience has reached the goal before them, that the methods, institutions, and viewpoints which have been un-

critically accepted have had more than a large nucleus of truth. It is only reasonable to expect that any system of thought or action which has commended itself to any large group of people for a considerable time, would have in it elements not lightly to be tossed aside. This expectation is as applicable to Fundamentalists as to Modernists, to Buddhists, Mohammedans, Bolsheviks, Rotarians, Pacifists, Militarists, the Ku Klux Klan, and the Knights of Columbus, as to the particular viewpoints any one group may find most congenial. Therefore, the scientific approach cannot rest satisfied with the assumption that there is something worth while in a practice because some people think so or thought so. The student of the ways of God may not assume in advance of his investigation which of the elements of tradition are true and valuable. Such findings may be his conclusion but a broad skepticism must be his starting point. True, the scientist in the field of religious education is not without assumptions. He assumes that God will reveal himself to those who seek in the manner in which God has chosen to be sought; that God is reliable rather than whimsical; that the most uninviting truth is better than the most heart-warming error; and that only by painstaking exactness can he hope to approximate to the real nature of the complicated relationships with which he deals.

### 3. PROBLEMS WHICH CANNOT BE ANSWERED BY EXPERIMENT

It is ever the fault of those who feel strongly the need for some point of view which has not been stressed by their fellowmen previously, that these pioneers become overzealous in the cause of their new ideas. Scientists have not always spoken to laymen with the same humility which would characterize these scientists in relation to the physical universe. Striplings in the world of ideas are apt to show a jaunty self-confidence which older and better established viewpoints do not need.

It would be unfortunate, but not surprising, if the scientific approach to problems of religious development were to claim more than it can fulfill. Because so much will be said about problems which can be best answered by experimentation it may be well to emphasize the fact that some of the most important

## 4 EXPERIMENTATION AND MEASUREMENT

questions of life cannot be settled in that way. Consider the following problems.

1. Do the people of Tennessee have a moral right to prohibit the teaching, to coming generations, of doctrines which are repugnant to the present generation?
2. What is truth? Reality? Objectivity?
3. What is beautiful?
4. What sort of life did Jesus live? What would he do in such a problem as mine?
5. Is militarism preferable to pacifism?
6. Do all men have an equal right to happiness?
7. Should a few suffer for the good of all?
8. Is western civilization with its rapid progress more satisfactory than the quieter meditations of oriental countries?
9. What is the aim of religious education?

Upon all such questions scientific experimentation may contribute data. None of them is completely answered when the report of the scientist has been made. Indeed in any matter involving human welfare the scientist can state only, "If this is done — these will be the consequences." Whether people choose to take one set of consequences or another rests eventually upon a value judgment or preference which goes beyond science, however much use it may make of scientific processes and findings.

### 4. EXPERIMENTS NEEDED TODAY

With due recognition of limitations, the impression still remains of the overwhelming need for accurate appraisal of the teachings and teaching of religion. In the appendix (pp. 256) are suggested a hundred problems many of which are crying for experimental investigation today. Upon most of them, practice is proceeding blindly. To some of them, such, for example, as the consequences of prayer for the sick, many people would feel the answer established. Yet none of them have had adequate scientific investigation. All of them are questions of fact which depend for answer upon the results of actual experiment, or the gathering up of evidence from the experiments which are constantly being performed in the natural course of living. None of them are capable of reliable solution by armchair or dialectic

methods. No decrees upon any of these questions, whether set forth by curriculum makers, church leaders, professors of religious education, or experienced church school workers are worthy of acceptance, except insofar as they rest upon careful descriptions of what has actually happened. Any statement upon any such problems at the present time should be accompanied by due consideration of the limits within which actual observations have been made and of the crudity of the measures of appraisal.

#### 5. PURE RESEARCH AND PRACTICAL RESEARCH

Faced by such a stupendous task, by questions which may not be answered for generations, the practical worker in religious education may well pause. What is to be done? Are children and young people to be left to their own resources while the army of religious workers turns with one accord to test tube and stop watch and statistical graph? Surely in some way business must be carried on at the old stand, while the more glorious building is being designed.

There are some workers who should be set free, and adequately equipped to carry forward tasks of pure research. They should be encouraged to give their lives to the exact and laborious search for the true answer to a detailed problem. By this method every firm coral island of science has been built. Out of such investigation emerges eventually an unambiguous answer to the chosen question, based upon evidence which any competent mind finds adequate. No restrictions or shackles should be placed upon such workers, save that they shall be rigorously scientific, impartial, thorough, and accurate. They should be protected from influence by men of affairs who are "practical" in a hard and limited fashion. Thorndike has well called the attention of educators to two great men who lived in England during the same generation. One was Thomas Arnold, who gave his life to the significant task of improving actual schoolroom practice. He stands out today as a great example in the quickly neglected past. But Francis Galton, who may never have taught in a single schoolroom, working in the realm of pure research, achieved results which permeate the theory and practice of almost every modern educational system.



## 6 EXPERIMENTATION AND MEASUREMENT

An equally great contribution to the progress of scientific religious education can, and should, be made by the teachers and administrators who are linked up with practical field situations. The greatest present hope for experimental results rests upon them. Their own advance in theory and improvement in technique depends in large degree upon their success as experimenters. It is to be expected that such workers will not launch out on any problem merely because it strikes their fancy. They will carefully analyze the work they are carrying on. They will find the points at which they are failing to achieve maximum results. They will select among possible devices the few which seem most likely to be successful. These will be put into operation. The success of the experiment will necessarily depend not so much upon the securing of a final theoretical answer as upon the practical progress of the program. A solution which is all right in theory but which does not work in practice, is no solution at all from this point of view.

### 6. SOURCES OF PROBLEMS FOR PRACTICAL RESEARCH

Experiments of this second sort are being performed constantly by every alert teacher and administrator. The worship service doesn't seem to go well. A song leader is introduced. Improvement follows. James insists upon disturbing the group. The teacher calls at his home. He appears worse the next week. He is sent to the superintendent. He comes back cowed, but takes no further constructive part in the class work. So it goes. Sometimes with groups and sometimes with individuals experiments are constantly under way. For the most part, however, they are measured only by the uncritical lump judgment of one teacher or administrator. Frequently the factors producing the change are so muddled together in a confused situation, that what seems to be effective in one case appears not to work in another. By careful control of the factors which might influence change, the investigator may discover the "key" to the situation. If he succeeds, he has found an operation which, under given conditions, will always and invariably produce given results. That this is not apt to be the result of the ordinary trial and error progress of the field worker is due in large degree to careless-