1928

Supervision of religious education through objective measurements

Ideal
Double Reversible
Manuscript Cover
PATENTED NOV. 15, 1898
Manufactured by
Adams, Cushing & Foster

28-6½

BOSTON UNIVERSITY

GRADUATE SCHOOL

Thesis

SUPERVISION OF RELIGIOUS EDUCATION
THROUGH OBJECTIVE MEASUREMENTS

Submitted by

Thomas M. Hutchin Cunningham

(A.B. Austin College, Texas, 1921)
(B.D. Austin Seminary, Texas, 1916)
(Th.M. Union Seminary, Virginia, 1926)

In partial fulfilment of requirements for

the degree of Master of Arts.

1928.

eligious education
ries

July 1929

37+91

TABLE OF CONTENTS

AND OUTLINE

page

Introduction - 1

Statement of problem

Purpose of this paper - 3

Limitation of its scope - 3

PART ONE

EDUCATIONAL SUPERVISION AND OBJECTIVE MEASUREMENTS

I. Educational supervision in its general aspects. -

A. Definition of Supervision - 6

1. The need for a clear definition - 6

2. Supervision not to be confused with administration - 6

3. A concise definition of supervision - - 7

4. A functional definition of supervision - 8

B. Principles of Supervision - 9

1. Supervision must be democratic - 9

2. Supervision must be a cooperative undertaking - 10

3. Supervision must distinguish between instruction and administration - 10

4. Supervision must be scientific - 11

5. Supervision must center upon the improvement of instruction - 11

6. Supervision must be a clearly defined program - 12

C. The scientific nature of supervision - 13

1. Science and education - 13

2. The development of the scientific method in education - 13

3. Steps in the scientific method - 14

a. The setting up of Goals - 14

b. Supplying the means for reaching the goals - 16

(1). Improved instruction

(2).Better curriculum

(3). More favorable physical equipment

c. Measuring progress to the goals - 17

(1) Starting point

(2) Amount of progress

(3) Difficulties in the way of progress

II. Objective Measurements in Educational Supervision

A. Definition of Educational Measurements - 20

1. Quantitative definition of education - 20

2. Definition of measurements - 21
3. Definition of educational measurements - 23
B. Methods of measurement in education - 23

1. Personal judgment - 23

2. Comparison - 25

3. Standardized measurement - 25

a. Strength - 25

b. Weakness - 26

C. Problems Demanding more careful measurement - 28

1. Administrative problems - 28

a. Classification and promotion - 28

b. Educational and vocational guidance - 30

c. Evaluation of school efficiency - 30

d. Evaluation of the efficiency of the teachers - 31

(1). Measurement of the products of teaching -32

(2). Teacher Rating - 32

e. Reports to patrons - 33

2. Instructional problems - 33

a. Standards for pupil progress needed - 33

b. Elements of Pupil Progress to be evaluated - 34

c. Diagnosis of pupil progress demanded - 35

d. Remedial measures to be applied - 37

3. Research Problems - 37

a. The evaluation of Teaching methods - 38

b. The evaluation of curriculum materials-38

c. Experimental work in personnel and organizational procedure - 38

D. Criticism of the use of educational measurements - 39

1. Measurements only a fad - 39

2. They place too much emphasis on facts - 39

3. They tend to produce uniformity - 40

4. They contribute to the mechanistic conception of our age - 40

5. Many outcomes of education are too intangible to measure - 40

6. Failure has been noted in their use - 41

E. Limitations to the use of measurements - 41

1. Human beings vary from day to day - 42

2. Measurements are not highly refined - 42

3. Measurement cannot take the place of personal equipment - 42

III. Objective measuring instruments

A. Tests - 45

1. Advantage of objective tests over old type examinations - 46

a. They are more comprehensive - 46

b. They are more carefully constructed - 46

c. They are more objective - 47

d. They provide greater uniformity in testing conditions - 47

e. They consume less time and energy - 47

2. Limitations of objective tests - 48

a. Limitations of standard tests - 48

(1). Cost

(2) Scarcity

(3) Specific needs not met

b. Limitations of unstandardized tests - 50

(1) They cannot take the place of standardized tests
(2) They are only relative in their measure
(3) They are not as objective as standardized tests
(4) They can only supplement

c. Advantages of unstandardized over standardized tests - - 5D

(1) Inexpensive

(2) Adaptable to emergencies of classroom

(3) May be constructed in any number as the teacher finds time

3. Traditional examinations made objective - 51

a. Setting up standards for scoring

b. Assigning relative values to the stand_ ards

c. Scoring.

B. Scales - 52

1. Test scales - measuring - 53

a. Rate of performance -53

b. Difficulty of performance - 53

c. Quality of performance - 53

2. Rating Scales - 54

a. Personal rating - 54

(1) Pupil rating -54

(2) Teacher rating - 55

(a) Personal and social qualities

(b) Cooperative qualities

(c) Leadership qualities

(d) Scientific and professional attitude

(e) Teaching ability

b. Score Cards - 57

(1) They force the investigator to explore every relevant part of process

(2) They are a reminder of relative values.

PART TWO

SUPERVISION OF RELIGIOUS EDUCATION THROUGH OBJECTIVE MEASUREMENTS

Introductory statement - - 60

IV. The scientific nature of supervision of religious education -

A. Definition of Religious Education - 62

1. In terms of religious changes produced - 62

2. In terms of analysis of elements of religious growth - 62

a. Knowledge of the materials of religion - 62

b. Attitudes and appreciation - 62

c. Skills in Christian living - 62

d. Conduct - 62

B. Objective measurements essential to effective supervision - 63

1. Measurement must go with supervision - 63

2. Subjective methods are inefficient - 63

3. Objective measurements facilitate supervisor's work - 63

a. They assist in establishing goals

b. They evaluate the conditions of progress

c. They furnish checks by which progress is made

C. Principles underlying supervision of religious education

1. Voluntary teachers

2. Indifferent public opinion

3 . The general attitude toward the scientific approach to religion.

V. Measurements and the supervision of religious education

A. Special problems - 69

1. Problems growing out of the limitations of present instruments - 69

a. Limited as to number available - 69

(1) Inadequate measure given

(2) Forces us to limited standards

(3) Necessitates undue emphasis on personal opinion .

b. Instruments are crude - 70

(1) Measure lower levels of educational attainment

(2) Measure spiritual traits only indirectly

(3) They secure but partial and inadequate results

2. Problems growing out of the nature of the process measured - 71

a. The intangible nature of spiritual factors - 71

b. The necessity at present for indirect approach - 72

c. Difficulty of securing performance consistent with ability - 72

(1) Difficult to control conditions

(2) Objectivity of scoring more difficult in this field.

B. Measurement of specific abilities - 74

1. Abilities classified - 74

a. The acquisition of content material - 74

(1) Subject matter

(2) Ethical and religious ideas

b. The acquisition of skills - 75

c. The acquisition of dynamic factors - 75

d. The acquisition of character traits - 76

2. Measuring instruments now available in religious education. (A list of 33 tests) - 78

VI. Administering a program of measurements
A. What a program of measurements is - 93
B. Preparation for the program - 94

1. Appreciation of the local attitudes

2. Must have sympathetic attitude toward teaching problems

C. Presenting the Program - 95

1. In the teacher training class - - 95

2. In the teachers' and officers' meetings - 96

3. In the Executive board meeting - 97

D. Active administration of the program - 98

1. Giving the tests

2. Tabulating results

3. Interpreting results and applying remedies.

GENERAL SUMMARY - -102

General Conclusion - -105

I. Supervision is an important task demanding skilled leadership. -105

II. The intelligent use of objective measurements is essential to the task.

INTRODUCTION

INTRODUCTION

Educational supervision has experienced a remarkable growth since the opening of the twentieth century. Functions once included in the office of administrator have now been set off and, together with new functions, placed in the hands of a separate officer called the supervisor. So vital are these functions thought to be that the supervisor is now ranked with the very first in the order of importance in educational leadership.

Along with the growth of the supervisory office, and largely responsible for that growth, has been the increasing recognition of the need for more accurate methods of handling educational problems which come under the supervisor's direction. The effort to meet this need has resulted in bringing into the service of the schools of America a system of objective measurements for facilitating the supervisor's task, and for making it, in many respects more scientific and exact. Haphazard methods based on the personal opinions of those in charge of educational processes have been proven to be inadequate. Dissatisfied leaders in the educational field were ready to welcome the introduction of any device that promised aid to the solution of problems growing out of the older methods of education. Accordingly, when the possibilities of making and using objective instruments for the measurement of educational products were discovered these instruments came into immediate popularity with those who were seeking new methods of approach.

If at first unwise application has been made of educational measurements on the part of enthusiasts, eliciting opposition to them from the more conservative leaders, it is hoped that now their true value is being generally recognized, and that their use, in proper limits, will forward the cause of education in all of its complex phases.

Religious education is likewise beginning to sense a need for a more careful evaluation of the outcome of its instruction. In spite of the impression in some quarters that religious growth cannot be measured, it is thought by those who have made it a study that a careful administration of a program of objective measurements in religious education can be made to contribute real help in the solving of difficult problems involved in the conduct of this work. Profiting by the experience of those who have pioneered in the field of measurements in secular edication, religious leaders are carefully and cautiously taking steps toward the development of a system of measuring instruments which seem to have real potential value in them for the accomplishment of their purposes. It is believed by these leaders that religion can be measured by the proper analysis of its constituent elements, and that a scientific basis can be established in such a procedure for making possible definite progress toward the goals of our American churches.

The purpose of this paper is to review the field of supervision in secular education and the use it makes of objective measurements, and upon this basis to discuss the uses of similar measurements in the supervision of religious

education. Part One is, therefore, a treatment of the former topic, and Part Two attempts to apply, as far as possible, the results of experience in this field to the latter. No attempt is made here to define the exact limits of the two fields, but we shall simply take for granted that religion has its distinct function to perform in the general scheme of education required for well-rounded citizenship. Although the two fields are closely related in both content and method, religious education has its own peculair problems to face over and above those which are common to both realms.

LIMITATIONS:

Two important limitations are imposed on this discussion:

1. It is a development of the subject of supervision in one particular phase of its activity, that phase which makes use of objective measurements for the accomplishment of its purposes. There are other phases of supervision which are not concerned directly with the application of such scientific tools to its work which do not come within the scope of this paper.

2. The primary objective in view is the field of religious education, not that of education in general. No exhaustive treatment of the field outlined in Part One is attempted, but an effort is made to lay a broad basis for the discussion of the comparatively new field of supervision in religious education. So meager is the work that has been actually accomplished in the latter office that it was thought advisable to

give a rather comprehensive view of the topic treated in Part One in order that a sound basis might be assured for the development of Part Two.

PART ONE

EDUCATIONAL SUPERVISION AND OBJECTIVE MEASUREMENTS

EDUCATIONAL SUPERVISION AND OBJECTIVE MEASUREMENTS

EDUCATIONAL SUPERVISION IN ITS GENERAL ASPECTS

A. Definition of Supervision

1. The need for a clear definition

As a separate function in education, supervision is comparatively new. It has arisen to meet the growing demands of a very complex process for a unifying element in that process. Formerly education consisted of a simpler process confined to a small group in one room with a very limited text material designed for simpler ends. As the educative process grew more complex with the growing complexity of life itself, the individual differnces of teachers expressed themselves in varying methods and techniques, until in the larger systems there seemed to be no end of methods in making education effective. The confusion was made all the more serious by the lack of unity of purpose.

Since the inevitable introduction of supervision into education, it has been severely critidised as an intruder on the sacred rights of the individual teacher, and as a duplication of the work of the administrator. This criticism has been due chiefly to the lack of a clear definition of the task of the supervisor, and its acceptance as a vital part of the educational procedure now depends on a definite marking off of its functions.

2. Supervision must not be confused with Administration.

Administration is a broader function and includes supervision. The whole task of school management and control

comes under the authority of the general administrator, while supervision has to do chiefly with the teaching problems of the school. Barr and Burton[1], in The Supervision of Instruction, p. 67, claim that "in the interest of efficiency" a distinction must be drawn, for, "the first step in getting together in cooperation is getting apart in definition." Pitman[2] says, "Administration has to do with setting up the organization, while supervision is the intelligent handling of the organization".

3. A concise definition of supervision.

Supervision is the improvement of teaching with the ultimate aim of improving the pupil. Nutt's two-fold purpose[3] of supervision is given thus: first, the attainment of increased effficiency on the part of the teacher, and second, the efficient education of the child. In all authoritative works on the subject the child is given a central place in the educational program. H.C. Morrison, The Practice of Teaching in Secondary Schools, Part IV. Chapter XXVIII[4], calls this emphasis on the child, "The Integrity of the School". The supervisor existe only for the sake of better teaching, and the teacher exists only for the sake of the child; the whole program of the school is bound up into a unity by the interests of the child. McCall uses a striking figure, "The pupil is the Sun, the teachers are satellites, and supervisors are moons of satellites."[5]

1- Barr, A.S. and Burton - The Supervision of Instruction N.Y. Appleton, 1926 1921
2- Pitman, M.S. The Value of School Supervision-Warwick and York,
3- Nutt,H.W.-The Supervision of Instruction-Houghton Mifflin,1920
4- Morrison,H.C.-Practice of Teaching in Secondary Schools Univ.of Chic.Press,1926
5- McCall, Wm.A. How to Measure in Education - Chap.V

The supreme purpose of supervision is that the education of children may be carried on under better conditions and that children may learn. All minor purposes must contribute to this end or be discarded. The improvement of the teacher merely from the standpoint of personality, or producing a pleasing appearance, in themselves, is not sufficient. Such improvements must be subjected to a rigid evaluation in the light of their contribution to the progress of the pupil. The supervisor should seek to eliminate all personal impressions of a teacher's work or any influence arising from the teacher's standing in the community, and to set up as his criterion of evaluation the results of her work in pupil development. This does not mean that he is to disregard those factors which do not seem to contribute to the pupil's learning process, but only that he is to eliminate them from his evaluation of the teacher's efficiency. There are times when, in the winning of the confidence of the staff, it will often be expedient to give notice to insignificant and irrelevant qualifications, but supervisory duties demand that all such minor factors be weeded out of strict measurement of a teacher's standing, and that judgments be purified of all criteria except that of pupil progress. This policy may lead to the changing, or even of dismissal, of certain teachers who were thought to be quite efficient, but if it results in the advancement of the pupils it is justified.

4. A Functional Definition of Supervision.

A number of working definitions of supervision have been offered. The one given by Barr and Burton[1] is suggestive

1-Barr, A.S. and Burton W.H. Supervision of Instruction, p.1730

of the best thought in this direction. We will mention the following: (1) The selection and organization of materials of instruction which includes items such as the setting up of objectives, experimental testing of materials, testing efficiency of instructional materials, etc; (2) Research and experimental study of the problems of teaching, using tests, record forms, and experimental studies, and perfecting the tools of research; (3) The evaluation of the effectiveness of teaching, by the use of rating cards, survey testing, objective standards and check lists; and (4) The performance of professional and semi-professional duties. Each of these functions reveals a demand for a scientific and impersonal attitude on the part of the supervisor in the efficient conduct of the task they impose.

B. Principles of Supervision

The development of a statement of principles of supervision has found its best expression in the six listed by Barr and Burton[1]:

1. Supervision must be Democratic. The supervisor may realize the ideals of democracy, (a) in providing for, and stimulating, initiative, self-reliance and responsibility on the part of every member of the staff; (b) in recognizing and dealing sympathetically with the problems of human relationships in the effort to improve instruction; and (c) in creating a sense of democratic leadership in a group of co-workers

---

1 -Barr and Burton - Supervision of Instruction, p. 83-85.

The supervisor may in this manner concentrate the efforts of a united body of workers to the end that the pupils of the schools may make the largest growth in desirable ideals, interests, knowledge, powers, and skills, with the least waste of energy, and the greatest amount of satisfaction to all concerned.

2. Supervision must be a Cooperative undertaking of Teachers, principles, and supervisors. This can be accomplished (a) through the recognition of the distinctive contribution of each to the improvement of instruction; (b) through a clear definition of responsibilities; (c) through the realization that not only teachers, but principals and supervisors as well, are learners in the study and observation of classroom problems; (d) through the general encouragement and conservation of contributions from teachers as well as from the heads of departments, principals and supervisors; and (e) through the organization, evaluation and coordination of the efforts of those concerned with the improvement of instruction. Careful observance of the details of these items should create a sense of unity and cooperation together with a growing sense of achievement which results from the power of a coordinate staff to accomplish the ends of education.

3. Supervision must distinguish between Instruction and Administration. This can be done (a) through the clearance of all administrative matters through administrative channels; (b) through the recognition of the principal as responsible for instructional conditions within his building; (c) through the recognition of instructional experts as responsible for

instructional conditions andfor the adequacy of this program in the several fields of instruction. It is important that the twophases of work be defined in clearly expressed statements, even though the two may be combined in one officer, as is so often the case. It is especially demanded where distribution of responsibility is made between a number of individuals on the staff.

4. Supervision must be Scientific. This should find expression through, (a) a common scientific background for principals, teachers and supervisors; (b) the development of definite, well-understood objective standards for judging and improving the quality of instruction (whatever standards are used should be known in advance by the one where teaching is to be observed); (c) an experimental and laboratory study of instructional problems; and (d) interpretive measurements of results. Special emphasis should be placed on the importance of this principle because of its essential relation to the development of the thesis of this paper. It is with the scientific viewpoint and methods in mind that our problem is attacked. In the treatment that follows this principle will constantly be assumed as fundamental to all that is said.

5. Supervision must center upon the improvement of instruction. That is, expert supervision, (a) should lead teachers to a broad vision of teaching problems, to a wide range of experience, so that the work of one grade may be seen in relation to the work of other grades, to an under-

standing of needed revisions, of necessary growth and of final outcomes of instruction; (b) should lead teachers to master the technique of classroom instruction, to establish good teaching habits, and to develop high standards of teaching efficiency; and (c) should lead teachersto persistent study of teaching problems, to experimentation, and to the use of classroom as a pedagogical laboratory in which to revise and improve methods of instruction. In this principle we discover the point of application for the scientific method, the object of chief concern in the supervisor's task - the improvement of teaching.

6. Supervision must be a clearly defined, definitely organized program. This program should be progressive from year to year, and should include, (a) a statement of the objectives of the program; (b) the means for arriving at the established goals; and (c) the checks and standards by which progress is to be measured. The planning of the supervisory program, implied in this principle, is fundamental and essential to success. It makes certain that the supervisor has thought through the more urgent problems of his task and has planned to solve some of them. It acts as a check on his judgment as to the conditions in his school. And finally, it enables the supervisor and his entire staff to formulate clear-cut aims for the future.

We shall constantly rely on those principles which have special relation to scientific methods in education as we proceed with this paper.

C. The Scientific Nature of Supervision

1. Science and Education

In both the definitions and the principles given above, we can readily see the scientific nature of supervision. Indeed, one of the principles given by Barr and Burton explicitely demands that supervision must be scientific, which means that it must approach its work with the scientific attitude of mind, carry on its duties with scientific care, and produce accurate scientific results. In an age of science when not only physical and material objects but also biological, psychological, and even sociological problems are being subjected to the analytical and factual tests of science, it is not strange that education should also be included. Indeed, in the interest of progress in education, nothing more wholesome or helpful could come to its aid.

2. The development of the scientific method in education.

The development of the scientific method has been an age-long process from the method of primitive man to our own day. In the main it has been by three steps, namely, Authority, Speculation, and Experimentation. Truth was first sought on the basis of authority alone; later the speculative spirit broke with authority and found truth elsewhere; and finally came the more exact method of experimentation which has busied itself with the task of verifying the truth authority and speculation laid claim to. All three processes still have their place in the modern world, but it has been found that authority and speculation forfeit their place when

they fail to base their judgments on experimental foundations. Education has had its stages of authority, speculation and experimentation, in keeping with the experience of the race. As in the more practical affairs of life, the world is growing impatient with "rainbow generalities" in education which have no factual basis for their claims, and education on its part is beginning to recognize this just demand for making the educational product a demonstrable, objective reality. In keeping with the scientific spirit of the times, it is attempting to discover the facts, not personal opinions alone; to eliminate the bias that always comes with personal judgment; to make use of exact methods in analyzing data and to establish procedure on a factual basis; to find solutions for problems that have been accurately determined; and to apply remedies that have been scientifically verified to have remedial values. Despite the fact that some educators fail to discipline themselves in the scientific method, there is every evidence of a steady tendency toward the scientific viewpoint.

3. Steps in the Scientific Method.

The scientific supervision of education proceeds in general by three definite steps:

a. The setting up of goals and objectives.

The educational process assumes that progress is desired toward some goal of educational achievement. Educational goals may be either general or specific. Two types of supervision may be used to illustrate the difference in the way goals are conceived. The one will have in

mind certain great and general values, such as truth, love, honor, self-sacrifice and the like, and he will be satisfied in contemplation of their synthetic values. The other demands an analysis of these virtues into specific elements so that he may evaluate each specific step in the attainment of the larger goal. The scientific viewpoint demands no less a synthetic evaluation of the greater goals of life, else the specific goals would be erroneous. It admits that philosophy must first produce the values out of which human goals are formulated. But once these values and general goals are recognized, it is the part of the scientist to break them up into smaller goals which are attainable in the more immediate future. So it is that the scientific supervisor strives to take the general goal, or goals, which philosophy has set up and divide them up into many constituent parts which are capable of systematic and definite attainment in an orderly fashion. The appreciation of a finished work of art is one thing; the mastering of the specific skills necessary for producing that piece of art is entirely another, and it is the mastering of the specific skills of human living that condition the attainment of the greater goals of life.

Not only will the scientific supervisor break up the goals, that he may see more clearly how they may be attained, but he will make them so specific and elemental that the child himself will readily grasp their meaning and see how to attain them. Otherwise, the goals are intangible and lack the vividness necessary for eliciting the child's

own initiative and interested participation. The supervisor will see that goals are set up in this painstaking and definite way for each class, and through the teacher, for each individual child. In doing this the basis for the maximum progress will be laid, and also the condition upon which progress may be measured will be established.

b. Supplying the means for reaching these goals.

The general conditions necessary for arriving at the educational goals may be classified under one or the other of the following:

(1). Improved instruction. The teacher is the most important factor in pupil learning. She comes into direct contact with the pupil daily in the most influential manner, and to a very large extent holds within her power the educational destiny of the child. Whatever else may enter into help or hinder the child reaching his goals, the teacher, by her efficiency, or lack of it, decides the crucial problem.

(2). Curriculum. The curriculum, in all its many phases, is an important conditioning factor in the attainment of specific goals of pupil learning. The supervisor will see that a well planned course of study and teaching materials is in the hands of every teacher, and that she knows how to make that course most effective in enabling the pupil to reach his objectives.

(3). Physical equipment. The location of the school, the environment, playgrounds, building, interior conditions, beauty, ventilation, lighting, and may other minor factors,

are receiving the attention of supervisors and educators because they are found to be important in the creation of favorable learning conditions.

Each one of these factors of pupil progress must be brought under careful and constant scrutiny, and tested as to its effectiveness in contributing its part to the progress of the pupil. Constant changes will be made in those parts which show lack of efficiency to this end.

c. Measuring the progress to the goals.

The third step in the scientific supervision of education is that of measuring the progress made toward the goals set up. In order to do this, it is necessary to have a starting point, (i.e. educational status at the beginning of the year), to measure the amount of progress from this point toward the goals, and to discover the difficulties in the way of attaining the goals.

(1). Starting point. By methods that will be mentioned later, a child's educational status may be determined which will both enable him to make greatest progress and render his progress measurable. The child's educational status, or starting point, at the beginning of the school year, is an important item in measuring his progress. This involves individual differences which arise out of heredity, environment, the attitudes of companions and many other lesser factors which have affected his life.

(2). Amount of progress. Techniques for measuring the amount of progress the child makes toward the goals of education will be discussed under the heading of Measuring In-

struments. With the use of these instruments, and given our goal and a starting point, the measurement of progress is made possible.

(3). The discovery of difficulties in the way of progress. This will be discussed more fully in connection with diagnosis and remedial treatment in a later section, so we merely mention it here as a logical step in the measuring of progress.

In summing up the scientific method, it might be stated that it is an effort to objectify education and to heighten its efficiency in the promoting of pupil progress.

CHAPTER SUMMARY

The general nature of educational supervision has been discussed in the following three points:

A. Definition of Supervision, especially as centering upon the improvement of teaching and pupil progress.

B. Principles of Supervision, emphasizing the principles which have to do with the scientific nature of supervision.

C. The Scientific Nature of Supervision, which is that phase of the subject which concerns the discussions of this paper most vitally.

OBJECTIVE MEASUREMENTS IN EDUCATIONAL SUPERVISION

OBJECTIVE MEASUREMENTS IN EDUCATIONAL SUPERVISION

In a program of scientific supervision, it is obvious, in the light of the discussion of the preceding chapter, that whatever contributes to the accuracy of educational procedure should be of great value. Chapters II and III deal with the subject of educational measurements as they contribute to this end. Standardized measuring instruments, rightly used, reveal to the supervisor many of the weak places in the educational system, and suggest ways of remedying them. Their continued application keeps him informed as to the efficacy of changes being tried out, and enables him to adjust and readjust procedures till a more effective method has been discovered. In our discussion of measurements in education, we shall first seek to define our terms.

A. Definition of Educational Measurements

1. A quantitative definition of education.

Before we may define educational measurements in understandable terms, it is necessary to first define education in quantitative terms. Thorndike[1] has given us just this sort of definition in the words, "Whatever exists at all, exists in some amount. To know it thoroughly involves knowing its quantity as well as its quality. Education is concerned with changes in human beings". This qualitative thing called education, then, is capable of be-

---

1. Quoted in Barr and Burton, Supervision of Instruction, p.99-100

ing defined in quantitative terms, and it is this quantitative definition which makes all the following discussion intelligible. "Changes" are measurable more or less objectively provided we are able to watch those changes with reference to a scale of definite units describing them.

3. A Definition of Measurements.

Measurement is a more exact description of an object, in all its particulars, in terms of previously known factors. Physical objects are measured by describing their qualities, such as length, weight, solidity, etc. in terms of inches, pounds, resistence, etc. The terms here used are arbitrary, but they are objective and sufficiently well-known to have meaning when used by the average man. We say that an object that has been measured in such previously known terms has been described more exactly than if only a general impression is obtained. General impressions are subjective and unreliable, especially when the object measured is to be compared with other objects, or if, for instance, a physical object is to be built into a structure requiring exact fittings with other objects of like nature. A more exact description of each detail of a building must be secured before those details will fulfill the requirements of the blue-print. The blue-print itself has been accurately measured and described in every detail in terms of the same previously known factors. It is only thus that scientific exactness of every part can be assured. It should be remembered, however, that slight errors enter into the measurement of the most carefully calculated structures. Allowance must be made for heat expansion and

other variables affecting the objectivity of physical measurements. Nevertheless, by means of the more refined instruments in the hands of skillful men, physical objects have been more exactly measured and greater accuracy has resulted than when these objects were handled by mere personal observation.

3. Definition of Educational Measurements.

By taking the definition of measurement, above, and applying it to the quantitative definition of education we find that educational measurement is the description of changes in the educational product in terms of previously known factors. As in the case of physical objects, educational changes are more accurately described by the use of refined and objective measuring instruments which eliminate, as far as possible, the personal equation, and give results that are virtually the same in the hands of any number of skilled persons. The units of measurement in education are not so well known to the average man as the units used in measuring physical objects; for that reason it requires more skill in handling them, and their use is therefore limited to the few who know how to apply them accurately. Moreover, the very nature of the educational product makes exact measurement more difficult.[1] The variable factors of human nature are much more numerous and of greater magnitude than those of physical things, and much greater allowance must be made for factors affecting the variability of education. But again, it is found that, even in education with all its

---

1 - Barr and Burton - Supervision of Instruction, Chap. IX

elusiveness and variability, refined and objective measuring instruments are being perfected which are eliminating much of the haphazard "guessing" that has heretofore characterized the evaluation of the educational product. Educational changes are being observed with much more care and precision in terms of definite units on objective scales, which, when used by different skilled persons, secure similar results.

B. Methods of Measurement in Education

Measurement of some sort has always been used in education. It is only recently that educational measurements have been sufficiently refined to render them objective and comparable with scores between schools and systems. Three methods have been followed in the evaluation of education. To some degree, all three methods are still in use, but the more refined method is gradually replacing the older ones with the result that the supervisor's task is becoming more and more and exact and reliable.

1. Personal Judgment.[1]

In the past, personal bias has played a large part in the scoring of pupil progress. The average teacher has been found to have a very unreliable sense of value when left to her own unguided opinion of a child's work. She is likely to be influenced by any number of irrelevant factors which unduly affect the score she assigns to a particular paper. Favor-

---

1 - Monroe, DeVoss and Kelly, Educational Tests and Measurements. Introduction by Cubberly.

itism unconsciously enters in to raise the score of one, and personal dislikes will lower that of another, no matter how carefully the teacher guards against these factors. A striking illustration of the unreliability of the personal judgment of teachers is found in the investigations made by Starch[1], who found that teachers' marks on the same piece of student work will vary from 35 to 65 points on the percentage scale. Not only do they differ in grading in one subject of school work, but their differences are equally wide in all the school subjects - Mathematics as in English, for instance. It is easy to see the unfairness of the method of marking which allows for so much personal bias. It is also clear that such judgments do not furnish reliable information upon which to proceed in the future treatment of the pupil.

When one adds to this personal bias the unreliability of the instruments that have traditionally been used he can get some idea of the need for more refined instruments of measurement by which to evaluate educational changes. The very nature of the traditional essay type of examination[2], with its lack of definite purpose, its poorly chosen content, its inequality of question values, its lack of objectivity in scoring, its uncontrolled conditions of pupil performance, its susceptibility to irrelevant factors, and its lack of standards for interpreting results, makes it imperative that a more objective technique be devised for both the teacher and the supervisor of education.

1- Starch, Daniel - Educational Measurements, Chap. II.
2- Barr and Burton- Supervision of Instruction, Chap. IX. p. 396

2. Comparison.

Another method of measurement is that of comparison, the basis of which is "common practice". Many men have attempted the solution of similar educational problems,and their methods of solution are capable of comparison with each other to a certain degree so that one school is found to be better than another, or one pupil is placed on a lower level of achievement than another. But at best this method gives us only relative measures and not absolute standards. It offers no goals which are based on objective standards; it has no way of telling us just how much has been accomplished educationally on a scale of standard efficiency; it is simply a relative standing based on what busy supervisors have thought to be useful, rather than an asbolute standard based on norms of pupil achievement. Obviously, this method of measurement will not suffice for a scientifically administered program of supervision, except where we lack a better method.

3. Standard Measurements.

Standardized measurements are so constructed and administered that the results will eliminate, to a large extent, the defects noticed in the two above mentioned methods of measurement. The strength and weakness of standardized measurements may be discussed as follows:

a. Strength[1]. This is seen in the following facts; (1) the educational product to be measured is carefully

---

1 - Barr and Burton - Supervision of Instruction, Chap. IX.

defined, and the measuring instrument accurately designed to measure just that product; (2) they break up the educational product into many definite and weighted units, each of which is more or less precisely measurable, and lend themselves to objective evaluation; (3) ambiguities of statement are carefully removed; (4) they are tested out with unselected groups or systems in order to discover further difficulties in their structure, and are corrected on the basis of these discoveries; (5) they are submitted to a large number of competent judges who give their verdict as to the objectivity and validity of the instruments for the specific purpose in view. (Again they are revised in the light of these criticisms); (6) finally, they are submitted to statistical treatment by which process they are standardized and ready for editing and use by the supervisor or teacher.

The results obtained from the use of standardized measurements are directly comparable with those of any other like use of them in any part of the country. Norms of pupil achievement enable the supervisor to place any given pupil in relation to those norms, and intelligent educational procedure may be followed out for the maximum progress of each pupil on the basis of this classfication and careful placing of the pupil.

b. Weakness[1].

It should be remembered, however, that the most objective instruments yet devised have not succeeded in

---

1 - Hanson, W.L. "Supervision of Religious Education through Objective Tests and Measurements", B.U. Bulletin, July 10, 1934. p. 4

removing all the personal factors which entered into the results of old methods. It was noticed that this was true, to a small extent, of instruments for measuring physical things. Any two persons measuring the dimensions of a house by the use of a standard foot rule will get slightly varying results. It would hardly be expected that complete objectivity could be achieved by the use of measurements in the more variable factors of education. Two reasons are assigned for the lack of objectivity in the measurement of education:

(1). Lack of refinement in the instruments themselves.

In comparison with the instruments used in the physical world, educational measurements necessarily fall short in respect to their refinement and exactness of measuring units. Experts are constantly refining these instruments, and as a result more objectivity and accuracy can be expected as time passes. However, it is doubtful if educational measurements may ever be made so exact as to eliminate altogether the personal element in their use.

(2). The product measured is, and always will be, variable.

Reference was made to the variability of the educational product under the definition of educational measurements above. We may simply add here that human nature being such as it is, instruments for measuring human changes must always lack objectivity to the degree that human beings are not the same from one day to the next,

or even from one moment to another.

However, admitting these weaknesses, objective and standardized measurements are coming more and more into popular use with educators for the reason that they actually do increase the efficiency of the educative process. They are indispensible to the scientific administration of a program of supervision; and both supervisors and teachers are rapidly sensing their value within limitations, and taking advantage of their aid.

C. School Problems Demanding more Careful Measurement for their Solution.

It will be helpful, in the consideration of measurements, to list the problems which arise in the process of education which seem to demand a more careful and objective use of measurements. These problems may be classified under three main captions in accordance with the three outstanding educational activities.[1]

1. Administrative and Supervisory Problems.

a. Classification and Promotion.

The administrator is annually, or semi-annually, faced with the difficulty of properly classifying and promoting the pupils in hisschool system. The old method of measuring pupil mentality and achievement, basic to proper classification, often led to blunders and misplacements which were never wholly overcome in future educational treat-

---

1 - Monroe, W.S. Theory of Educational Measurement. Chap.III N.Y. Houghton Mifflin, 1923.

ment, no matter how effective the teaching might have been. Two types of measurement are used in the classification and promotion of pupils. Mental measurement will fairly accurately determine the child's ability to achieve, though it will not give any measure of what he will do. Authorities agree, that the intelligence test is not a pedagogical guide, though it locates the pupil in a general way. School achievement is affected in marked degree by factors other than intelligence.[1] In order to get a fair and unbiased picture of the pupil for classification, it is necessary to supplement mental measurement with achievement measurement. Indeed, Kelley,[2] in his statistical treatment of achievement tests, shows that an accurate basis for pupil placement may be obtained by standard achievement tests alone. Whatever may be the method of classification, it is plain that the problem calls for the use of some sort of objective measurement for efficient solution.

Frequently in a class of pupils, a spread of mental ability of five or six years has been found. Such a condition, when revealed, presents a serious problem impossible of solution without the use of measurements that will locate each child, as nearly as may be possible, under the circumstances, where he belongs.

Regarding the non-promotion of pupils, Judd[3] discovered that the problem shows a lack of organization and

---

1 -Barr and Burton - Supervision of Instruction - p. 302 f.
2 -Kelley, Truman L.- Statistical Method, p. 214f
3 -Judd, Charles H.- Measuring the Work of Public Schools, Chap.I.

unity within the system, and upon this lack he bases his contention for a "new type of scientific supervision". Careful measuring of the situation will enable the supervisor or the administrator to present facts to the teachers, and to offer a plan for adopting uniform standards of promotion. Classification and promotion of pupils have a direct relation to each other. The lack of proper promotion at the close of the school term vitally affects the proper classification of the pupils at the beginning of the next term. In like manner, a careful classification of pupils at the beginning facilitates the normal processes of education on the basis of which promotions are made.

b. Educational and Vocational Guidance.

The multiplication of bureaus for the guidance of pupils both in their immediate study needs and in the choice of a life vocation is requiring a special technique in this field, and the administrators are seeking the aid of objective measurements in making this guidance of the most benefit to those who need it. The pupils cannot be left to chance guidance and be expected to find their way to their best course in life. Vocational courses are becoming numerous in the public school curriculum, and merely to allow the child to elect a certain course because it appeals to his immediate fancy would be to load these courses with as much potential harm as good.

c. Evaluation of School Efficiency.

A general survey of the school system to determine its efficiency, again, requires exact methods. It is not

necessary to enumerate all the points in a survey of a school in which measurements are required. It is enough to note that the administrator, or the supervisor acting in administrative functions, must have an accurate picture of the school based on accurate measurements, accurate records, and scientifically interpreted reports in order to proceed wisely in the building up of school efficiency. Exact data as to pupil progress in each unit of the system, accurate diagnosis of the weak points, whether it be the teaching, or curriculum, or something outside of the school, is demanded. Otherwise chance remedies will result and future policies will be rather blindly made and carried out.

d. The Evaluation of the Efficiency of Teachers

An estimate of the efficiency of the teachers in a school system is required for several reasons. The administrator must have a basis for the payment of salaries, for the promotion of teachers, for the changing of a teacher from one school to another, and for the removal of a teacher altogether when necessary. Objective measurements furnish data of an impersonal nature which makes the conduct of administration at this point much more satisfactory and accurate.

The evaluation of teacher efficiency is also of great concern to the supervisor, and together with the administrator, in case the two are separate, the problem may be faced cooperatively. The supervisor's object may be slightly different from that of the administrator, in that his aim is to discover the best teachers and keep them on the staff, and to improve the teaching practice of all the

teachers. Two ways of evaluating the effficiency of teachers have been in use.

(1). The measurement of the products of teaching.

Objective tests used in the measurement of pupil progress over a period of time under conditions which are comparable to those of other teachers and their pupils will give results which enable school officials to rate a teacher as to her efficiency. Teachers may not be compared on this basis unless the pupil groups (a) are of approximately the same mental ability; (b) are approximately on the same level of achievement; (c) are being taught similar curriculum; and (d) are thoroughly measured for discovering significant outcomes.

(2). Teacher Rating.

Here we are concerned, not so much with the product of teaching, as with the teacher herself. The teacher's personal equipment, qualifications, teaching methods, etc. are rated by comparing to a rating scale which analyzes these traits for more careful and detailed observation. The theory underlying the rating of a teacher is that the characteristics rated, and every detailed element in the analysis of those characteristics, makes some contribution to the efficiency of the teacher in securing pupil progress, provided those elements and traits are found in proper balance. Teacher rating scales assign due weight to each element found in the analysis of personal equipment, and a certain amount of objectivity is secured by directing attention singly to each item listed. They are found to be valuable as a device

for stimulating self-improvement on the part of the teacher.

e. Reports to Patrons.

This is a very acute problem involving the possibility of hurting the feelings of both pupils and parents. On the basis of traditional methods of evaluating the pupil's work reports were often unfair, and at times merited the criticism they received. The introduction of objective measurement into the work of the school not only enables school officials to correct these mistakes, but also to appeal to the improved system of measurements for verification and vindication of their reports in case of misunderstanding.

2. Instructional Problems Requiring Objective Measurements.

To the problems of classroom instruction and of pupil progress, objective measurements make their most important contribution.

a. Standards for Pupil Progress Needed.

Standardized measurements furnish the following helps to the solution of this problem, (1) they establish norms of achievement; (2) they assist in setting up objectives for pupil achievement; (3) they furnish checks by which the progress of the pupil is measured; and (4) they enable the teacher to guide the pupil in study procedures, both in the formulation of study directions and in the use of practice tests.

By "norms of achievement" is meant the average achievement of a large number of unselected children of a given age on a given test. By comparing the score of any pupil (or class), with this average, or normal, score his educational

standing may be estimated in objective terms. Superintendent Brooks[1] asserts that the haphazard opinion of teachers as to what the pupil "ought to know" is extremely unreliable as a basis for judging the educational changes that have been procuded in the pupil.

In the formulation of study directions, measurements are useful in analyzing and objectifying the chief points to be mastered in a given subject matter. Used in this manner they stimulate interest on the part of the pupil. By "Practice tests" is meant tests which are designed for individual use to enable the pupil to improve some skill he has been found to lack. Individuals are excused from further practice in these tests when they demonstrate their ability to accomplish normal work. This saves the time of the teacher enables the pupil to make maximum progress, and simplifies instruction in the classroom.

b. Elements of Pupil Progress to be Evaluated.

Standardized measurements help in the solution of this problem in (1) measuring pupil's knowledge of subject matter; (2) evaluating pupil's ability to think in terms of materials of instruction; and (3) in analyzing and measuring the pupil's interests and motives. Because of the comparative ease of construction and use of tests and measurements for pupil knowledge, the type of instrument which measures pupil knowledge has been most numerous. Ascending into the

---

1 -Brooks, S.S. - Improving Schools by Standardized Tests. Chapt. VI.

realm of the more elusive traits dealing with the reasoning powers and with the emotions of the pupil, the problem has been much more difficult of solution. However, experiments are proving that nothing is too elusive to lend itself to a more or less objective evaluation when the conditions are properly controlled.

c. Diagnosis of Pupil Difficulties Demanded.

To be able to diagnose and remove pupil difficulties is of prime importance to those interested in the progress of the pupil. The results of diagnosis will affect both general instruction and individual treatment, and upon these results will depend the formulation of special instructional policies for the future. There are several conditioning factors in diagnosis which objective measurements are called upon to differentiate and control in order that diagnosis may be wisely made. These are:

(1) One grade differs from another in difficulty of diagnosis. Children too young to read and write naturally cannot follow out instructions in written tests, and are therefore very difficult of objective diagnosis. Beginning with Grade III, however, there is an increase in difficulty as we go upward toward the higher grades, reaching the greatest difficulty in the High School classes. The multiplication of interests, the complexity of conditions, and the large number of contacts makes the process a much more difficult one here. Measurements however, are assisting materially in the diagnosis of difficulties in all ages.

(2). The content of the materials of instruction calls for varying methods of diagnosis. Tool subjects are comparatively more easy to diagnose than are the content subjects. Without objective measurements the tendency is to take the line of least resistence and teach those subjects which are easiest to the neglect of the more complex subjects, resulting in peaks and depressions in the graph card. Brooks[1] insists there must be a levelling up process, and objective measurements help both in revealing the unevenness of the learning process and in suggesting remedies.

(3). Teachers differ in their emphasis. Not all teachers emphasize the same subjects, and this imposes the task of getting behind the actual school marks and discovering the teacher emphasis in instruction. Teachers have their hobbies though they may strive to give due attention to all phases of the curriculum. Objective measurements will reveal the unevenness of emphasis and enable the teacher to redirect her attention to the coordination of her work with the aim of the school as a whole.

(4). Individual pupils differ in ability. Again measurements are called on to analyze the differences. The score of a pupil taken alone means very little unless seen in the light of his background out of which arise those factors which make his chances for learning altogether different from those of another pupil. Individual differences arise out of two general sources, heredity and environment. Heredity was formerly con-

---

1 -Brooks, S.S. - Improving Schools by Standardized Tests. Chap. V.

sidered the only source of individual difference, but increased accuracy has demanded an evaluation of all the factors entering into this problem. Environment has been found to have an equal, if not a greater, influence on the differences of children. A detailed analysis of home conditions, neighborhood conditions, play-mates, and school environment must be taken into account and accurately measured before the educational achievement of a pupil can be justly estimated. Growing out of these conditions are the factors of interest, effort, training, maturity, health, and intelligence, all of which should be measured and re-measured many times in order to evaluate them for pupil diagnosis.

d. Remedial Measures to be Applied.

After pupil difficulties have been dissected and carefully diagnosed it remains to apply remedial treatment. Perhaps the simple guidance of the pupil in his study habits is the thing needed to set him right. Other remedies varying all the way from correcting some physical defect to the treatment of mental disorders make their demands on scientific measurements to point the way to more accurate treatment.

3. Research Problems.

In research work, we are interested in determining school procedure through the measurement of outcomes under controlled conditions. This work requires the application of carefully tested methods in the hands of highly skilled persons, and exact measuring instruments are demanded. The following research objectives call for the assistance of objective measurements in their realization:

a. The evaluation of teaching methods. The outcomes of two teaching methods in any given subject will, other things being equal, determine the method to be used in school practice in connection with the same subject. A large number of teachers are given relatively uniform conditions, allowing for only one variable, and that the method under investigation, and thus securing results which determine the value of that method.

b. The evaluation of curriculum materials. Curriculum materials vary in their susceptibility to the application of objective measurements. Unwritten materials used in the problem-project method, for instance, are nearly always of a nature to elude the measuring rod. Life situations and their solutions are so varied and complex that any attempt at measuring them would involve an enormous outlay of time and money. This is only one example of the type of materials found difficult to measure. However, in the field of written materials designed for intellectual mastery, we find objective measurements to have immediate and definite values.

c. Experimental work in Personnel and Organizational Procedure. Research laboratories have for many years been utilized by business institutions for the proper selection and placing of leaders. In the educational field where personal relationships play such an important part in the smoothness of organizational management this is a major problem. To locate each leader where he will be able to do the most satisfying work, and where the personal relations are the most congenial, is no small task. It is important enough to

justify the utmost care and insight based on laboratory experiments of the most refined nature available.

D. Criticisms of the Use of Objective Measurements in Education.[1]

The introduction of objective measurements into the work of the school has given rise to criticism on the part of those not well informed as to their uses. Such criticism has too often been just. The unwise use made of measuring instruments by enthusiasts merits the opposition of conservative leaders, whether in the past when these instruments were new, or in the future when their use becomes more widespread. Some of the objections raised are:

1. Measurements are only a fad. Those who make this criticism are not aware of the basic demand for more accurate investigation into the various phases of educational procedures. It very likely arises out of a lack of understanding of the nature and values of the measuring instruments themselves. As long as education calls for economy, efficiency, and definite results commensurate with the native abilities of children, there will be required some objective way of evaluating the process of education in all its phases.

2. They place too much emphasis on facts and not enough on ideas and judgments. It is true that instruments which measure mere facts have far outnumbered those which evaluated the ideational elements of knowledge. But the latter are

---

1 - McCall, Wm.A.-How to Measure in Education. Chap. I.

becoming more numerous as the research bureaus are able to produce them. On the other hand, there is a basic value in determining the factual content of pupil knowledge, for facts are the raw material out of which ideas and judgments grow.

3. They tend to produce a "deadly uniformity" in educational results. Such a criticism, again, ignores the nature of standardized measurements. Standardization applies to the measurements and not to the child. Instead of producing standardized character they are themselves made standard by the normal abilities and needs of the children. Actual use of these measurements have been found to arouse personal interest and to inspire individual initiative in the child, and where such results are brought out there need be no fear of their producing uniformity of results.

4. They contribute to the mechanistic conceptions of our age. The implication here is that they give undue attention to the scientific accuracy of an educational product to the neglect of the appreciation of more subtle values of life. When measurements are used as an end in themselves this is a real danger. But like the mechanic who takes the machine apart and then puts it all together again to serve more largely than before, the scientific educator must also be able to put the parts back into a synthetic whole to effect a greater value than before. Measurements are only a tool for effecting greater values in the child than would otherwise be possible.

5. Many of the outcomes of education are so intangible as to make it impossible ever to measure them. This may be

true. Still, it has been found that the indirect approach to these "intangible factors" is an effective method of observing their processes.[1] Some degree of objectivity is attained in this method even with those qualities which seem to elude the scientists tools. For instance, a person's vocabulary in a given subject has been found to indicate fairly his ability to think or reason in that subject.

6. Failure has been noted in the use of measurements. Over-zealous experimenters, indeed, have failed in certain outstanding cases because they have neglected to consider all the conditions essential to the successful use of measurements. For a number of years after the introduction of these instruments into the educational field, there was a tendency on the part of many to hail them as the solution for every important problem, and too often hasty and immature use was made of them. The application of physical measurements in this hasty fashion likewise results in failure. It must be emphasized that objective measurements must be valued for what they are as tools to be used with extreme caution by competent persons. They are not ends in themselves. A carpenter's saw is useless as shelter from the storms, but it is a very valuable instrument to aid in constructing a house which will shelter.

E. Limitations to the Use of Objective Measurements

The criticisms given above may be successfully

---

1 - Wood, Benn D. Measurement in Higher Education

answered. Nevertheless, there remain certain limitations to the usefulness of objective measurements which may not be so easily met. Because measurements have been found to effect useful results we have no warrant for mailing them as a substitute for all the other valuable means of obtaining educational results. Their limitations may be stated thus:

1. Human beings vary from one day to another, and a measurement taken today will be different from one taken yesterday. However, this limitation may be partially overcome by carefully controlled conditions, repeated measurements, and the use of various types of measuring instruments Still, it remains a basic fact that human variations in emotional attitudes and tides of interests puts a real limitation upon educational measurements which is not so evident in measurements used in the physical realm.

2. The measuring instruments themselves have not been highly refined. Especially is this true in the upper levels of human achievement. The units on the scales of our present measuring instruments are, as yet, far from accurate and uniform in their discrimination of human performance. They may, perhaps, never approach the fine discriminations which are accomplished by the instruments of science in the physical world, but with all their present crudity, they are nevertheless proving useful in the evaluation of the results of education today.

3. Measurements cannot take the place of skill, enthu-

siasm, and the other personality traits and qualifications of a good teacher. Indeed, measurements are subservient to these, to be used as tools to supplement the basic powers of teaching. Their results must always be supplemented by the personal and pedagogical judgment of the teacher whose trained mind catches the more subtle factors involved.

SUMMARY

The discussions of this chapter are organized around the following points:

A. Definition of educational measurements

B. The Methods of Measurements in Education

C. School Problems Calling for more careful Measurement

D. Criticisms of Objective Measurements

E. Limitations to the Use of Measurements.

CHAPTER III.

OBJECTIVE MEASURING INSTRUMENTS

OBJECTIVE MEASURING INSTRUMENTS

In this chapter, a brief outline treatment of the nature, advantages and disadvantages of the different types of measuring instruments will be given. Strictly speaking, there are only two types of measuring instruments in use for educational purposes. These are (1) Tests, used in measuring the intelligence and achievement of the pupil; and (2) Scales, generally used for describing the speed, difficulty, or quality of pupil performance in terms of definite units, or for evaluating an educational process such as a recitation, or to examine the efficiency of a text book or school building.

A. Tests

Tests are classified as standardized and unstandardized. There are a number of points of similarity between these two classes such as comprehensiveness of subject tested, the care essential to the construction of the test, objectivity of scoring, ease of administration, uniformity of testing conditions provided, and saving of time and energy for the more important task of instruction. The chief differences between standardized and unstandardized tests lie in such matters as the purpose for which constructed and the statistical treatment to which the former are subjected, by which treatment they are standardized. The uses made of these tests have already been discussed in Chapter II, and will not be treated here.

1. Advantages of Objective Tests over Old-type Examinations[1]

Certain advantages are to be seen in the use of objective tests over the traditional type of examination;

a. They are more comprehensive. Every important phase of the subject matter is included by means of a large number of short exercises, whereas, in the old type examination with its essay form of response only a very few questions can be handled, limiting materially the range of material covered.

b. They are more carefully constructed. Rigid care is exercised in the selection of the test exercises, in the statement of the questions, in the determination of the variables to be measured, in the control of conditions of testing, insofar as the structure of the test may control conditions, and in the formulation and adherence to a specific purpose throughout the test. Especially is the force of this advantage seen in the construction of standardized tests which are submitted to competent judges for evaluation, thoroughly tested out in unselected groups, submitted to statistical treatment, norms discovered for age-groups and grades, and the matter of reliability and validity of the tests carefully estimated. The essay type of examination usually requires no more time for its construction than it takes the teacher to thoughtfully write it out - a matter, usually, of only an hour or two. The personal equation enters into the selection of questions in a marked degree when such a haphazard method is used.

---

1 - Barr and Burton - Supervision of Instruction, Chap. IX

c. They are more objective. Test exercises are so constructed and arranged on the test sheet as that the marks of the pupil are quickly seen and scored as either right or wrong. That is, the question being one of fact, there is only one answer possible; hence any two teachers will agree in scoring the papers. All irrelevant factors are ruled out, and the specific ability being measured is objectively revealed. The traditional examination, as has already been pointed out in Chapter II, admits of a wide difference in scoring by different teachers, due to the personal judgment exercised in evaluating pupils' answers.

d. They provide for greater uniformity in testing conditions. Tests strive to eliminate the personal influence of the teacher in getting pupil response, they attempt to control other conditions such as time, seating, equipment, and in every way to make the response as near normal as it is possible to do so. Furthermore, in the case of standardized tests, the accompanying instructions attempt to control these conditions by making them uniform wherever the test is given. This is in order that results obtained may be compared, and the standing of a pupil or class may be accurately estimated in relation to norms which have been established on the basis of each test.

e. They consume less time and energy. On the part of the pupil, little effort is demanded in the mechanical responses such as writing, etc. so that his attention may be given to the thought necessary for marking the right answers. In this way a large number of exercises may be covered in a

small amount of time, and with relative ease. On the part of the teacher, the papers can be scored much faster and with greater ease than is possible in the case of traditional examinations. In the latter it is necessary to read through a great many papers each of which usually consists of a large number of illegibly written pages which are lacking in uniformity of arrangement, and which may contain elements of merit and demerit hard to locate. This necessitates so much time and energy for scoring that the teacher is limited to the use of only a few such examinations during the course of a term. The new type of test solves the problem in several ways: (1) Many tests can be administered in the process of a course of study, thereby getting a truer measure of the process measured; (2) The time and energy saved may be used by the teacher in teaching a greater amount of material in a given time; (3) Teaching may be directed to better purpose since the situation may be adequately measured with repeated tests.

2. Limitations of Objective Tests.

There are definite limitations to the use of objective tests which we must not overlook in this discussion. In order to be more specific we shall treat these limitations under the separate types of test.

a. Limitations of the standardized test. The standardized test has certain limitations which do not apply to the unstandardized test, a few of which are as follows: (1) cost; (2) comparatively few in number; and (3) specific needs of

the class not always met.

Cost. Because standardized tests are produced at considerable expense they must be sold to users at a price that is often prohibitive, except in very limited numbers. Such a limitation of the number of tests used lessens their value greatly. To secure the best results, tests must be repeated often during the course of a term of school, several tests being necessary to measure one function pertaining to the mastery of only one subject. The cost per pupil, when adequate testing is administered, runs up into figures that stagger the ordinary school board.

Small number of Tests. Standardized tests are growing in number every year. Yet for a given subject, and for certain processes, a comparatively limited number are available at present. The teacher is constantly facing new emergencies which call for the use of a type of test not available on the market, and if she were limited only to standardized tests, she could get no help in such emergencies.

Specific needs of the class not met. Growing out of the above limitation is the fact that standardized tests do not meet many of the specific needs of classroom instruction. These needs are often temporary and subordinate to the greater needs that are, indeed, met by standard tests, but they are nevertheless felt needs for which the teacher must look elsewhere for aid. There are times in the conduct of her class work when the teacher needs a quick testing instrument, not necessarily standard nor absolute in its

results, but simply a teaching device which will stimulate her pupils to greater and more sustained effort. For the solution of these smaller, though insistent, problems we turn to unstandardized tests for help.

b. Limitations of the unstandardized tests. For a treatment of the limitations of the unstandardized tests in their various forms, reference is made to the able treatment of them given in Paterson's little book, The Preparation and Use of New-Type Examinations.[1] Among their limitations may be mentioned the following: (1) They cannot take the place of standardized tests; (2) They are only relative in their measure; (3) They are not as objective as standardized tests; (4) they are only a supplement to standardized tests.

However, there are a number of definite advantages to be seen in the use of unstandardized tests which correspond to the limitations noted under the discussion of standardized tests. These advantages, which do not require elaboration, are, (1) They are inexpensive, and therefore equally practical in all types of schools; (2) They are relatively flexible as to purpose and construction to meet definite emergencies; (3) They can be constructed in any number as the teacher sees the need for them in her work.

---

1 - Paterson, Donald G.- Preparation and Use of New-Type Examinations. Chap. V

3. Traditional Examinations made Objective.[1]

Closely related to the subject of objective tests, and deserving consideration at this point, is the use of traditional type of examinations which have been treated in such a way as to secure a rather satisfactory degree of objectivity. The following steps have been suggested for making Examinations objective: (1) setting up standards for scoring; (2) assigning relative values to the standards such as : 1 point for neatness, 2 points for organization of facts, 3 points for pertinency of facts to topic, and 4 points for validity of the argument, making a total of 10 possible points for each question; (3) and scoring. In scoring this type of examination it is necessary first, to consider each question separately throughout the entire stack of papers, thus decreasing the subjectivity of scoring; second, to take the standards one by one and evaluate each answer on the basis of these standards; third, when one question has been carefully evaluated in this way through all the papers, pass to the next question and go through the same process. Such steps tend to eliminate the personal element in judging the work of the pupils.

It is claimed that this method of handling the essay type examination makes it just as objective as the new type tests. Special care should be exercised, however, in the construction of the questions in order that a ready

---

1 -Russell, Chas. - Classroom Tests, Chap. VIII

analysis of the pupil's answers may be made. The chief advantage of this type of examination is in the opportunity it gives the pupil for independent thought and organization of material. Even so, the exclusive use of essay examinations would be very unwise in the light of the limitations already noted in connection with their use. The most important drawback is in the enormous amount of time and labor involved in the scoring of papers, the papers of one typical class requiring several days for objective scoring. Their greatest value is seen in using them along with some form of unstandardized test or as a part of a battery of tests.

The three types of examination discussed above should be used interchangeably in the classroom to meet specific needs of the pupils. When standardized tests are used they should be used alone, and with a specific purpose in view. The various forms of informal tests and the essay type examination may be used in combination in a single testing problem. This enables the teacher to approach the same situation from several angles at the same time. It also enables the pupil to give the best account of himself since he may respond more readily to one type of exercise than to another.

B. Scales

Scales describe performance or an educational process in terms of uniform or graduated units. They are used in two different ways; first, as applied to the results of tests, and second as rating scales applied to personal qualities and

qualifications, or to recitations or other educational processes, or to school plants and curriculum materials.

1. Test Scales. The difference[1] between a test and a scale is that a test secures the performance of the pupil in certain definite ways, while the scale measures and describes that performance according to definite units. Scales describe performance in three ways:

a. The Rate of Performance. In this case units of equal value are indicated on the scale, and the number of units a pupil is able to accomplish in a given time describes his rate of performance. A scale of this nature may be illustrated as follows: / / / / / / / / / / / / / / / / / /. Each unit has the same value, and it is a matter of speed of performance as to what the pupil's score will be.

b. The Difficulty of performance. Here the units are graduated upon an increasing scale of difficulty, thus:

/ / / / / / / /

The pupil who can perform up to the first unit only is rated as very low in comparison to the one who is able to pass 6 or 7 units, and each unit gives a certain difficulty value to his performance.

c. The Quality of Performance. Here we have a scale consisting of a series of sample products ranging all the way from a very poor to a very excellent example, with a value

---

1 - Monroe, W.S. -*Theory of Educational Measurements*, Chap. II

assigned to each. The Ayre's Handwriting scale is one of this type. By comparing a pupil's handwriting with the scale and finding the sample it is most alike on the scale a quantitative value may be assigned to a performance which is qualititative.

These three measurements of performance are not always of equal importance in a test, and sometimes only one of them is measured, sometimes two, and sometimes all three, according as the examiner sees the need for evaluating them.

There is some chance for difference of opinion in the rating of a quality performance by comparing it with the scale, and therefore this particular type of scale is not so objective as the others. Nevertheless, it is very useful in furnishing a standard of achievement, and in estimating the emphasis which should be placed on that particular part of classroom instruction.

3. Rating Scales

There are in general, two types of rating scales: Personal rating and Score cards.

a. Personal rating.

(1). Pupil-rating scales. In these, we have usually a scale for each school habit or trait consisting of three degrees of excellence from very poor, through medium, to excellent and the rating of a child for any particular trait is estimated along the scale according to the best opinion of the scorer, not necessarily directly on one of the three points given, but shading between them as in the opinion of the scorer the child possesses more or less of a trait.

(2). Teacher-rating scales. These have been devised for the improvement of teachers, and cover every phase of teacher personality which make for success in teaching. A good example of rating scales for teachers is found in one constructed by T.H. Shutte, which is composed of the following general qualities: (a) Personal and social qualities; (b) cooperative qualities; (c) Leadership; (d) Scientific and professional attitudes; and (e) Teaching ability. In the scale mentioned each of these general traits is analyzed by means of from 11 to 27 specific questions each calling for a rating in a single aspect of the general trait. The rating in the general trait is the average of the specific ratings. A combined rating may be found which is the average of the ratings in the five general traits.

In making practical use of teacher-rating scales care must be exercised lest the teacher misunderstand their real purpose and become offended at this effort to "criticise her weaknesses." It is advised at the start the teacher be given a scale on which to rate herself without requiring that she submit her rating to the supervisor. The supervisor may even have his staff rate him as a supervisor, announcing the composite scores for their interest. The subjective element enters into even the most careful rating. Nevertheless the scales serve a very useful purpose in two definite ways, first, in setting up a high standard, and second in evaluating the elements of this standard. The objectivity of scoring or the lack of it, makes little difference to the values

which come to the teachers from their use. The value of teacher-rating scales is seen both as a guide to improvement and as a measure of present qualities.

What has been said of teacher-rating scales may be said of pupil-rating scales in general. Both are designed primarily for stimulating self-improvement by holding up certain well analyzed ideals for achievement. They are sadly abused if applied to individuals in an adverse and critical way, or for getting a strictly objective rating of the individual, and the supervisor should understand that their chief use is in self-guidance and self-improvement.

Schutte[1] gives the following uses to be made of the teacher-rating scale: (1) To single out qualities for improvement; (2) To analyze a teacher for employment and promotion; (3) As an aid in the supervision of teachers by providing a concrete basis for conference; (4) For use by classes in education to point the way to analysis; (5) As a guide in observing others for self-improvement; (6) As a guide and basis for discussion in teachers' meetings to produce consciousness of details and teaching merit.

The limitations of these scales have already been brought out in the above discussion, but two definite limitations should be singled out here before leaving the subject; (1) They are much less objective than tests. In no sense can they take the place of tests in evaluating personal

---

1 - Schutte - Scale for Rating Teachers.

traits and achievements. (2) They are peculiarly subject to the "halo" weakness. The scorer will invariably tend to rate high the person to whom he may be attracted personally, even though the personal traits in the scale have been carefully analyzed for more objective observation.

b. Score-cards.

The score-card is a "systematic enumeration of the elements that make up any whole so that in passing judgment on this whole attention is directed to every significant feature that enters into it."[1] A score-card is composed of a series of scales designed for the measuring of complex situations and processes,[2] such as school plants, an educational system, teaching equipment, materials of study, or the process of a recitation. Two outstanding values are to be realized in the careful use of the score-card:

(1). It forces the investigator to explore every relevant phase of the process or situation. Personal interest in any specific phase is thus checked up, and the scorer seeks out many things he otherwise might possibly have overlooked.

(2). It is a reminder of relative values. Standard values are assigned each item in the score-card and composite values of various sections have been carefully balanced in accordance with expert opinion. An investigator who has been in the habit of assigning undue value to one phase to

---

1- The Indiana Survey-Vol.II. Part III,Chap.VII,p.106.C.C.Peters
2- Hanson, W.L. -Supervision of Religious Education Through Objective Tests and Measurements, B.U.Bulletin July 10, 1924. p. 4

the discredit of another may refer to the values assigned on the score-card for checking up his own personal bias.

SUMMARY

Two general classes of measuring instruments have been discussed in this chapter, with their varieties of types:

A. Tests

1. Standardized Tests
2. Unstandardized Tests
3. Traditional type examinations made objective

B. Scales

1. Test Scales
2. Rating Scales

It has been the purpose of this chapter to point out both the value and the limitations of the instruments of educational measurements. Both the critic and the enthusiast would do well to weigh carefully the arguments of both sides of the question of the practical use of tests and measurement before making a final decision in regard to them.

PART TWO

SUPERVISION OF RELIGIOUS EDUCATION THROUGH OBJECTIVE MEASUREMENTS

INTRODUCTORY STATEMENT:

A full treatment of the subject of Part Two would necessitate the repetition of much of the content of Part One, for the two fields are closely related. Therefore, the procedure in Part Two will be largely that of facing problems that are peculiar to Religious Education, taking for granted that the common problems have been adequately presented already. Repetitions will occur, but an effort has been made to repeat only in cases of needed emphasis.

It is assumed that Religious Education is a vital part of the general scheme of education and we pass by that phase of discussion for the central problem as indicated in the title of Part Two.

CHAPTER IV.

THE SCIENTIFIC NATURE OF THE SUPERVISION OF RELIGIOUS EDUCATION

THE SCIENTIFIC NATURE OF THE SUPERVISION OF RELIGIOUS EDUCATION

A. Definition of Religious Education

1. In terms of Religious changes produced in the child.

In Part One a definition of education was given based on Thorndike's quantitative view of the process, and it was pointed out how useful this definition was as a basis for the discussion of educational measurements. A similar advantage is sought in Part Two by defining religious education in quantitative terms. Such a definition, stated simply, would be the production of religious changes in the child. The amount of change in a given time is more or less subject to objective measurement, and may be scientifically treated.

2. In terms of an analysis of the elements of religious education in which definite changes are sought. Such an analysis might be made in several different ways, but the following is here suggested:

a. Knowledge of the materials of religious education, (1) Biblical knowledge; (2) Extra-biblical knowledge.

b. Attitudes and appreciation, (1) Faith in God and reliance on him; (2) Devotion to Christ; (3) Attitude toward life - a quest or adventure; (4) Appreciation of literature and art.

c. Skills in Christian living, (1) Prayer; (2) leadership; (3) Cooperation, etc.

d. Conduct. Personal responsibility and social self-control.

B. Objective measurements Essential to effective supervision of religious education.

1. A certain amount of measuring must go with all supervision. Consciously or unconsciously, the supervisor applies his own method of evaluation to every educational Situation. Old methods of education produced many good results, but it was due to the fact that some one or more in the school was able to "size up" the problems and take advantage of proven values and methods. No progress is possible without some degree of measurement.

2. Subjective methods are inefficient.

The traditional method of measuring educational progress in religion has been even more subjective and biased than that in secular education. The latter has at least demanded papers and examinations by which to evaluate achievement. The Sunday Schools of America, which have borne the chief burden of religious education in the past, have been so very haphazard in their methods that they have commanded the respect and loyalty of only a small proportion of the population, and have proven entirely inadequate to meet the demands of the modern world.

3. Objective Measurements facilitate the work of supervision. This is done by making his evaluations factual, definite, impersonal, and more reliable as a guide to the improvement of school procedure. Three definite steps in scientific supervision were named in Part One, (which I here repeat in a brief way) as made possible by objective measurements: (a) They assist in establishing goals for

school achievement; (b) they evaluate the conditions of progress; (c) they furnish checks by which progress is measured.

C. Principles of Supervision in Religious Education

In Part One a comprehensive discussion of the principles underlying educational supervision was given. The same principles apply to supervision in the field of religious education as given there. But this particular field presents peculiar problems not discovered in secular education and for that reason attention should be directed to these problems as they affect supervision.

1. Voluntary teachers. In the public schools the teachers are paid a definite salary on the basis of which considerable pressure may be brought to bear by the supervisor for the improvement of instruction. A scale of graduated salaries stimulates the teachers to a desire for advancement and larger salaries. No such advantage exists in religious education, and the supervisor must appeal to other motives, as social approval, love of children, desire to serve, etc. His task is in purifying these motives, winning the teachers over to the higher motives only, and on the dynamic of these motives to inspire them to utilize every help in the efficient conduct of their work.

2. Indifferent Public Opinion. This is an administrative, as well as a supervisory problem. But it touches supervision in the attitude of the pupils to their work. The

public schools of our country are supported by an interested public opinion, and the pupils are generally affected by this to the extent of possessing a desire for real achievement in their work. But in religious education, public opinion is lacking except with a very few families in each community who support the church. Even here the enthusiasm and moral backing is often decidedly less than that given public education. This attitude reflects itself in the attitude of the children who very naturally sense the difference in value placed on the two forms of education by the parents.

The supervisor has this handicap to overcome in religious education in that both the parents' and the pupils' interest must be aroused to the end that the pupils may make more satisfactory progress in their work.

3. The general attitude toward the scientific approach to religion. This is a real problem in many places for the reason that religion is thought to be too intangible and spiritual for such an approach. The supervisor's task here is to so handle this phase of his work as to conserve every value held dear to the church and to show that this approach, rather than destroys, enhances the method of realizing those values.

SUMMARY

Summarizing the discussion of Chapter IV., we are bold to assert that the supervision of religious education is a scientific task, and that its procedure must be similar in every way to that of the supervision of secular education, as far as the general considerations of definition, principles,

and their use of measurements are concerned. Succeeding chapters will take up the more unique problems of religious education as they differ from those of the secular field of education.

CHAPTER V.

MEASUREMENTS AND THE SUPERVISION OF RELIGIOUS EDUCATION

MEASUREMENTS AND THE SUPERVISION OF RELIGIOUS EDUCATION

A. Special problems in the use of measurements in Religious Education.

Having discussed the scientific nature of supervision in religious education and finding it to be comparable to that of secular education in general, we come now to certain definite problems that are more or less peculiar to religious education as it aims to make use of educational measurements.

1. Problems growing out of the limitations of the present measuring instruments.

Two serious limitations are here noted, which are rather expected by reason of the relatively short time that has been given to measurements in this field. Both limitations should diminish in the future, though, perhaps, they may not altogether disappear.

a. The instruments are limited as to the number available. Goodwin B. Watson[1] gives a summary of the tests now available in the field of religious education in his recent work in this field. This does not include other types of measurements, such as is contained in the Indiana Survey, Volume II for use in evaluating the curriculum, church plants, etc. but it does reveal the fact that the instruments now at hand are comparatively limited. Such a limitation as this narrows the usefulness of measurements in several ways:

---

1 - Watson, G.B.-Experimentation and Measurement in Religious Education. Association Press, N.Y. 1937

(1). A very inadequate measure of the educational product is given. We have already had occasion to say that an adequate investigation of all the traits, or even of one trait, in a child with all his changing moods and emotions requires very many uses of similar tests over a period of time. The standardized tests now ready for use will do no more than start us in such an investigation.

(2). We are forced to a limited and uncertain standard of educational instruction for the present. The standards of Religious Education are now in a state offlux pending further investigation and research into the problem from many angles.

(3). It necessitates undue emphasis on informal tests and personal judgments. Objectivity of treatment is thus curtailed to a large extent, although the use of teacher's classroom tests, with their objective scoring, is valuable.

b. Measuring instruments in religion are crude at present.

(1). They now measure the lower levels of religious growth, such as the acquisition of factual materials, though a few measure ethical judgments and attitudes. Only one attempts to measure conduct, and in that one the measurement is limited to a single virtue. The field is in need of much intensive research work and the construction of measuring instruments which will further objectify the whole task of religious education.

(2). They measure spiritual traits only indirectly. Perhaps the time will never come when direct measurement of spiritual traits will be accomplished. But as the situation

now stands,too few contacts are being made even indirectly, to get any true picture of the spiritual growth of a child. There is a need for many more of these indirect measures, in the absence of direct observations, so that a more composite view may be obtained of the central and controlling forces governing the life of the child.

(3). They secure but partial and inadequate results. The question of relativity comes in here to modify this statement, as, indeed, in all the above points. In comparison with old methods, the results obtained from the few measurements we now have are superior and much more adequate, but as compared to the ideal of standardization and objectivity, as applied to the whole scope of religious education, the present instruments appear hopelessly inadequate for the task imposed on them.

2. Problems growing out of the nature of the process measured.

a. The intangible nature of spiritual factors.

It is extremely difficult to analyze and objectively investigate spiritual characteristics. For instance, just what is reverence? Psychologists attempt to give us some of the factors entering into reverence, but the psychologists leave us just where we were before in using terms that are themselves just as hard to analyze for measurement. Our best approach at the present time is to analyze the specific acts and attitudes of reverence.

b. The necessity at present for an indirect approach to spiritual factors. This is accomplished through specific skills and abilities which make up those factors. The child's ability to think in terms of the concept "God", his acquired attitudes toward God and prayer, his ability to express himself reverently, together with other basic elements entering into his acts of reverence, seem to be the only basis at present for evaluating the trait called reverence. This limitation runs throughout the entire list of traits found in the goals of religious education.

c. The difficulty of securing performance consistent with the ability or attitude being measured.

(1). In the first place, it is hard to control the conditions of measurements so that a functional relation is constantly maintained between performance and the thing we are attempting to measure. The pupil will be influenced by the least turn of affairs surrounding the measuring project. Often the unsuspected attitude of the teacher will cause the pupil to perform very differently from that of his normal and spontaneous reaction. To eliminate this difficulty the test is often made in a way that the pupil will not be aware he is being tested at the time. This may seem unethical at first sight, but do we not procede along just this line in our personal judgments? After all, our objective is to get the real state of mind of the child in his normal, everyday reactions to life situations, and not the highly intellectual attitude assumed when he thinks he is being tested, or is to be graded for his performance.

(2). Objectivity of scoring becomes more difficult as we pass into the more elusive levels of human achievements. Religious traits are so complex, and are often found in company with each other so closely related and intertwined that it is very difficult to mark off the limits of a single trait and say we have just so much of it and no more. Therefore on any scale that may be devised, one scorer will differ from another in evaluating the same product or process of religion. The best that can be accomplished for the present is to analyze, as best we can, every factor to be taken into account, break them up into as many parts as they are capable of, and then attempt to train teachers and leaders to evaluate them as nearly as possible in accordance with standards that have been agreed on by those who are competent to evaluate them.

It is highly necessary that we recognize the limitations and the problems that face us in planning the use of measurements, so that we may modify our testing program from time to time in the light of these problems. Some will say that the problems are too great and the limitations are too serious to allow for any worth while results at all from the use of measurements, and therefore there is no sense in wasting time with them. Others are likely to become too enthusiastic and fail to see the limitations and the problems involved, and try to make them a "cure-all" for every educational trouble. It is cowardly to take the former position, and extremely foolhardy to assume the latter. The avoidance of either of these extremes will be necessary for the successful supervisor of religious education.

B. The Measurement of Specific Abilities Conditioning Religious Growth.

In the discussion of measurements as applied to religious education, we shall find it a help to specify certain definite abilities which condition religious growth.

1. Abilities Classified.

Authorities classify the basic abilities conditioning religious development in various ways. The classification given here is merely suggestive, and does not claim to be final in any sense.

a. The acquisition of content material.

(1). Biblical and Extra-Biblical subject matter.

The acquisition of the factual information contained in a given course of study is the simplest form of education, and it is comparatively easy to analyze and measure the performance of a pupil on this level. The importance of this sort of knowledge, however, has been underestimated by many who would discredit the method of measurement proposed here. The facts of religious knowledge are basic to forming religious judgments and to the building of Christian character, and they deserve their rightful place in the order of things religious.

(2). Ethical and Religious Ideas.

The ability of the pupil to "think in"[1] the subject matter is not so simple as "knowledge about" that

---

1- Paterson, Donald G. -Preparation and Use of New Type Examinations. p. 6

material. It is harder and correspondingly more important to evaluate the ideational elements of the child's knowledge of a course than to measure his mastery of the facts. The number and nature of the pupil's ethical and religious ideas, the ability to relate the facts he has learned to concepts, and the intelligent use of a vocabulary in religious and ethical ideas, all are important steps in his religious training, and are therefore worthy of careful measurement.

b. The acquisition of skills.

It cannot be doubted that a child should be given training and guidance in making judgments vital to his ethical and religious development, and in how to form decisions in the light of his judgments. He will need to learn how to make discriminations between one course of action and another, how to forsee possible consequences of each choice, and how to use his best judgment in deciding the course of action to pursue. The ability to pray, to sing, to speak intelligently on religious subjects, to participate in discussions, and to assist in the solution of problems, are evidences of the acquisition of religious and moral skills which are essential to the full development of the religious life.

c. The acquisition of dynamic factors.

There are many emotional and dynamic elements in the development of the religious experience. The child's interest in various things, his attitudes socially and religiously, his ideals and motives from which spring all the conduct of life, his prejudices, his likes and dislikes, everything that makes up his emotional experience, whether in school or out, power-

fully influence his religious development. It is comparatively recently that any efforts have been made to measure these psychological phenomena and now that science has led the way, religion is taking advantage of the facts of scientific discovery to make its work more effective and lasting. Again, it is more difficult to measure dynamic qualities objectively than the more factual attainments of the memory, but its importance makes the effort, however imperfect, well worthwhile.

d. The acquisition of character traits.

This is a synthesis of all the above points, as character is vitally concerned with each of the factors under discussion above. And so we have now come to the most complex of all - the process of uniting the intellectual, emotional, and the performance elements of human nature in order to erect conduct controls that will result in the full rounded Christian character.[1] However, with all its complexity, specialists are attempting to measure character in its entirety as well as in its parts. The Committee on Character Education Inquiry of Teacher's College, Columbia University, led by Professors Hartshorne and May, is attempting research work in this difficult field at present. The report of this committee before the Religious Education Convention in Philadelphia, Pennsylvania, on March 9, 1928, showed significant progress toward its goal.

---

1 - Religious Education, Vol.XXIII, March, 1928. "Taking our Religious Measure", by A. Holmes.

3. Measuring Instruments now available for use in Religious Education.[1]

Slowly those who are experimenting in the measurement of religious characteristics and skills are evolving a system of standardized measuring instruments to meet the growing demand in the field. Early achievements in this direction were, as would naturally be supposed, in the simpler processes, such as the knowledge of facts, and other intellectual skills. But we are aware that "out of the heart are the issues of life", and there is a more insistent demand for accurate evaluation of the motives, emotions, and purposes of the child.[2] In the catalogue of instruments given here we have followed, in general, the order of difficulty, from simple to complex, from the informational type to the emotional and character type of tests. A warning is given by Watson[3] concerning the use of certain types of instrument dealing with the emotional factors in child life, especially in the attempt to cure unstable cases. Always the advice, if not the personal assistance, of a trained specialist should be secured when such instruments are used in the ordinary Sunday School.

In the following list of instruments; we have depended largely on the thirty instruments now available given by Watson.[4]

1- Religious Education, Vol.XXIII, March, 1928, "What Tests can we use in Church Schools", by Goodwin By.Watson
2- Teacher's College, Contribution to Education, No.176 "The Measurement of Fairmindedness". p.1, G.B. Watson
3- Religious Education, Vol. XXIII, March, 1928, "What Tests can we use in Church Schools", p.217, By G.B.Watson
4- Watson, Goodwin B. Experimentation and Measurement in Religious Education, Chap.IV.

a. True Tests, of Pencil and Paper Type.

(1). The Whitley Biblical Knowledge Tests.

These consist of one form for the Old Testament and two forms for the New Testament, and have for their purpose the measurement of the knowledge about the Bible. The exercises are in the form of questions or statements with a multiple choice answer printed opposite for the child to mark. The scoring is made objective by counting only the answers that are right.

No norms are available on these tests, but they have been given to a large number of unselected groups of children of all ages; the scores have been statistically treated, and they are in the process of being standardized as to reliability and validity. Ages to which they may be given are from ten up, and the time required is 30 minutes for each form.

(2). Giles Sunday School Examination A.

This is a True-False test, the purpose of which is to test the Biblical information based on the Uniform Lessons of the Sunday Schools. It consists of 3 parts: Old Testament knowledge, New Testament knowledge, and Ethical Ideas. These tests are accompanied with norms from 2,000 pupils and the reliability as shown between Old Testament and New Testament sections is .50. It is interesting to note an almost total lack of correlation between either the Old Testament or New Testament Tests on the one hand and the ethical judgment test on the other, which seems to indicate a lack of carry-over from knowledgment of the Bible to its ethical

application in life.[1]

(3). Church School Examination Alpha.

This is the Boston University revision of the Giles Sunday School Examination A, using the same material but altering its form to eliminate the objectionable features discovered in True-False testing of biblical knowledge. Its purpose is to test the knowledge of biblical facts as contained in the Uniform Lessons of the Sunday School, and to test the knowledge of ethical principles which are expected to result from a proper mastery of this biblical material. It consists of 25 exercises in the Old Testament, 25 in the New Testament, and 25 on ethical principles, all in multiple choice form. It may be given to children 10 years of age and upward. This test has not been standardized.

(4). Laycock Test of Biblical Information.

The purpose of this test is, as the name indicates, to test the knowledge of statements made in the Bible, without regard for creed, church, mental, or moral ability. It is multiple choice in form and is in 3 sections of 25 exercises each. It is designed for use with ages 12 to 16, and each section should be timed. A scoring sheet accompanies the test, and objective results are obtained. Norms are established on the basis of testing 1115 cases in Canada. The mean scores of the total number are given, as also those of the girls alone and boys alone and those of each age group, 12 to 16.

1- Indiana Survey, Vol. II. Discussion on this subject. p. 396

(5). Multiple-choice Test of Religious Ideas.

This is produced by Boston University, School of Religious Education. The purpose is to ascertain the actual religious ideas of children or adults, whether in the church school or out of it. It contains 10 questions, each having 15 suggested answers. The pupil is to check the 5 best in each group, and thus give an idea of his mental equipment in religious ideas. The scoring and reliability of this test are based on the criticism of a number of religious leaders. No norms are furnished.

(6). Union Tests of Religious Ideas.

The purpose of this test is to measure intellectual factors in religious development. It consists of questions usually answered by "Yes" or "No", on ideas about God, Jesus, Prayer, etc. Also, it contains a completion exercise giving the story of the Bible testing the knowledge of its contents and relationships. Form I is designed for grades III to VIII, and Form II for the High School and adults. Score sheets are furnished with the tests. A reliability of .90 was obtained for form II, and norms are being established. The validity shows a correlation of .25 with intelligence.

(7). Union Test of Ethical Judgment.

The purpose is to measure the development of enlightened social and ethical standards, especially for religious educational groups. It contains exercises and questions of the "yes-no" type, or exercuses of the "excellent-fair,poor" type, dealing with choosing courses in school, finding life work, and many other everyday solutions of life

situations. Form I is for grades III to VIII, and form II. is for High School and Adults. Score sheets are furnished with the tests. The reliability of form II is.92, and the validity shows a correlation of .23 with intelligence.

(8). Fernald Ethical Perception Test, No.27,105.

The purpose is to evaluate the possession of Knowledge of right and wrong. This is important in placing responsibility, but has little worth in measuring actual conduct. It consists of 10 ethical questions, 7 of which are of the "yes-no" type, and 3 answered by taking one or the other of alternate courses suggested. It is scored by computing the percentage of correct answers.

(9). Fernald Ethical Discrimination Test, No.36,035.

The purpose of this test is to furnish a way of judging the moral work of the pupil in order to understand his mentality and to correct judgment of his responsibility. Some insight into the moral "make-up" of a pupil is obtained from a study of the arrangement in the order of their worth of ethical offences from 1 to 10. The test consists of one large card and ten slips of cardboard containing offenses to be arranged in a series from the least to the greatest in the order of gravity. Directions for giving are furnished with the test.

(10). Brotemarkle Comparison Test.

The purpose of this test is to find, by comparative arrangement of moral concepts, the pupil's estimate of its meaning. It consists of 7 test exercises,7 words in each with varying degrees of meaning, with the two extremes lo-

cated at each end of a line. The pupil is to place the remaining 5 words in the order he thinks right along the line between these extremes. This is designed to be administered to High School students, or to adults.

(11). A Brief Test in Religious Education

The purpose of the test is to give in brief time, a fair idea of religious and ethical development of groups in the Church School. It consists of 20 "yes-no" type questions on the Bible, God, Jesus, etc. and 10 instances from every day life in which the pupil is asked to judge which of 4 or 5 alternatives would be the best thing to do or say. It may be given to any age above 9 years, and the time required is about 15 minutes for pupils in the 8th grade. Self-correlation of the test shows r= .45. The validity is estimated on the basis of the judgment of leaders of Boys' Clubs.

(12). Koh's Ethical Discrimination Test.

The purpose of this test is to measure significant ethical knowledge and ability to make essential moral judgments, interpretations and decisions. It consists of exercises giving social and moral situations from life, and the pupil is asked to make his choice of alternate decisions. It is to be given to pupils 11 years old and up. Each exercise must be carefully timed, and 20 minutes given to the entire test. Scoring directions are given in an accompanying manual.

(13). Chassell Foresight of Consequences Test.

The purpose: To ascertain the young person's ability to foresee possible consequences of certain given courses of action and to assist in developing proper decisions in life in the light of foreseen consequences. It consists of stories giving concrete situations with possible reactions to each on the part of the pupil, and possible consequences which might follow. The pupil marks the consequences he thinks important and his choice of action. It is designed for ages 12 to 16.

b. Tests - Performance type

Voelker Tests of Trustworthiness.

The purpose of these tests is to create situations that are just as natural as may be obtained, in which the boy shall respond in a normal fashion to an environment which has been carefully designed to bring out the degree of trustworthiness he possesses. It consists of 10 tests, such as giving a boy some money to make a purchase at the store and noting whether he returns the right amount of change; and placing a boy at an electric signal button with instructions to press at short intervals for a period of 20 minutes and noting on a recording device how closely he adhered to instructions. These tests were discovered by Voelker to have a very satisfactory correlation. Some have questioned the validity of these tests, and it is significant that they have been omitted from Watson's exhaustixe description of tests in this field.

c. Questionnaires or Surveys of Interests, Attitudes, etc.

(1). A survey of Public Opinion on Some Religious and Economical issues. (Watson Prejudice Test).

The purpose is to measure fairmindedness as contrasted with prejudice; to show amount of prejudice in agreement with certain typical points of view, conservative and radical, within religious, moral, and economical questions. It consists of 6 forms, each of which is designed to draw out the judgment of the pupil on questions about which considerable prejudice now exists. It should be used with seniors in High School or above. Scoring directions accompany the test. Pupil's prejudice is estimated by comparing with fairmindedness standard. Norms for various groups are given and compared to fairminded score.

(2). Hart Test of Social Attitudes and Interests.

The purpose of this test is to show predominant likes and dislikes, attitudes and points of view. It consists of lists of conditions or activities which the pupil marks "plus" or "minus" as he likes or dislikes each. It is to be given to ages 13 or over, and requires about 40 minutes to administer. No time limit is placed on it. No norms have been established.

(3). Bogardus Social Distance Test.

The purpose of this test is to measure the extent to which persons of other races, nations, or economic classes are welcomed into the fellowships of social life. In the racial test the name of a race is given, followed by 7 columns representing 7 stages of intimacy ranging from will-

ingness to admit to kinship by marriage down to willingness only to admit as aliens. It is to be given to adults and requires about 30 minutes to administer. In scoring, a group profile may be made showing the willingness of the average in the group for admitting various races and classes into national life.

(4). Test of Racial Attitudes. (Watson)

The purpose here is to measure changes in attitudes toward persons of other races. It is composed of 36 statements to which the pupil adds one of the following words to complete the sense as he views it: "All, most, many, few, no -- Jews will try to get the best of a bargain even if they have to cheat to do so."

(5). Test of Social Relations. (Ream)

The purpose is to reveal interests in terms of acquaintance with vocabulary of certain activities such as baseball, poker, hymns, popular music, literature, etc.

(6). Downey Will-temperament Tests.

The purpose is to determine the temperamental traits of individuals through a series of motor reactions. This test is based on handwriting to a large extent. Various speeds of writing are called for, copying a model, writing with the eyes closed, and while counting taps. It is designed for ages 15 and up. Time required is 50 minutes. The scoring is a long process, but gives a profile showing speed of movement, etc. The reliability is from .05 to.40 The validity shows a correlation with ratings of -.65 to.54, mean .0 to -.25

(7). Lundholm Emotional Crosscut Tests

The purpose is to find out if there is a standard reaction to certain mental content in the human individual and if there are sex differences in this response. Also it aims to study the influence of social inhibitions on responses. It may develop as a complex indicator in the mentally deranged.

(8). Pressey x-o Tests for Investigating the Emotions.

Purpose to reveal individual differences in emotional and affective make-up, associational tendencies, moral sentiments, and anxiety tendencies. It consists of 4 parts with 600 elements for crossing out or choosing from among lists of words. The scoring takes notice of affectivity, or the relation of words crossed out to norms for college students, and idiosyncrasies, or relation of the number of words encircled which are not usually encircled. Norms based on 114 college students are available for this test.

(9). Woodworth-Matthews Personal Data Sheet.

The purpose is to show the general emotionality, nervous, and mental stability, of adolescents. It consists of 100 questions such as, "Do you sleep poorly? yes-no". This data should be carefully handled by one who is scientifically trained in emotional analysis of individuals. The reliability is .55 to .90 depending on investigation. Validity is .12 to .66. Norms are available, at extra cost.

(10). Kent-Rosanoff Free Association Test.

The purpose of this test is to reveal abnormalities of emotional reactions. It is composed of words to which the frequency of each of the common responses is known.

The test is given to individuals alone, noting the exact time for each reaction. In the scoring of the test unusual time for reactions indicates conflict. Any unusual response-deviation from the norm -- shows the degree of idiosyncracy. Norms are based on 1000 adult responses giving frequency of each association.

(11). Emotional History Record. (Experimental purposes only).

Purpose: To reveal students needing help in emotional adjustment, and to study relationships between past experiences and present symptoms. It consists of 4 parts. Parts I and II reveal neurotic symptoms with interests of the subject. Parts III and IV give a record of past experiences which may have influenced emotional make-up. This test is to be given to adults and it requires two hours. It should not be administered by amateurs. Scores are best used as a picture of the present status of the individual.

(12). Colgate Emotional Hygiene Test.

The purpose is to discover unhealthy emotional outlets and emotional types among college students, and to help in vocational guidance and personality adjustments. It consist of a series of elements, each followed by graphic scales. Traits grouped according to clinical entities such as psychoasthenoid, schiziod, neurasthenoid, hysteroid, introversion and extroversion. Two schedules are used in combined time of one hour and it is self-administered. It is

useful in personnel work in industrial plants and in colleges. Tests are scored by means of stencil which indicates the deviation from the norm.

(13). Hart Personnel Assayer

The purpose is to give information about a person's interests and attitudes which might be helpful in vocational guidance, to differentiate certain types of people from other types. It consists of 19 lists, 15 items each, to be marked "yes" if liked and "no" if disliked. Then the 5 things about which one feels most strongly are underlined, and the one thing which arouses most feeling is double underlined. It is to be used with ages 12 and up. The time required for its administration is 40 minutes or more. There is no time limit. This test is in the process of standardization.

(14). Miners' Analysis of Work Interest Blank.

The purpose of this test is to help discover special interests and abilities by suggesting how to observe one's own likes and dislikes. It consists of 38 paired contrasts, such as; indoor-outdoor, etc. It is to be administered to High School pupils and adults. The reliability is .70 for all interests but one. The validity is in separating women taking secretarial course from those taking home economics, men interested in machine construction. Norms are based largely on 800 students at the University of Kentucky.

(15). Interest Analysis.

Purpose, to discover occupational interests. It consists of four parts. Part I lists 70 occupations, to respond with 5 degrees of liking. Part II lists a variety of words to be marked liked or disliked. Part III deals with belief in superstition. Part IV, the Pressey Cross-out list, to cross out those which are unpleasant. To be administered to High School groups and over. The time required is 50 minutes.

(16). Freyd's Occupational Interest Blanks.

This consists of a list of 80 occupations for men (or 67 for women) with a chance to register any of five attitudes toward each.

(17). Fernald Achievement Capacity Test, No.19,433.

The purpose of this test is to discover that function of the mind called will, persistency, determination, pluck, or spunk in terms of muscle fatigue in units of time. A person stands with heels elevated from floor on a device equipped for recording involuntary tremors of the muscles on fatigue. The scoring is by noting the time, in seconds, a person persists in standing in the same position.

(18). Thurston and Chave Test of attitudes toward the church.

This test aims to place a student along a scale of varying degrees of favor toward the church from the extreme of disgust for it to that of loyalty to it. It consists of a large number of statements about the church which cover this range of attitudes and which are weighted by submitting

them to competent judges. These statements are arranged in chance order, and the student marks those which express his own attitudes. Care has been exercised to get a normal response. The results give a profile of an entire class, and interesting comparisons are made between Freshman, Sophomore, Junior, and Senior classes in a college, from which may be interpreted the influence of college life on a student's attitude to the church.

d. Scales and Score Cards.

There are a number of rating scales and score cards for use in Church schools. Among them are (1) the Chassell Teachers' Rating Scale for Sunday School Pupils; (2) The Drew Measurement Chart for Sunday School Juniors; (3) The Mendenhall Self-Measurement Scale for High School Pupils; (4) The Mendenhall Self-Measurement Scale for Children in Grades V to VIII; (5) Coe's Teacher Rating Scale; (6) Score-Card for Measuring Curriculum Material in the Church School, by Peters; and (7) The Interchurch Standards and Score-Card for Measuring Church and Religious Education Plants, by N.L. Englehardt and E.S. Evenden.

CONCLUSION

The relation of measurement to the Supervision of Religious Education has been discussed as follows:

A. Special Problems in the Use of Measurements.

B. The Measurement of Specific Abilities conditioning religious growth.

It was shown that, in spite of the rigid limitations

which beset this problem at present, a creditable beginning has been made toward the production of a system of measurements in religious education.

CHAPTER V I.

ADMINISTERING A PROGRAM OF MEASUREMENTS IN A SYSTEM OF RELIGIOUS EDUCATION

ADMINISTERING A PROGRAM OF MEASUREMENTS IN A SYSTEM OF RELIGIOUS EDUCATION

In the present chapter we arrive at the practical stage of our discussion. Being made aware of the value and limitations of objective measurements, and of the painstaking care necessary in using them in a program of supervision; having reviewed the measuring instruments available, their use, structure, and method of administration; and having noted in some detail the theory underlying the use of measurements in religious education, we are now ready to consider the introduction and administration of a program of measurements in a local school of religion. A suggested procedure is offered in the chapter before us.

A. What a Program of Measurements is.

The purposes and uses of measurements as given in former discussions of this paper lie at the basis of any program of measurements. In planning a program for a local situation, however, it is necessary to consider the particular needs of the school for which the program is planned. The one outstanding need may be singled out and made the basis of the first program planned. Following this, a clear statement of the purpose of the program should be formulated, and specific purposes should be clearly stated with regard to each test administered. Each teacher and supervisor concerned should be thoroughly informed of the purposes to be served in the program. The steps to be taken in the administration of a program of measurements will be given further on.

B. Preparation for the Program.

1. The supervisor must first understand and appreciate the local attitude toward the subject of measurements. Likely there will be so little knowledge of educational measurements that little opposition will be manifest, at the same time this lack of knowledge presents a problem of acquainting the people with them. Whatever may be the situation, the first duty of the director or supervisor is to know the minds of the people on the subject, or what might be their reaction when they have been informed of the nature and use of measurements. What is their attitude toward innovations, etc?

2. He must have a sympathetic attitude toward the teachers as they face problems in the school. By saying nothing directly about measurements themselves, and at the same time showing an insight into the actual problems of the various teachers, avoiding undue haste in introducing something they do not understand, the supervisor may gain the confidence of his staff. It cannot be urged too often, nor too emphatically, that the whole-hearted confidence and loyalty of the teachers is of primary importance and basic to any constructive endeavor. The supervisor who sees and understands the difficulties of his teachers, and then tactfully points out possible solutions, may sooner or later suggest measurements as an aid to just such problems as they are facing. Many opportunities will arise for suggesting solutions as he comes into contact with the work of the teacher in her classroom and elsewhere. The perplexity of the teacher as to the real cause of lack of interest among

her pupils for instance, may open the way to a discussion of the problem of classification, and lead to a possible suggestion to study her children with reference to the spread of ability among them. A hint, possibly definite suggestions, may be given in teachers' meetings as the supervisor senses the readiness of his staff to receive them. In many ways the preparation may go on by means of a silent campaign at first, and then with definite effort in that direction, while, in the meantime, he must keep his hold on the loyalty and cooperation of the teaching staff.

C. Presenting the Program

1. In the Teacher-Training classes.

From the very first the supervisor may introduce measurements in the courses of teacher training for young people. They are in an attitude and frame of mind for whatever is offered in the course that they may prepare themselves the better for teaching. Morever, they have fewer preconceived ideas and prejudices to prevent them from seeing the practical uses of such a program in solving the problems of the school. Young people may enter into their work with this background and quickly acquire a scientific proficiency in teaching surpassing those of many years' experience, simply because they have had the advantages of superior training on this point. S.S. Brooks relates an experience[1] of just this sort in the public school field

---

1 - Brooks, S.S.-Improving Schools by Standardized Tests. Chap. I

when he was carrying out his experiments in measurements. The young prospective teacher training class is, therefore, the easiest and most logical place to begin an active presentation of a measuring program. The subject may be tactfully taught, also, in classes of older, active teachers in service.

3. Teachers' and Officers' Meetings.

At an appropriate time, when the supervisor has sensed the time is ripe for it, a thoroughgoing presentation of the program may be made. Care must be exercised in being assured the staff are unanimous, practically so, in wanting it presented. The actual presentation should be planned with great care and detail, with a speaker from the outside on the program if desired.

The preparation preceding the final presentation may be of nature of a series of meetings in which by practical demonstration the old methods of haphazard evaluation can be exposed as inadequate. It will be all the more effectively done if some teacher who is interested in the use of measurements can make a demonstration of the value of measurements with her own class. Other teachers may visit that class to watch the demonstration and hear the results reported, and thus there will be a number who have already seen the value of the program first hand and are ready to back up the supervisor in his presentation.

Whatever the preparation previous to the meeting when the final presentation is made, the program should, on this important occasion, be so clearly and definitely visualized that all doubts and misgivings will be dismissed,

and an enthusiastic demand for its adoption will come from the teachers themselves. Graphs, charts and pictures may be utilized with telling effect in making a visual presentation of the values of the program. Both the values and the limitations of measurements should be clearly set forth, and ample time given for questions and discussions. Perhaps one or more further meetings of the staff as a whole, and one or more with the department groups separately would be the practical outcome of the first meeting. This will afford time for digesting and assimilating the particulars, and understanding the procedures necessary for making the program effective. It may be that, even then, not more than one department will be ready for the program, or even one class. Time must be given to thoroughly plan the details of the program with the teachers and supervisory staff. Cooperation is necessary, and it takes time for some minds to get into action even after they have visualized the program, and have approved it.

3. Executive Board Meeting.

Meanwhile, the minds of the Board members are also being prepared for the program in a general way. No detailed preparation is needed here, as in the case of the teaching staff, but in general a certain amount of education may be carried on with this authoritative body. Doubtless some of its members will come from the teaching staff, and should be already growing enthusiastic over the program. Such members will form a valuable nucleus for educating the remaining members. Since the authority of this board must be behind

every move made by the supervisor, they are due a presentation of the program to which they are to give official approval.

D. The Active Administration of the Program

1. Careful consideration should be given the details of procedure as the program is being set up, attention being given to one Department at a time until the program is securely on its way. Some steps in the setting up of the program have been suggested by Barr and Burton:[1] (1) Giving the tests; (2) tabulating results; (3) interpreting results; and (4) prescribing and carrying out remedial treatment.

a. Directions for giving the tests, which generally accompany the tests, should be followed out very carefully in order to get results that are comparable to the results of other classes using the same test.

Teachers should acquire a certain skill in handling a testing program for themselves so that the supervisor may be relieved to give his attention elsewhere. It will likely be necessary for the supervisor to provide systematic training for a period of time in order that a number of the teachers may profit from the expert advice he may give. It is wise to encourage the teachers to adopt the use of tests as a routine part of their work, and some understanding of the meaning of their results should be acquired.

As to scoring the tests, the teachers may do this themselves or they may have the pupils do it while they read

---

1 - Barr and Burton-Supervision of Instruction, Chap.IX

out the correct answers themselves. This not only saves time for the teacher, but it has the added value of impressing the pupils with the correct answers just at the time when their minds are most alert to know those answers.

b. Tabulating Results.

The tabulation of results and the arrangement of them into distribution tables is a long process if carried out to the limit of statistical methods. The teacher should at least know how to handle the results in this way, though she should have others do most of the statistical work if possible, so that she will have her time free for instructional duties. Whoever makes up the tables should supply the teacher with the figures and tables which have been made, so that she may appreciate their significance.

c. Interpretation of results and the prescription of remedial measures.

These two steps are so closely related that they are treated together. This is the stage of the process in which the greatest value of tests are to be found. They are not primarily teaching devices, but revealers of trouble in the learning process of the child, and upon the facts revealed the teacher and supervisor will proceed with their policies of guidance in the learning procedures. Many unexpected facts will be revealed in the results of a testing project with a class or school, regarding their knowledge of the Bible and other teaching materials, and regarding the correlations obtained between the results in one direction and those in another. It has been found that there is very

little correlation, for example, between the knowledge of the Bible and ethical judgments. Evidently, the children who were tested in this case had not been taught the ethical content of the Bible as it should have been taught, or else it shows that the Bible alone is not suffident for the supplying of proper ethical content in religious education.

The needs of a class should be located definitely and in case of individuals who show extreme difficulty, the individual needs must be investigated, and the remedies prescribed and applied. Probable causes of difficulty should be listed and each class located under one or the other of these causes. It may be a combination of causes have united to produce a certain situation, and the diagnosing of each contributing factor may prove a complex task. Still the more difficult the task, the more there is need for working it out thoroughly and applying the proper remedies. It is at this point that the educational physician saves the spiritual life of his patients and starts them on the way to normal growth and wholesome character formation.

SUMMARY

Chapter VI is a practical adaptation of the theory of educational measurements to the actual work of the school in the local community. Three general steps were indicated in this adaptation:

A. Preparation for the Program

B. Presentation of the Program to the school authorities

C. The Administration of the Program

GENERAL SUMMARY

GENERAL SUMMARY

The development of thought in this discussion has been on the basis of the general principle of proceeding from the known to the unknown. Relatively speaking, the field covered in Part One is familiar to all. The known findings as to the use of measurements in that field have been taken over as a basis for the development of the arguments in Part Two, which is a relatively unknown field.

Part One.

The Supervision of Education and the Use of Objective Measurements

This broad field was treated in outline fashion under the general topics of:

I. The General nature of Supervision and its scientific basis. Here the attempt was to show, through a careful definition of supervision and by means of certain principles underlying this function, that a general scientific basis is demanded for effectively meeting the obligations of that office. The scientific method was then reviewed and outlined, and the question of a demand for objective measurements was raised.

II. Objective Measurements in Educational Supervision.

A definition of education, and of educational measurements revealed the quantitative nature of the educational "changes" to be measured. Our subject was then attacked from the standpoint of the school problems demanding more accurate measurements for their effective solution.

III. Objective Measuring Instruments, their uses and limitations. By means of a brief classification and description of the instruments in use, along with their values and limitations, a more practical approach to the topics of former discussions was attempted.

Part Two

Supervision of Religious Education Through Objective Measurements

Here we encountered our primary problem, but since the underlying discussions had already been taken up in Part One, our task was shortened, and we confined ourselves to the specific problems arising from the supervisory uses of measurements in this particular field.

I. The Scientific nature of Supervision of Religious Education. A definition of religious education in quantitative terms was found to be needed in order that objective measurements might be intelligently applied to its processes. Certain specific principles were then given, which led conclusively to the demand for a more scientific approach to the task of supervision.

II. Measurements and the Supervision of Religious Education. At this point, assuming all that had been said in Part One on this topic in relation to secular education, we proceeded to point out certain specific problems to be faced in the use of measurements in religious education. Specific capacities basic to religious growth were then analyzed as to their capability for being measured, and some of the instruments now available were listed.

III. Administering a Program of Measurements in a System of Religious Education.

As a final and practical attack on the entire subject, it was here undertaken to show how a program of measurements might be carried out in a local situation, and how to find some means by which the greatest good might be realized from the contributions which have been made in this field by the specialists. In this step alone lies the hope for making connections between the potentialities of research workers and their findings, and the everyday voluntary leaders who are in direct contact with the children for whom these findings were so laboriously worked out.

GENERAL CONCLUSION

The task of education has become exceedingly complex during the generation that has just past. The growing number of interests of life are making increasing demands upon the school, requiring radical changes in the curriculum and a more thorough-going system of management and supervision of its enlarging problems. Out of this complex situation, has emerged, by necessity, the separate office of supervisor whose task is to unify and systematize the entire educative process to the end that the younger generation may be trained in the ideals of American democracy and citizenship.

Religious education, so closely related to secular education, faces the same complex situation in the modern demands made upon its services to the nation and to the world. Both the conservation of energy, time and money, and the concentration of attention upon the essentials of religion call for a far more efficient organization of the work of religious education than has been achieved as yet. The following points sum up the demand for a more scientific and thorough-going approach to the problem.

I. The supervision of education, in both the secular and religious fields, is a most important task, demanding a highly specialized leadership which is able to attack its work with scientific accuracy.

II. A knowledge of the uses and the technique of objective measurements is an essential part of the supervisor's equipment, if he is to meet his obligations effectively.

It will be a long time before any ideal solution to the present problems is reached, but it is confidently hoped that an increasing number of religious leaders may be willing to attack their work in a systematic and scientific manner and that definite progress toward ideal conditions may soon be evident.

BIBLIOGRAPHY

Athearn, W.S. - The Indiana Survey, Vol. II, N.Y. Doran, 1924

Barr, A.S. and Burton, W.H., The Supervision of Instruction New York, Appleton, 1926

Betts, G.H. - How to Teach Religion

Bower, W.C.- Survey of Religious Education in the Local Church.

Brooks, S.S.- Improving Standardized Tests, Boston, New York, Houghton & Mifflin, 1922

_____ "Measuring the Efficiency of Teachers by the Use of Standardized Tests," Journal of Educational Research, Nov. 1921

Burton and Matthews
Burton, Ernest Dewitt, and Matthews, Shailer, Principles and Ideals for the Sunday School. New York, Appleton, 1926

Case, Adelaide, "Children's Ideas of God", Religious Education Magazine, Vol. XVI, June, 1921

Chassell, C.F. "Multiple Choice Test of Children's Ideas About God". Indiana Survey, Vol. II

Freeland, Geo. E.-The Improvement of Teaching, New York, MacMillan, 1924.

Gregory, C.A. Fundamentals of Educational Measurement, New York, Appleton, 1923.

Hanson, W,L. "Supervision of Religious Education Through Objective Tests and Measurements", Boston University Bulletin, Vol. XIII, July, 1924.

_____ "Boston University Revision of Giles Sunday School Examination A". Indiana Survey, Vol. II

Hartshorne, Hugh, "Measurement of Growth in Religion", Religious Education, June, 1919.

_____ "Cooperative Study of the Religious Life of Children", Religious Education, Vol. XVI, Dec. 1921

Judd, Charles H., Measuring the Work of Public Schools, 1921

Kelley, T.L. Statistical Method, New York, MacMillan Co. 1923

Laycock, S.R. "Laycock Test of Biblical Information", Journal of Educational Psychology, May, 1926

Lincoln, E.A. Beginnings in Educational Measurement

McCall, Wm.A. How to Measure in Education, New York, MacMillan 1923

Monroe, W.S. Theory of Educational Measurements, Boston, New York, Houghton and Mifflin, 1923.

Monroe, W.S., DeVoss, J.C. and Kelly, F.J.- Educational Tests and Measurements, New York, Houghton Mifflin Co. 1917.

Morrison, H.C. The Practice of Teaching in Secondary Schools, University of Chicago Press, 1926.

Nutt, H.W. Teaching High School Pupils, New York, Century Co. 1923.

Paterson, D.G. Preparation and Use of New Type Examinations New York World Book Co.

Pressey, S.L. and Pressey, L.C. Introduction to the Use of Standard Tests, New York, World Book Co. 1923

Strayer, Geo.D. and Englehardt, N.S. The Classroom Teacher

Russell, Chas. Classroom Tests, N.Y. Ginn and Co. 1926

Trabue, M.R. Measuring Results in Education

Voelker, P.F. The Function of Ideals and Attitudes in Social Education, New York, Teacher's College, 1921.

Watson, Goodwin B. Experimentation and Measurement in Religious Education. N.Y. Association Press, 1927

———— The Measurement of Fairmindedness. Teacher's College, Contributions to Education, No.176 N.Y. Columbia University, 1925.

Wood, Ben D. Measurement in Higher Education

Starch, Daniel, Educational Measurements, N.Y. MacMillan Co. 1923.