

Statement of Permission to Copy

In presenting this professional paper in partial fulfillment of the requirements for an advanced degree at Montana State University, I agree that the Library shall make it freely available for inspection. I further agree that permission for extensive copying of this professional paper for scholarly purposes may be granted by my major professor, or, in his absence by the Director of Libraries. It is understood that any copying or publication of this professional paper for financial gain shall not be allowed without any written permission.

Signature Brent Richard Pfeiffer
Date May 26, 1972

A MODEL STANDARDIZED TESTING PROGRAM FOR K-12 DEVELOPED FOR THE NORTH
PACIFIC UNION CONFERENCE OF SEVENTH-DAY ADVENTIST'S EDUCATION SYSTEM

by

BRENT RICHARD PFEIFFER

A professional paper submitted to the Graduate Faculty in
partial fulfillment of the requirements for the degree

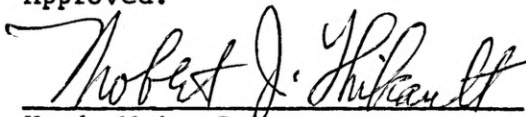
of

MASTER OF EDUCATION

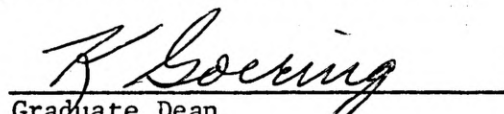
with concentration in

Counseling

Approved:


Head, Major Department


Chairman, Examining Committee


Graduate Dean

MONTANA STATE UNIVERSITY
Bozeman, Montana

June, 1972

TABLE OF CONTENTS

	Page
ABSTRACT.	v
 Chapter	
I. INTRODUCTION	1
Statement of the Problem	1
Need for the Study	1
General Questions to be Answered	2
General Procedures	2
Limitations.	3
Summary.	3
II. REVIEW OF LITERATURE	4
Standardized Scholastic Achievement Tests.	4
Achievement Testing: K-3	6
Achievement Testing: Grades 3-12	11
Standardized Intelligence Tests.	13
Tests for Mental Ability - Grades K-12	13
Standardized Vocational Aptitude Tests - Grades 9-12	13
Vocational Interest Tests.	15
III. INTRODUCTION.	18
Population Description	18
Method of Data Collected	18
Presentation of Data	19
IV. A PROPOSED STANDARDIZED TESTING PROGRAM.	25
Introduction.	25
Rationale for a Standardized Testing Program	25
Proposed Tests and Schedules	26
Review of Recommended Tests.	27
Scoring, Recording and Dissemination of Test Results	32
Summary.	34
V. CONCLUSIONS AND RECOMMENDATIONS.	36

	Page
Conclusions.	36
Recommendations.	36
SELECTED REFERENCES.	38
APPENDIX	41
Appendix A	42
Appendix B	44
Appendix C	45
Appendix D	47
Appendix E	48
Appendix F	49

ABSTRACT

The purpose of this study was to recommend a model standardized testing program for grades K-12 in Seventh-day Adventist operated schools in the North Pacific Union Conference. This conference includes Washington, Oregon, Idaho and Montana.

To obtain information to complete this study questionnaires were sent to the entire population of elementary schools (102) and secondary schools (9) to determine teachers' and counselors' feelings regarding standardized testing programs. A review of literature pertaining to standardized testing and testing programs was also done. The questionnaire and the review provided the foundation upon which to build the recommended model testing program. The questionnaire sought the following information:

1. What are the components of a satisfactory standardized testing program?
2. What do you view as the objectives of a standardized testing program?
3. What areas should be covered by testing?
4. Is there a need for a unified testing program?

Results of the questionnaire are as follows:

1. Components--a standardized testing program should yield prompt results that are easily understood by teachers, students and parents.
2. Objectives--a standardized testing program should
 - (a) indicate student progress over a period of time,
 - (b) indicate student and teacher strengths and weaknesses,
 - (c) assist teachers in overall planning,
 - (d) provide for educational and vocational guidance.
3. Areas covered--a standardized testing program should cover
 - (a) subject area achievement,
 - (b) subject area aptitude,
 - (c) mechanical aptitude,
 - (d) interest inventories,
 - (e) personality inventories,
 - (f) vocational aptitudes.
4. Unified program--sixty-three percent of the elementary school respondents and 100% of the secondary school respondents favored a unified testing program.

It was found that the results of the questionnaire for the most part agreed with literature reviewed.

The investigator built his recommended standardized testing program around these results and information obtained from his review of literature. The investigator realizes that any one testing program will not satisfy all people but feels that this recommended standardized testing program can be used as a solid framework on which to build a testing program as other specific needs and objectives are located.

Chapter I

INTRODUCTION

The administering of standardized tests is a part of practically any school system. How the testing program is conceived and how it is perceived has much to do with whether it is a very real part of the school system and whether or not it will be an effective program. If the testing program "just grewed," then it has a good chance of being of low quality and low effectiveness. However, if some serious planning is done before the program is born, then it has a much better chance of surviving to enhance each component of the school system, i.e. students, teachers, parents, administrators. Anyone making use of an established standardized testing program certainly should expect to derive a very positive benefit from it.

Statement of the Problem

The purpose of this study will be to construct a model standardized testing program for kindergarten through high school for the North Pacific Union Conference of Seventh-day Adventists' educational system.

Need for the Study

There is at present a great deal of expressed dissatisfaction with the standardized achievement test battery often used. There exists a definite lack of unification in the secondary level testing program.

In view of these two shortcomings it is felt that a hard look at the testing program is in order and that recommendations are needed to make it a more meaningful standardized testing program.

General Questions to be Answered

This investigation will attempt to provide answers for the following questions.

1. What is the general feeling of current authors and researchers regarding standardized testing in the areas of achievement, intelligence, aptitude and interest?
2. What are the complaints regarding testing in the individual system involved in the investigation?
3. What standardized tests currently available would best suit the purposes of the educational system involved in this investigation?

General Procedures

The investigator will do four things to develop this model program and to seek answers to the previously stated questions.

First, a review of literature dealing with standardized tests, and standardized testing programs will be conducted. Second, by using a questionnaire, opinions of test administrators in the educational system will be sought. Third, by use of Buros' Mental Measurement

Yearbook and other studies and sources, specific tests will be recommended for this model program. Fourth, a complete standardized testing program will be recommended.

Limitations

This paper will deal only with standardized tests in the areas of achievement, intelligence, vocational aptitude and vocational interests. It will not consider personality tests or individually administered intelligence or achievement tests.

Summary

The investigator will bring together existing information regarding the establishing of a standardized testing program and standardized tests and attempt to establish a standardized testing program which will meet the needs of the previously mentioned school system.

Chapter II

REVIEW OF LITERATURE

Every student within an educational system has the right to expect as much assistance as he needs from that educational system to become prepared to live and lead a life acceptable to both himself and to society. One of the ways educational systems have planned to achieve this goal is by establishing standardized testing programs which will enable teachers to be of greater assistance to this student and which will enable the student to better understand himself. This chapter will focus on standardized tests in the areas of scholastic achievement, intelligence, vocational aptitude and vocational interest. It will present a review of current literature dealing with these various areas and the recommendations made in this literature.

Standardized Scholastic Achievement Tests

Standardized achievement tests were the result of significant advance in statistical calculation and scientific measurement of a person's abilities that were developed in psychological laboratories. Their acceptance by the teacher as a guide of instructional improvement and by the public in general as a valid undertaking allowed them to continue being refined until these scholastic achievement tests have become sophisticated to a degree beyond that considered by their original designers.

Standardized achievement tests today fall into three types.

The survey test batteries are comprehensive and are the mainstays of any school testing program. The specific subject tests are limited to a detailed coverage of specific topics and can be used to supplement teacher-made tests or to gain further detailed information about student strengths or weaknesses indicated in the survey test battery. Diagnostic achievement tests provide exercises and problems to maximize the possibility of error in order to point out the specific areas which are weak. It is not advised to use subtest scores of achievement batteries because "...in most instances the items were not constructed to measure the components involved in a particular skill...." (Mehrens and Lehman, 141, 1965).

Chauncey and Dobbin (1963) paraphrased the general conclusions of various studies conducted in 1920's, 1930's, and 1940's and these conclusions will serve as the philosophy behind the existence of standardized achievement tests.

The purpose of schools is to educate each young person in such a way that he will be able, to the limits of his natural ability, to contribute to the welfare and strength of American society and to realize a full development of his own potential as a human being. In order to do this, the school teaches each child certain subject matter content, such as English, history and mathematics, so that he may have the knowledge with which to solve his problems; the learning of this content is the immediate goal of instruction. When he has learned some subject matter and mastered some skills, the school teaches the student to apply them in a variety of new situations, so that he will be able to use them when they are needed in nontextbook circumstances; learning to use school-learned knowledge in a variety of ways is

the intermediate goal of instruction. When he knows the subject matter and how to apply it, the ultimate goal of education is that he will apply his learning, bettering his own life and the lives of his fellows (Chauncey, 1963, 15).

The immediate and intermediate goals are now testable but it is still not possible to test the ultimate goal.

Achievement Testing: K-3

This group of students falls into a special class when it comes to being administered standardized tests. The majority of authors feel that standardized tests must be limited in use with this group for the following reasons: Skills learned by this group are viewed as being additive rather than independent, teachers are more intimately involved with pupils achievement, pupils are unable to handle separate answer sheets thus raising the cost of testing, test scores may be highly reliable at the time of testing but their long-term reliability and validity leave much to be desired (Baurenfeind, 1963).

These limitations however do not mean that there is no place for achievement testing at this level. Tests having a place at this level of education are: reading readiness tests which assist in administrative decisions for entrance or non-entrance and for the formation of reading groups and reading and arithmetic achievement tests which

...provide perspective on the teacher grades, assist in placement at the next grade level, ...helps in identifying strengths and weaknesses in the school program (Baurenfeind).

Since reading and arithmetic are the two major learning areas

in these early grades and since there are limitations in currently available tests to measure general achievement accurately and since testing which serves no purpose and is expensive, general achievement testing should be avoided. That is unless a school wishes to administer a battery on an experimental basis.

Achievement Testing - Grades 3-12

It is to this group of students that achievement testing applies and has the most validity, however, some authors will debate the two extreme ends. The lower end because of the previously mentioned limitations and the upper end because scores are usually of little use to anyone. Generally speaking these instruments are not dependent upon specific courses, textbooks or type of construction but provide a broad measure of skills and understanding.

It is generally felt that there are real values in giving educational achievement tests in the school program. Baurenfeind lists eight major values and possible uses of achievement tests, "a list," as he points out, "not equaled by any other type of measuring device" (Bauernfeind, 1963, 136).

1.have immediate interest and relevance for all classroom teachers.
2.usually have immediate interest and relevance for the student.
3.have immediate interest for parents.

4.have immediate interest and relevance for school administrators.

5.have immediate interest and relevance for curriculum specialist.

6.substantially overlap measures of general abilities but in a relatively non-threatening and action oriented way.

7.often afford substantial predictive validities with respect to future educational attainments.

8.often prove useful in vocational counseling at high school level (Bauernfeind, 1963, 136-141).

The case for using standardized achievement tests is well settled and these tests are used widely by a great percentage of schools at all levels.

Anne Anastasi lists the following uses for achievement tests:

1.attainment of minimum performance standards
2.selection (hiring)
3.placement and classification
4.counseling
5.remedial teaching program
6.evaluation of teaching
7.improvement of instructional techniques (Anastasi 1961, 227).

An even more exhaustive listing of major functions of achievement testing is given by Payne who lists fifteen functions falling into the general categories of:

1. Directing curriculum emphasis
2. Providing for educational guidance of pupils
3. Stimulating the learning activities of pupils
4. Directing and motivating supervisory and administrative efforts (Payne, 1968).

While the standardized achievement test battery is able to accomplish the above the teacher grades cannot hope to make the same assessment for several reasons. Measures are obtained for the student only in areas in which he is taking course, grades are too restricted in meaning and teacher grades often include more than pupil achievement and aptitude for a given course.

When establishing an achievement testing program it must be decided whether the program is to provide continuity from grades three or four to grade twelve, enabling the school system to follow the students educational progress throughout the elementary and secondary grades or whether there is enough difference in the purposes of the elementary and secondary schools to require a different kind of achievement battery at each of the two levels. An example of an achievement testing program following the first decision would be STEP program, the California achievement program or the Metropolitan achievement program. If it is felt there is a difference between the purposes of the two levels the Iowa Test of Basic Skills (grades 3-9) and the Iowa Test of Educational Development (grades 9-12) might serve to measure these

differences.

Is the purpose of the test for diagnosis of student weaknesses or evaluation of student progress? Deciding between these two purposes largely determines when the tests will be administered. If the test scores are to be diagnostic, early fall administration is best. If the test scores are to be evaluative, late spring administration is best.

Some testing experts feel combination diagnostic-evaluative results can be obtained by testing in late fall. Late fall testing allows planning for second semester work, and allows the teacher to better judge the accuracy and implications of the test because they are more familiar with the student.

Once achievement test scores have been obtained it must be decided how they will be viewed. All tests come with some kind of national norms but a student who scores "average" according to national norms may not be "average" if his school or school district established their own norms. National norms may be far too low, or high, for the socio-economic status of the pupil or his school. It is therefore imperative that national norms be considered in how they relate to the school and the community in which the pupils exist (Baurenfeind, 1963 Mehren and Lehmann, 1969). Karmel and Thorndike and Hagen in their books on standardized testing all recommend establishing of local norms in the form of percentiles, percentile bands, stanines or grade placement, whichever the school or district feels is most applicable to their

situation in order to make the test scores more realistic.

Whatever choices are made regarding the administering of achievement test it is clear that the case for using standardized achievement tests is well settled and they will continue to be used and, unfortunately, misused in American education.

Standardized Intelligence Tests

With some difficulty intelligence tests can be divided into three classifications: (1) Achievement loaded tests, (2) Achievement free tests and (3) Wide spectrum tests. If a school has an existing achievement testing program little is to be gained by administering an achievement loaded intelligence test. They may wish to supplement their achievement testing program with an achievement-free test "...if they have facilities or the time to work with individual children who might then be legitimately identical as 'under-achievers'" (Sipkens and Brown, 1972).

The wide spectrum tests do not give a complete measure of any phase of learning but only a partial one and therefore their use is discouraged.

Cronbach has established criteria for comparing test of scholastic aptitude. He lists tests on a continuum from maximum educational loading to minimum educational loading. "The functions of the tests at the two ends of the spectrum are different. Those toward the top

are designed for cold-blooded prediction of future school success..." while those at the bottom are designed "...to locate undeveloped potential that novel treatment may bring out" (Cronbach, 1949, 234-36).

Intelligence tests are administered today mainly to estimate the capacity of the student for school learning and are based upon the intellectual skills that the student has acquired over a long period of time and from many sources such as home and community as well as school. But in many circles they are unable to live down their tainted past and are viewed with suspicion.

Orville Brim, Jr. reports that 25-50% of adults are against using intelligence tests for selective purposes and that 53-62% of high school students are opposed to intelligence testing. Neulinger, attempting to measure similar feelings, however, concluded from his study that attitudes toward fairness showed a difference of opinion, that anti-test sentiment is neither ubiquitous or consistent and that inconsistencies are determined by the respondents social background and personality characteristics.

This inconsistent feeling carries over to those writers who are concerned with school standardized testing programs. Louis Karmel feels that they are of utmost importance while Bauernfeind states that they merely duplicate information gained by a comprehensive achievement test battery. All writers do agree, however, that intelligence test scores are valid only if they are interpreted correctly by the people involved

in interpreting these test scores and that they are dangerous and indeed invalid if interpreted incorrectly.

Tests for Mental Ability - Grades K-12

Research indicates that intelligence test results are no more reliable in the early school years, K-3, than are achievement test scores. Louis Karmel, in his book Measurement and Evaluation in the School, states "a single measure at this level (7-12) is more likely to represent the true ability of the student than a single score obtained at the elementary level" (Karmel, 435). He adds that if finances are a problem of your testing program a single test administered in the eighth or ninth grade is most appropriate, however, if finances are no problem an intelligence test at the beginning of the seventh, end of the eighth and beginning of the tenth is recommended. No rationale for this test schedule is given.

It would seem that since intelligence or mental ability tests enjoy a rather tenuous position at best in the school testing program that the intelligence test chosen should be one that does not overlap to a great extent achievement tests that are administered, should be one that is easily understood and interpreted by the teacher to minimize incorrect and therefore dangerous interpretations.

Standardized Vocational Aptitude Tests - Grades 9-12

Aptitude tests were originally devised to measure special

abilities and enable psychologists to predict vocational success for an individual, thus giving the tests a practical purpose. If these aptitude tests could discover unusual talents that may not have been identified in the school achievement tests and obtain information that would be helpful in educational and vocational counseling then they should by all means be part of the school testing program. As a result of the attempt to accomplish the above mentioned tasks and a result of Thurstones' defining eight fairly independent ability areas, multiple-aptitude test batteries have been developed. "It was hoped after Thurstones' early work ...that the breakdown of intellect into independent measures would help us to predict different types of criteria" (Bauernfeind, 193, 1963). This has not been found to be the case. It has also been found that multiple aptitude tests do not measure success in school as well as general achievement test batteries. Then too, multiple aptitude tests are of uncertain value in predicting vocational success. Unfortunately this last complex puzzle has not yet been solved.

Why then should a school testing program bother with a multiple-aptitude test if it has these shortcomings? The following reasons might be listed: (1) They give a more complete picture of the individual, (2) they may indicate where remedial programs would be of benefit, (3) future educational or vocational plans may, with caution, be built around the students strengths and weaknesses, (4) the counselor may use the individuals test profile to initiate a rewarding counseling

session (Bauernfeind, 1963).

It is not until the junior high school years that students are beginning to consider career decisions and plans for further education because abilities and interests are becoming more differentiated with maturity and experience. In light of this the multiple-aptitude batteries are not recommended prior to junior high school and often not until the last year of junior high, (the ninth grade). Some authors recommend, depending on the geographical area, that these test batteries not be administered until just prior to the drop-out age in hopes that it may encourage some potential drop-outs to continue their education.

The multiple-aptitude test has a part in the school testing program because of its ability to add more definition to the individual. Even in spite of its shortcomings, if it can do this, it belongs.

Vocational Interest Tests

Information regarding a person's interests may be obtained in a number of ways. His interests may be "expressed interests (verbalized to the counselor), manifested interests, tested interests and inventorial interests..." and "...be applicable to the measurement of the eight basic interest groups defined by Super and Crites (1962): scientific, social welfare, literary, material, systematic, contact, aesthetic expression and aesthetic interpretations" (Aiken, 210, 1971)

Interest inventories were developed to indicate to the student,

What am I like, and are an obvious companion to the vocational aptitude test. These inventories were developed along two different lines, the forced-choice and free-response. There are arguments for and against the use of both. In the forced-choice inventory the person has to choose between two or three possible items, any one, none or all of which may be interests. It is argued that this type of inventory is more difficult to fake and that it parallels real-life choices more closely because a person cannot do all the things he wants to do. The forced-choice item inventory can prove to be frustrating to the student because it does not allow him to tell to what degree he likes or dislikes the selected item. The free-response item inventory may: (1) enable the counselor to come to know the student better; (2) encourage the student to begin exploring career possibilities because of the enthusiasm the student may have. Neither type of test is able to predict levels of satisfaction in certain courses or careers nor to predict the level of success in certain courses or careers (Bauernfeind, 1963). Two cautions to the counselor are that: (1) the inventories yield scores that are highly reliable at the time of administration but these same scores may not be relevant over a period of time and (2) that the results indicate areas of interest only and should no way be interpreted as saying the student would do well in these interest areas. There are too many other variables involved to make this kind of interpretation.

It can be seen from this chapter that there are many reasons for testing. Achievement tests have been the backbone of all testing programs with other types of tests added to them to cover as wide a scope as possible. Intelligence testing continues to have a cloud over its respectability but can add to a testing program if the intelligence test is chosen wisely. Vocational aptitude and interest test continue to provide viable information for student, teacher and counselor if their results are correctly understood and used.

Chapter III

INTRODUCTION

In this investigation information will be sought which will adequately answer the questions posed in Chapter one and which will provide information which will serve to construct a testing program from grades K-12.

This chapter will include a description of the population receiving questionnaires, the method of data collection, and an analysis of the data collected which will include information and recommendations obtained from the returned questionnaires.

Population Description

The population will include all principals of elementary schools and all counselors in secondary schools in the North Pacific Union Conference of Seventh-day Adventists. This union conference consists of 102 elementary schools and nine secondary schools distributed throughout the states of Washington, Oregon, Idaho and Montana. The names and addresses of these administrators and counselors will be obtained from the Superintendent of Adventist Education in Montana.

Method of Data Collection

A questionnaire (see appendix) designed by the investigator will be sent to the entire population. They will be given a specified time in which to return the questionnaire. A cover letter (see appendix)

will accompany the questionnaire explaining the purpose of the study.

The questionnaire attempts to measure the satisfaction or dissatisfaction of the population regarding the existing standardized testing program and their recommendations of what should be included in a standardized testing program in order to make it an acceptable program. The investigator feels that this information will be valuable in building a standardized testing program that will be accepted by those who must use it and the results obtained from it.

All information received from each respondent is assumed to be accurate and specific data that is reported will be reported on a descriptive basis.

Presentation of Data

Seventy-one percent of the population responded to the questionnaire. The following information was taken from those questionnaires. The percentage figures are based on the total number of returned questionnaires and unless specifically designated elementary and secondary figures are combined.

1. Do you feel your schools standardized testing program yields satisfactory information to benefit the:

	<u>Yes</u>	<u>No</u>
a. student	53%	37%
b. teacher	65%	25%
c. administrator	54%	23%
d. parents	65%	26%

The next section asked those who responded no to any portion of question one to state how this might be improved. It was found that suggestions could be grouped in five areas. The number in parentheses indicate the percentage making the suggestion.

1. Results need to be returned more promptly. (35%)
2. Students should be able to understand reports. (7%)
3. Parents should be able to understand reports. (16%)
4. Teachers should be able to understand reports. (5%)
5. Teachers should score the tests to insure that they see the results. (3%)

It was reported that on the elementary level 80% of the classroom teachers were called upon to explain test results. Five percent reported the administrator explained test results while 3 1/2% reported no one explained test results. On the secondary level 55% reported counselors explained test results, 27% reported this task fell to the administrator and 18% reported test explanations to be the task of the classroom teacher.

The following were most frequently listed as being important objectives of the school's standardized testing program. Again, the number in parentheses is the percentage of respondents naming this item.

1. To indicate students weaknesses or strengths. (66%)
2. To indicate student progress over a given period of time. (21%)

3. To indicate teacher weaknesses or strengths. (19%)
4. To help teachers in overall planning. (12%)
5. To evaluate the school. (12%)
6. To provide guidance for the student, either educational or vocational. (14%)
7. To rank students in comparison with other students. (7%)

Five percent reported that they didn't know what would be considered test objectives.

When asked to indicate what tests, if any, could be dropped from the testing program, five percent felt all tests could be done away with, one respondent desired to drop the SCAT-STEP series, the rest felt none of the existing tests should be dropped unless they were replaced.

It was felt that there were nine areas not adequately covered by the school testing program. These were:

1) Subject area achievement, science, social studies and reading were mentioned specifically, 2) algebra aptitude, 3) subject area aptitude, 4) mechanical aptitude, 5) interests inventories, 6) personality inventories, 7) vocational aptitude.

Fifty percent reported that there were no areas that needed to be covered more adequately.

Sixty-three percent of the respondents favor a unified standardized testing program, 5% indicated they didn't favor a unified program,

20% did not respond to this question. One hundred percent of the secondary respondents favored a unified testing program.

From studying this information it appears there are numerous inferences one could make regarding testing. This investigator will make only those inferences that he feels will apply to the standardized testing program to be recommended later in this paper.

It appears that the existing testing program is satisfactory to a majority of the people using it. The only major criticism of it was not with the tests themselves but the delay in attaining tests results when the tests were machine scored. If this delay cannot be overcome either by hand scoring or improvement of the machine scoring service the test program will continue to be severely weakened.

Since a large majority of the classroom teachers, especially on the elementary level, are called upon to interpret the test results to themselves, the students and the students' parents, it would seem that the criticism of not being able to understand the reports or to explain them might be considered. This could be remedied by in-service training sessions for the teachers, simplified test results and explanations for the student and parents. The simplification of test results and explanations could well be used on the secondary level as well.

There seems to be a real lack of understanding when it comes to testing objectives. Two-thirds of the respondents said the major

objective was to indicate student weaknesses and/or strengths but other mentioned objectives received support from one-fifth or less of the respondents.

Supplemental areas recommended to be covered mostly applied to the elementary grades. While tests in the first three areas could conceivably be added tests in the latter three areas have not proved to be realistic for the elementary level.

It can be seen then that while the elementary standardized testing program is unified, all schools administer the CTBS and CTMM, it is not satisfactorily serving the people it is intended to serve. On the secondary level each school has gone its own way in developing its testing program which may make it satisfactory to the individual school but provides for little unification.

Chapter IV

A PROPOSED STANDARDIZED TESTING PROGRAM

Introduction

It is very evident to the investigator, especially after doing research for this paper, talking to others interested in testing and analyzing results of the questionnaire, that there is no one testing program that will please all people or answer all questions. It is felt however that the program below would be of value to the Seventh-day Adventist elementary and secondary schools in the North Pacific Union Conference if the test results are interpreted as intended by their authors.

Rationale for a Standardized Testing Program

Any standardized testing program must have a raison' d'etre otherwise it shouldn't exist. Karmel (1970, 434) states that standardized tests are to serve three basic purposes: (1) to provide a diagnosis of the readiness each child displays for engaging in future educational activities, (2) to provide a periodic measure of pupil progress toward the educational objectives of the school, (3) to assist in determining occupational, vocational, technical, professional goals and programs for students, including those whose abilities, achievement, motivation, background and self-concept do not lend themselves to traditional objectives.

At its fullest development a standardized testing program may consist of the basic testing program which is made up of the tests to be administered, the dates they are to be administered, guidelines as to how they are to be administered, provision for immediate and accurate scoring, reporting and recording in a useful manner; the supplemental testing program which provides for optional group testing for additional information; and a special testing and assessment program which consists of individualized testing for diagnostic purposes in the areas of academics, vocation, intelligence, personality, etc. (Karmel, 1970) The testing program to be recommended in this investigation is to be a basic one as it is felt the supplemental and special testing programs are ones which need to be more fully adjusted to local school needs. Also special testing programs need specially qualified people to administer and interpret the tests and the elementary schools in this school system do not have access to these specialists.

A good testing program is one in which there is a close alignment between the test materials and the testing purpose (Bauernfeind, 1963:8). It is felt by the investigator that the following recommended testing program follows this advice.

Proposed Tests and Schedules.

ELEMENTARY PROGRAM

Grade Level	Test*	Purpose	When Administered
K or pre school	Murphy-Durrell Reading Readiness (1) +	First-Grade Readiness	May or June before entry of 1st grade
3	C.T.B.S.-Arithmetic Level 1 (2) SRA Achievement: Reading Subtest (3)	Check on these two basic areas	September
4-6	Science Research Assoc. Achievement Series (4)	Check on Achievement	September
7	Lorge-Thorndike Intelligence Test (5)	Check on Achievement	September
8	SRA Achievement Series	Achievement Survey More valid for entry to Academy	April

SECONDARY PROGRAM

9	Differential Aptitude Test (6)	Class ability Study General Aptitude	September - required (to Academy students only)
	California Occupa- tional Preference Survey (7)	Vocational Interest	September (Required)
	Orleans Hanna Algebra Prognosis Test (8)	Math Placement	September (Required)
	Stanford Achieve- ment Test High School Reading (9)	Diagnostic	September (Required)

SECONDARY PROGRAM continued.....

Grade Level	Test	Purpose	When Administered
10	D.A.T.	New Students only	September (Required)
11	Washington Pre-College	College Exper- ience	April (Required)
12	Flanagan Aptitude Classification test (11)	Assist non- college bound student	November (optional)
	Strong Vocational Interest Blank (12)	Identify Special Job Interests	September (Required)

*Tests at all elementary levels are required

*See review of recommended tests.

Review of Recommended Tests

In the statements below regarding the proposed tests the investigator has presented the strengths of the tests as reviewed in Buros Seventh Mental Measurements Yearbook. The investigator feels that these points of strength are valid reasons why these tests should be part of the proposed standardized testing program.

1. The Murphy-Durrell Reading Readiness Test was selected because this is a well constructed and well standardized test. "The

Phonemes and Letter Names Test possess adequate reliability. The Learning Rate test is well designed but reliability and validity coefficients do not recommend its use for other than diagnostic purposes (Buros, 1972:1178)."

Scores: sound recognition, letter names (small and capital, total), learning rate, total.

Time: 60 minutes (2 sessions)

Cost: \$6.20 for 35 tests

Publisher: Harcourt Brace and Jovanovich

2, Comprehensive Test of Basic Skills - Arithmetic: Level 1

"The items of this arithmetic test are representative of a middle-of-the-road new math curriculum. Test 6 is well enough normed to provide some useful data on achievement of skills in basic operations (Buros, 1972:912)."

Scores: computation, concepts, applications, total

Time: 95 minutes

Cost: \$5.35 for 35 tests

Publisher: McGraw-Hill

3. Science Research Associates, Achievement Series: Reading

Subtest. "Usability of this test is substantial. Teachers or administrator without special training can easily administer and interpret the test (Buros, 1972:1098)."

Scores: Comprehension, vocabulary, total

Time: 120 minutes

Cost: \$15.00 for 35 tests

Publisher: Houghton Mifflin

6. Differential Aptitude Test

Reviews in the past have justifiably dubbed the D.A.T. as the best available instrument of its kind. The effectiveness of the D.A.T. is enhanced by such supportive data as given in Counseling in Profiles and the recurring pedagogic emphasis on the technical manual. It so far has not lived up to the promise of a differential instrument but no other instrument for the given age range has yet succeeded in this respect (Buros, 1972:1049-52).

The D.A.T. can serve several purposes in the ninth grade. (1) It can serve as an early vocational counseling guide, (2) the VR + NA scores can be compared to the abstract reasoning score if one wants to conduct another search for underachievers, (3) the DAT scores can help predict success in nonacademic courses such as shop (Sipkens and Brown, 1972:38).

Scores: Book I: verbal reasoning, numerical ability, total, abstract reasoning, clerical speed and accuracy.
Book II: mechanical reasoning, space relations, spelling and grammar.

Time: Book I - 120 minutes
Book II - 115 minutes

Cost: Booklet I - \$8.80 for 25 tests
Booklet II - \$8.80 for 25 tests

Publisher: Psychological Corporation

7. California Occupational Preference Survey

...the instrument will be useful to students who desire to learn about their own interests. Construction of the scales was carried out completely; coverage of both professions and skills is excellent; and the test and manual are set up so as to be highly instructive to the student. It is an honest questionnaire with straight forward items (Buros 1972: 12-12-13).

It is also a free-response test which may be more useful to some than the forced-choice interest test and may prove an interesting comparison for the student and counselor.

Scores: agriculture, home economics, literature and journalism, fine arts, social science, physical science, biological science, foreign language, business administration, accounting, teaching, civil, mechanical and electrical engineering, law, total.

Time: 30 minutes

Cost: \$20.00 for 25 tests

Publisher: Educational and Industrial Testing Services

8. Orleans-Hanna Algebra Prognosis Test

In general this appears to be an excellent instrument. Major new features include excellent supporting documents for machine scoring, a very good and readable manual, some "new math" items and a questionnaire section (Buros, 1972:510).

Time: 55-60 minutes

Cost: \$8.70 for 35 tests

Publisher: Harcourt-Brace and Jovanovich

9. Stanford Achievement Test: High School Reading

This appears to be a very satisfactory test of reading comprehension for high school students. In spite of a few limitations and with the understanding that this test... measures only a specific type of reading achievement this test ranks as one of the very best survey tests of secondary school reading achievement available today.

It adequately serves its objectives as a rough measure of reading achievement for comparative purposes and as an instrument of identification upon which further evaluation of reading strengths and weaknesses may be based (Buros, 1972:70 -9).

This test should be administered when remedial reading is available to the students.

Time: 55 minutes

Cost: \$8.20 for 35 tests

Publisher: Harcourt-Brace and Jovanovich

10. Washington Pre-college Test

This test is required for admission to any Washington college or university. It predicts student achievement in major areas in the college curriculum and has proven to be a reliable instrument.

Scores: English usage, spelling, reading, mechanical reasoning, spatial ability, applied mathematics, vocabulary, quantitative skills, mathematics achievement, vocational interest inventory.

Time: 320 minutes in two sessions

Cost \$7.00 per student

11. Flanagan Aptitude Classification Test

While being difficult to administer there are several features that recommend this test to the counselor for use in vocational counseling. Foremost among these is the continued research aimed at validating the claim of importance of matching job element with aptitude clusters. Secondly, the quality of the testing materials and the aids to counselors are superior (Buros, 1972:1054-55).

Scores: Inspection, coding, memory, precision assembly, scales, coordination, judgment and comprehension, arithmetic patterns, components, tables, mechanics, expression, reasoning, ingenuity.

Time: 388 minutes in two sessions

Cost: \$5.10 for 25 tests

Publisher: Educational Testing Service.

12. Strong Vocational Interest Blank

The new revision has been handled with circumspection....He [Campbell] seems to be trying to preserve the structure of the instrument that Strong built....at the same time he has modernized and streamlined the instrument, has made it more adaptable and has laid the foundation for more extensive innovations in the future (Buros, 1972:1462-63).

Scores: 22 basic interest, 55 occupational, 8 non-occupational and 6 administrative

Time: 30 to 60 minutes

Cost: \$6.00 for 25 tests

Publisher: Stanford University Press

It should be noted that the costs listed include only the test manual not the answer sheet, scoring keys, manuals for interpretation, etc. Also the time given for each test is the amount of time for administration not working time.

The investigator realizes that many different times for administration of these tests might be established but it was felt that the major testing objective was for discovering the strengths and weaknesses of the student and it was felt that this should be done as soon in the school year as possible.

Scoring, Recording and Dissemination of Test Results

With the advent of machine scoring of tests it has become relatively easy for the classroom teacher. Where at all feasible, machine scoring is recommended if the following criteria can be met:

1. Rapid and accurate scoring and return of test results are a must if the test program is to be usable.

2. Listing of local as well as national standings of each student tested is a much more benefit to the teacher. Student and school then merely comparing the student to the national figures.

3. Easily understood results are a must if the teachers are expected to use them.

Even if machine sorting is not possible, the above criteria are still important and must be a part of the scoring demands. If the classroom teacher doesn't have the time to hand-score the tests provisions should be made for extra secretarial help in doing so. The number of extra persons needed would of course depend on the size of the tested population and how fast the test results were desired.

Not only are fast, accurate, easily understood results important to a testing program, but so are the kinds of scores. Bauernfeind (1963), Karmel (1970) and Mehrens and Lehman (1971) all recommend that scores based on local norms as well as national norms are important. While national percentiles or stanines or grade levels may be interesting, they may not be accurate or appropriate for the school or school system. Local percentiles and stanines are most appropriate for conveying test results to students and parents as these have more meaning.

Recording test scores so that they are usable is an important part of any standardized testing program. These test scores should

appear on one sheet (card stock) and include the student's name and other personal information and have an appropriately designated place for recording the name of the test, the level of the test, the date given and the students score in percentile or stanine form as compared to national as well as local norms, (see appendix). This record sheet for each student should be kept in a place where each teacher, counselor, and/or administrator will have access to it as they need it.

When it comes to dissemination of test results to student and parent, the teacher or counselor must take utmost care to make certain the results are readily and easily understood. If possible a written explanation as well as a conference might be used to make the test results more meaningful. Two of the major criticisms of standardized testing has been either the unaccessibility of the test results to parents and students and the uncomprehensible way results are presented (Brim, 1965). (For an example of a written explanation of the Lorge-Thorndike Intelligence test see the appendix).

Summary

This chapter has presented the rationale behind a standardized testing program, the recommended standardized tests to be administered and the proposed methods for scoring, recording and dissemination of results. It is recognized that there are many variations that could be proposed and arguments presented for and against the proposed program.

It is felt by the investigator that this program provides a solid framework on which to build a testing program as other specific needs and objectives are located.

Chapter V

CONCLUSIONS AND RECOMMENDATIONS

Conclusions

The problem of this paper was to propose a standardized testing program for the elementary and secondary schools of the North Pacific Union of Seventh-day Adventists. Results from the questionnaire led the investigator to make some conclusions as to how educators in this school system felt about standardized testing in general and about the existing standardized testing program. Information gleaned from reading the ideas of testing experts and researchers also assisted him in reaching some important decisions as far as standardized tests and standardized testing programs are concerned. These conclusions and decisions helped the investigator propose a complete standardized testing program.

Recommendations

The recommendations of the investigator are ones that he is certain the reader could arrive at from studying the previous pages.

These recommendations are:

1. That a unified standardized testing program be put into operation on both the elementary and secondary levels in the North Pacific Union Conference.

2. That the proposed tests and testing schedule found in this paper be accepted as a foundation for the standardized testing program.

3. That a uniform system of record-keeping of test scores be adopted.

4. That this record of test scores follow the student through school and be kept in the permanent records of the school he last attends.

5. That there be consideration given to making qualified test personnel available to the elementary schools to administer individual intelligence, aptitude and personality tests, and

6. That an in-service training program be undertaken to assist teachers and administrators in administering and using test results.

Summary

The investigator feels very strongly that a standardized testing program can be of a very positive benefit to the school system. He also feels very strongly that in order for it to be a benefit, its users must see the need for it and put it to use. If this cannot be done, if it cannot be put to use to benefit all concerned with it then it serves no purpose and would be better off scrapped. It is hoped that this will not be done.

SELECTED REFERENCES

SELECTED REFERENCES

- Aiken, Lewis R. Jr. Psychological and Educational Testing. Allyn and Bacon, Inc. Boston: 1971.
- Anastasi, Anne. Psychological Testing. Macmillan Co. New York: 1961.
- Bauernfeind, Robert H. Building a School Testing Program. Houghton Mifflin Co. Boston: 1963.
- Brim, Orville Jr., American Psychologist. "American Attitudes Toward Intelligence Testing" Vol. 20. #2. February 1965.
- Brown, Toby and John Sipkens. Planning a School Testing Program. Montana State University, 1972. Unpublished paper.
- Buros, Oskar K. The Seventh Mental Measurement Yearbook. Highland Park, New Jersey: The Gryphon Press, 1972.
- Chauncey, Henry and John E. Dobbin. Testing: Its Place in Education Today. Harper and Row. New York: 1963.
- Cronback, Lee J. Essentials of Psychological Testing. Harper and Row. New York: 1960.
- Dyer, Henry S. "Needed Change to Sweeten the Impact of Testing." Personnel and Guidance Journal. V. 45 No. 8. April 1967.
- Goslin, David A. The Search for Ability. Russell Sage Foundation. New York: 1963.
- Karmel, Louis J. Measurement and Evaluation in the Schools. MacMillan Co., Toronto: 1970.
- Loretan, Joseph. "The Decline and Fall of Group Intelligence Testing." Teachers College Record. V. 67. No. 1. October, 1965.
- Lyman, Howard B. Test Scores and What They Mean. Prentice Hall, Inc. Englewood Cliffs, New Jersey: 1963.
- Mehren, William A. and Irvin J. Lehmann. Standardized Tests on Education. Holt, Rinehart and Winston, Inc. New York: 1969.

Neulinger, John. "Attitudes of American Secondary School Students Toward the Use of Intelligence Tests." Personal and Guidance Journal. V. 45, No. 4. December, 1966.

Nunally, J. C. Educational Measurement and Evaluation. McGraw Hill Book Co. New York: 1964.

Payne, David A. The Specification and Measurement of Learning Outcomes. Blaesdell Publishing Co. Walthans, Mass. 1968

Thorndike, Robert L. and Elizabeth Hagen. Measurement and Evaluation in Psychology and Education. John Wiley and Sons. New York: 1955.

APPENDIX

Please complete this form by checking the appropriate blanks which apply to your situation or by filling in the blanks with the information required.

Please list the standardized tests in the areas of achievement, intelligence, aptitude and interest which are given in your school.

Name of test and form given	Grade level	Date of testing	How scored? hand machine	

(Please use reverse side of page if necessary)

1. Do you feel your school's standardized testing program yields satisfactory information to:

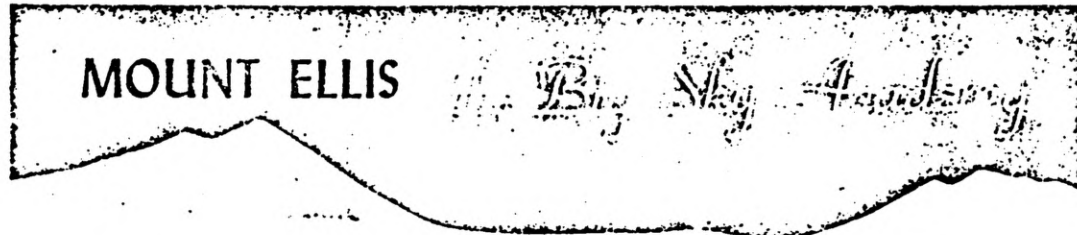
	Yes	No
a. benefit the student	_____	_____
b. benefit the teacher	_____	_____
c. benefit the administrator	_____	_____
d. benefit the parents	_____	_____

2. If you answer No to any of the above, would you state how this might be improved.

3. Who most often is called upon to explain test results?
_____classroom teacher _____counselor _____administrator
4. Please indicate below the two most important objectives of your school's standardized testing program.
- 1.
- 2.
5. Which tests, if any, do you feel could be dropped from your testing program?
6. Are there any areas that are not now adequately covered by your testing program that you feel should be covered?
_____yes _____no If yes please list areas.
7. Would you favor a unified standardized testing program for K-12 in the North Pacific Union? _____yes _____no.

Additional comments:

If you would like the results of this study sent to you please put your name and address below. Thank you.



RT. 2 BOX 136
BOZEMAN, MONTANA 59715

TELEPHONES
AREA CODE 406

PRINCIPAL	596-2648
REGISTRAR	596-5746
TREASURER	596-5740
DEAN OF BOYS	596-5778
DEAN OF GIRLS	596-5773
BOYS' DORM	596-5946
GIRLS' DORM	596-5948

April 7, 1972

Dear Colleague:

I have been authorized by the Board of Education of the North Pacific Union to, "...conduct a definitive study and recommend a standardized testing program for kindergarten through the twelfth grade." This study partially fulfills the requirement for a Master's of Education degree at Montana State University.

In order to be able to recommend an acceptable program, I am asking you to take a few minutes of your time to fill out the enclosed questionnaire and return it to me in the enclosed, self-addressed, stamped envelope by April 22, 1972.

Your immediate attention to this questionnaire and its prompt return will greatly facilitate this study. All information will be held as confidential and will not be seen by anyone except me.

Sincerely,

Brent R. Pfeiffer
Registrar

TACOMA PUBLIC SCHOOLS
Joan England
Jim Laurent

PUPIL PERSONNEL SERVICES

RESEARCH DEPARTMENT

YOUR LORGE-THORNDIKE SCORES AND WHAT THEY MEAN

WHAT IS THE LORGE-THORNDIKE?

A few weeks ago you and all the other seventh graders in Tacoma Public Schools took the Lorge-Thorndike Intelligence Test. The test questions asked you to solve many different kinds of problems. Some required word knowledge and reading ability; these comprised the Verbal test. Some used only drawings or numbers; these made up the Nonverbal test.

VERBAL measures ability to reason and see relationships among words and ideas expressed in words. This ability relates to skills in reading, writing and speaking. You might think of it as general scholastic aptitude. It is important for school subjects and occupations in which the use of words is important.

NONVERBAL measures ability to reason and see relationships among numbers and ideas expressed in diagrams and drawings. This kind of test affords an opportunity to demonstrate reasoning ability without using words. Ability such as this is important for school subjects and occupations in which one works with objects and materials rather than with words.

WHY DID WE TAKE THE LORGE-THORNDIKE?

The Lorge-Thorndike is part of the Basic Testing Program in the Tacoma Public Schools. This program consists of nationally-standardized tests of aptitudes and achievements which are given to all pupils at certain grade levels in Tacoma schools. You may remember another such test, the SCAT-STEP from earlier grades.

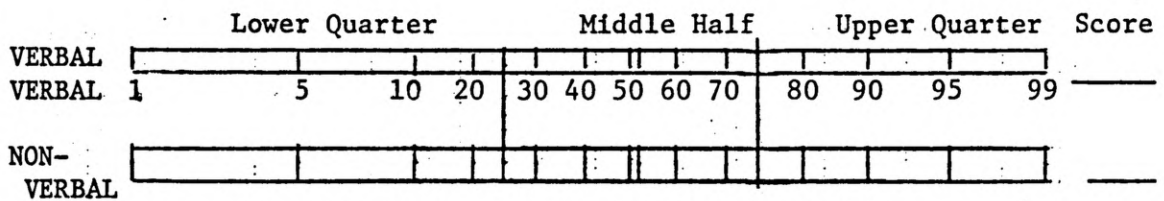
Most school districts have basic testing programs. These represent an attempt to measure systematically the developing aptitudes and achievements of each pupil as he progresses from kindergarten through high school. Results of these objective measurements, when properly interpreted, can help teachers and parents to better understand pupils-and pupils to better understand themselves.

HOW DID I DO?

Any measurement is an estimate, and the Lorge-Thorndike is one kind of measurement. If you took this test over and over again your scores would not always be the same; sometimes they would be a little higher, sometimes lower. Over a period of years they might be quite different. The reported percentile band shows the limits within y

Your Lorge-Thorndike Scores and What They Mean (continued)

not always be the same; sometimes they would be a little higher, sometimes lower. Over a period of years they might be quite different. The reported percentile band shows the limits within which your scores would be likely to fall most of the time. They show how your performance on the Lorge-Thorndike compares with that of seventh grade boys and girls from all over the United States.



NATIONAL SEVENTH GRADE PERCENTILE SCORES

Remember that these scores represent only two items of information. As you study this sheet keep in mind your previous standardized test scores, grades, interest, aspirations and everything else you know about yourself. Never look just at a stst score by itself. Your counselor will be happy to discuss this test and these results with you and help you relate them to your educational planning for the future.

JAL/jh
9-68

SECONDARY TEST RECORD

Name: _____

Test: Differential Aptitude Date & Grade		Sept.	9th	California Occupational Preference Survey	Sept.	9th	Washington Pre- Course Test	May	11th	Other Test Results	
Verbal Reasoning		_____		Total:		Summary:					
%tile		_____		Summary:							
Local		_____									
National		_____									
Numerical Ability		_____									
%tile		_____									
Total		_____									
Abstract Reasoning		_____									
Clerical Speed and Accuracy		_____									
Mechanical Reasoning		_____		Orleans-Hanna Alg Prognosis Test		9th		S.V.I.B.		12th	
Space Relations		_____		Score:				Summary:		High School Attend- ance Record	
Spelling & Grammar		_____		Remarks:						Freshman:	
Remarks:				S.A.T. Reading Test		9th				Sophomore:	
				Score:						Junior:	
				Remarks:						Senior:	

ELEMENTARY TEST RECORD

Student Name: Murphy-Durrell Birthdate: May 1972 Social Security No.: Grade K Elementary School
 Fathere Name: Address Street No. City State Attendance Record

S.R.A. Reading		SRA Achievement	SRA Achievement	SRA Achievement	SRA Achievement
Subtest	/3rd	/4th	/5th	/6th	8th
Letternames Total Score					
Xtyle	Local National				
Test Total:					
Xtyle	Local National				
Remarks:					
Comprehension:					
Xtyle	Local National				
Vocabulary:					
Xtyle	Local National				
Total:					
Xtyle	Local National				
Remarks:					
Language Arts					
Xtyle	Local National				
English Usage					
Xtyle	Local National				
Arithmetic Usage:					
Xtyle	Local National				
Reading Total					
Xtyle	Local National				
Remarks:					
Large-Thorn-dike IQ.					
Verbal:					
Xtyle	Local Nat'l				
Nonverbal:					
Xtyle	Local Nat'l				
Total:					
Xtyle	Local Nat'l				
Remarks:					

SOURCES FOR EDUCATIONAL AND PSYCHOLOGICAL TESTS

Harcourt, Brace and Jovanovich
757 Third Avenue
New York, New York 10017

McGraw-Hill Book Co.
330 West 42nd.
New York, N. Y. 10036

Science Research Associates
259 East Erie
Chicago, Illinois 60611

Houghton, Mifflin Co.
110 Tremont St.
Boston, Mass. 02107

Psychological Corporation
304 East 45th St.
New York, N. Y. 10017

Educational and Industrial Testing Service
P. O. Box 7234
San Diego, California 92107

Educational Testing Service
Princeton, New Jersey 08540

Stanford University Press
Stanford, California 94305