A SURVEY OF TESTING PROGRAMS
IN MONTANA HIGH SCHOOLS

BY

EVERETT A. AARESTAD

Submitted in partial fulfillment of the requirements for the Master of Education degree in the Department of Education
Montana State College
June, 1964
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Statement of the Problem</td>
<td>2</td>
</tr>
<tr>
<td>Procedure</td>
<td>3</td>
</tr>
<tr>
<td>Definitions</td>
<td>3</td>
</tr>
<tr>
<td>Limitations</td>
<td>4</td>
</tr>
<tr>
<td>II. REVIEW OF RELATED LITERATURE</td>
<td>6</td>
</tr>
<tr>
<td>Origin and History of the Testing Movement</td>
<td>7</td>
</tr>
<tr>
<td>Present Status</td>
<td>14</td>
</tr>
<tr>
<td>Aspects of the Future</td>
<td>17</td>
</tr>
<tr>
<td>Summary</td>
<td>19</td>
</tr>
<tr>
<td>III. RESULTS OF THE SURVEY QUESTIONNAIRE</td>
<td>20</td>
</tr>
<tr>
<td>IV. SUMMARY, CONCLUSIONS AND RECOMMENDATIONS</td>
<td>28</td>
</tr>
<tr>
<td>Summary</td>
<td>28</td>
</tr>
<tr>
<td>Conclusions</td>
<td>23</td>
</tr>
<tr>
<td>Recommendations</td>
<td>29</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>31</td>
</tr>
<tr>
<td>APPENDIX</td>
<td>35</td>
</tr>
<tr>
<td>Appendix A: Sample of the Questionnaire</td>
<td>36</td>
</tr>
<tr>
<td>Table Number</td>
<td>Description</td>
</tr>
<tr>
<td>-------------</td>
<td>---------------------------------------------------------</td>
</tr>
<tr>
<td>I.</td>
<td>Rank Order of Achievement Tests</td>
</tr>
<tr>
<td>II.</td>
<td>Rank Order of Aptitude Tests</td>
</tr>
<tr>
<td>III.</td>
<td>Rank Order of Interest Inventories</td>
</tr>
<tr>
<td>IV.</td>
<td>Rank Order of Personality Tests</td>
</tr>
<tr>
<td>V.</td>
<td>Rank Order of Vocational Tests</td>
</tr>
<tr>
<td>VI.</td>
<td>Rank Order and Percentages of Uses of Test Results</td>
</tr>
</tbody>
</table>
A Survey of Testing Programs In Montana High Schools

Everett A. Anestad

Statement of the Problem:

A survey was conducted to determine what standardized tests are most commonly used by high school guidance departments, and for what purposes test results are used.

Procedures:

A survey questionnaire was constructed which consisted of two parts: (1) A list of commonly used standardized tests, with space provided for additional tests used. (2) A list of 25 statements relative to the uses of test results. A questionnaire was sent to all public high schools who had a guidance department as determined by the State Department of Instruction. One hundred twenty-eight out of 157 schools participated.

The standardized tests as reported on the survey questionnaire were grouped into five categories of achievement, aptitude, interest, personality and vocational. Each category was ranked according to the frequency with which the tests had been reported.

The statements relative to the use of test results were ranked in order and a percentage computed on the basis of frequency of being reported.

Summary and Conclusions:

The following statements are representative summaries and conclusions of this study.

1. The Iowa Test of Educational Development is the most popular standardized achievement series.
2. More achievement tests are used than any other type test.
3. The Differential Aptitude Test is used more often than other aptitude type tests.
4. The most commonly used intelligence test is the Otis Mental Ability Test.
5. Few personality tests are being used in high school guidance departments.
6. The Kuder Preference Record was reported most often of all vocational tests.
7. Test results often aid in planning and selecting student curriculums.
8. Test results often form the framework for counseling interviews.
CHAPTER I

INTRODUCTION

Studies of individual differences began toward the end of the 19th century, and were the primary impetus for the testing movement as it exists today. It had its roots of origin in the work of such notables as Binet, Wechsler, Galton, Thorndike, Terman and others. In exploring their ideas of measurement, they recognized the fact that innate differences must exist, both in psychological and physiological functioning. Development from this original impetus resulted in a multitude of measuring instruments, each designed in its own way, each intended to accomplish its purpose.

With today's emphasis on the testing of individuals, it is not strange that one should find this program firmly entrenched in the school systems. Schools assume direct responsibility for guidance and assistance to students in helping them to achieve their goals.

It is natural that the schools should employ some measure of the characteristics of an individual in assisting him to better understand himself, as often these factors are deeply hidden behind various kinds of outward reactions. It is the endeavor of any measuring instrument to bring these factors to the attention of the individual, and to persons whose qualifications will assist that person to more effectively deal with his environment. It should not be implied that tests are a means to an end, as they are not all inclusive. Their qualities lie rather in the dissemination of information about a person, and their help in determining the strengths and weaknesses of a total personality picture, for better adjustment of the individual.

---

Statement of the Problem

In addition to the influencing factors of the home, the church and the community, the major responsibility of initiating and carrying out the functions of organized guidance falls upon many of our public schools. In the association with schools, the student is administered various tests which purport to measure some aspect of that individual. These results lend themselves to psychiatric and impressionistic conclusions which may aid the person in a more accurate self-appraisal.

It was the intention of this author to ascertain the answers to these and other questions.

1. What specific standardized tests are most popularly used by public school counselors in Montana?

2. Of the five categories of tests, achievement, aptitude, intelligence, interest and personality, which group is most extensively used by guidance departments in Montana?

3. What are the main uses of test results?
   a. Are they used in determining curriculum offerings?
   b. Do they aid students in college or vocational preparation?
   c. Are they used for prediction, selection, classification or grouping of students?
   d. Do test results form the framework for the counseling interview?
   e. Are test results used during parent-teacher conferences?

---

Procedure

With assistance from Mr. Albert Suvak, Assistant Professor of Counseling and Testing, a survey questionnaire was constructed to include a representative number of commonly used tests in guidance practices. Twenty-five statements relative to the use of test results were included, and the counselor was instructed to mark which tests he used, with additional space provided for writing in other tests used, and to indicate which statement was most appropriate for his use of test results.

The survey questionnaires were sent to all public schools in the state who have a guidance department as determined by the State Department of Instruction.

The compiled data was treated as follows:

1. Tests Used: The tests as reported on the survey questionnaire were grouped into respective categories, achievement, aptitude, intelligence, personality and vocational according to groupings as used by Euron. Next they were ranked in order on the basis of their frequency.

2. Test Results: The statements as reported on the survey questionnaire were ranked in descending order on the basis of their frequency and a percentage computed.

Definitions

The term guidance department in this paper refers to an organization operating within the limits of an institution of learning for the express purpose of providing assistance by personally qualified and adequately trained people to an individual of any age to help him manage his own activities.

---

develop his own points of view, make his own decisions and carry his own burdens.  

A standardized test is one in which the procedure, apparatus and scoring have been fixed so that precisely the same test can be given at different times and places.  

A school counselor is any person who is trained as a specialist in human relations who involves in a cooperative enterprise all those who affect the development of the child.  

Limitations  

This survey was limited to public schools having a recognized guidance department as determined by the State Department of Instruction, as it was the desire of the writer to ascertain information pertinent only within the State of Montana. Guidance departments in Montana schools vary greatly in the number of personnel and number of students served. Some operate only in the high school whereas others include elementary schools. This variation in the departments surveyed produced a limitation in that some statements may apply to larger departments while others may be more appropriate for smaller organizations.

The counselors were not asked to state how often any particular test was used, nor to qualify their answers in any other manner. It must therefore be assumed that interpretation of each question could vary. This would
constitute a limitation of the validity of the survey.

Recognition of some tests by more than one name or set of initials could have caused duplication or omission as to the most commonly used tests. In actual testing programs, some counselors may have reported tests not actually used by them, but administered to students under their jurisdiction by other persons or organizations. These constitute additional limitations to this study.
CHAPTER II

REVIEW OF RELATED LITERATURE

Anything that exists at all exists in some quantity, and anything that exists in some quantity is capable of being measured.¹

This statement was made by Edward L. Thorndike, a prominent figure in the early field of measurement. It does not imply that one is able to measure all things. It simply states any existing quantity is capable of being measured. Measurement of any kind is a matter of determining how much or how little, how great or how small, how much more than or how much less than. In more refined terms, to measure means to observe or determine the magnitude of a variety.²

Measuring should be as accurate as is necessary to accomplish its intended purpose. Sometimes a rough measure will suffice, as in the case of a person asking for a piece of string about three feet long. Someone may state the next town is about thirty miles away.

Measurement does not always include the use of an instrument. Many times individuals are rated by someone, or by themselves. Under these conditions, the person is measured according to a pre-conceived set of criteria.³

Such a brief discussion of measurement would not be complete without giving some thought to errors. All types of measurement can be considered


valuable if they serve their purpose with a limited amount of error. In
general, it is said that the sources of error in measurement are primarily
caused by imperfections, either in the measuring instruments, or in the method
in which they were used. While both sources of error are subject to consider-
able amount of control, neither can be entirely eliminated. Three methods of
controlling errors in measurement are: (1) Improve existing instruments,
(2) devise adequate methods of estimating or allowing for these errors,
(3) develop skills in applying the instruments of measurement.

The classical history of the testing movement dates back to the time
before Christ. The next topic concerns these origins.

Origin and History of the Testing Movement

Perhaps the earliest records indicating use of any type test were found
in Chinese history about the year 200 B.C. They were credited with having
used tests in the selection of civil servants. Socrates used oral quizzes
about 400 years before Christ. As is true in many movements it is difficult
to express an exact date for a formal beginning of the test movement. Practic¬
ally all of the work which has been done in conjunction with standardized
tests has occurred in the last fifty years. Prior to this, measurement existed

---


5 Ibid., p. 12.

6 Ibid., p. 13.

7 Ibid., pp. 27-28.

in various fields, astronomy being one of the older and more highly developed. In 1795 an astronomer's assistant named Kinneybrook was found to consistently differ from his superior in reaction time of reporting astronomical observations. Astronomers as well as other scientists became interested in the causes of different reaction times.

The report of a school committee which visited the English High School in Boston in 1845 records that members of a committee examined pupils in algebra, geometry and French at that time.

In 1837, Horace Mann had been appointed Secretary of the Massachusetts State Board of Education. His job included observations of the "common" schools, in which he pointed out weaknesses. This did not meet with approval of the schoolmasters which resulted in a committee being formed for resisting Mann. A quarrel followed which ended in an agreement to prepare written examinations in history, arithmetic, geography, grammar, natural philosophy and astronomy. This incident was a major unifying factor among the common schools of the area at that time.

About 1864 an English schoolmaster, Reverend George Fisher constructed a "scale book" which contained samples of typical questions and of various degrees of proficiency in answering the questions in several school subjects.

---

9 Haus and Stanley, op. cit., p. 6.


12 McCanna, Stanley J. and Glick, Marvin D., Evaluating Pupil Growth, Allyn and Bacon, Boston, 1953, p. 15.

These questions were intended to serve as models for future examinations.

At approximately the same time, Sir Francis Galton in England became active in the exploration of measuring individual differences, mental abilities and in the use of statistical methods. During the years 1865 to 1878 the Regents' Examinations were inaugurated in New York State.

Shortly after Galton's work in England, the measurement movement in the United States began, getting much of its start from James McKeen Cattell, an American psychologist. He became interested in the problem of individual differences and did a number of experiments in sensory-motor abilities. He assumed the theory that differences in sensory keenness, speed of reaction and similar abilities would reflect differences in intelligence. Cattell became very disappointed when no clear relationships were found to exist between these variables. He is credited with being one of the testing movement pioneers and was the first to use the term, "mental tests".

The next era of progress occurred in the field of intelligence. The French had long been leaders in abnormal psychology and this resulted in their attempts to classify and treat the mentally defective, some of the work being done by Alfred Binet. His studies of law, medicine and biology gave him a

---

17. Ibid., p. 19.
18. Ibid., p. 12.
broad background for studying intelligence although he was never quite sure exactly what he was trying to measure. His main goal was to find a simple, rapid and precise method for identifying mentally retarded children and measuring the extent of retardation. After much study and prolonged experimentation, Binet and his assistant, Theodore Simon, presented a series of tests which reputedly would measure the intelligence level of children. This resulted in the immediate interest of workers in other countries to begin experimenting and communicating with Binet and his tests. Three years later he published an improved version of his scale and in 1911 still another. His death occurred shortly after. It is little wonder that Binet did not know exactly for what he was searching. He could be compared to a hunter going into the woods to find an animal no one had ever seen. Most everyone was sure the beast existed for it had left "sign".

Binet knew there must be something like intelligence, since he could see its effects, but could not describe it. One idea basic to his findings was that some children were bright and some dull. He discovered that those who were best on tests of judgment were also superior in attention, memory, vocabulary and related fields. These associations led him to believe there existed some common causal factor. Through a process of trial and error, Binet came to his description of intelligence. He described its essence as

20bid., p. 29.
23bid., p. 160.
"the tendency to take and maintain a definite direction; the capacity to make adaptations for the purpose of attaining a desired end, and the power of auto
criticism. 24

Credited as the real inventor of comparative tests in America was Dr.
J. M. Rice who, in 1934, began his work on achievement tests. 25 He had
studied in Germany and came under the influence of experimental psychologists.
In the United States the major share of his efforts had to do with a uniform
spelling test which he administered to pupils in many cities, and the con-
struction and administration of tests in arithmetic and language. Rice's
work met with passive acceptance by other educators and it was not until 15
years later that he received proper recognition. 25

In 1934 appeared one of the first books dealing directly with mental
measurement, entitled An Introduction to the Theory of Mental and Social
Measurement, authored by Edward L. Thorndike. 27 It contained statistical
methods and fundamental principles of testing, and became popular as a text-
book. It served to stimulate interest and activity in the field of measure-
ment, remaining in such a capacity for the next ten years. The honor of being
called the "father of the testing movement" goes to Thorndike, for no other
person has touched the measurement movement at so many points or made so many
contributions. 28

24 Ibid., p. 160.
25 Ibid., p. 33.
26 Ibid., p. 33.
27 Ibid., p. 33.
Contribution to improvements in the testing movement was the realization that existing measurements varied. A series of studies were conducted at the University of Missouri and the University of Chicago in which the inequities of present ranking systems were demonstrated. 29

Starch and Elliot conducted a study in 1913 at which time they had 116 high school teachers grade the same geometry papers. The values assigned ranged from a low of 23 to a high of 92. The manifestation was if high school teachers could not agree any better than that on such an objective subject, the situation was indeed bad. 30

The Binet-Simon scales had not as yet received much attention in the United States until 1908 when H. H. Goddard translated an accurate version. 31 In 1911 he revised it, and in 1912 Kuhlman published another revision. 32 This sufficed for short duration and was overshadowed by the 1916 revision. This work, done by Louis Terman at Stanford University, began in 1912 but did not become readily available until 1916. 33

Stern and Kuhlman had suggested a ratio of mental age to chronological age as a good method of determining the status of a person's intelligence. The device attracted little attention until 1916 when Terman made the term I. Q.

29 Ibid., p. 39.
31 Ross, op. cit., p. 33.
32 Ibid., p. 33.
popular by incorporating it as a basic feature of a mental test. The Stanford Binet 1916 Revision had many other features which made it superior over any of its predecessors. For the first time in history, direct administrative rules were provided rather than general instructions. Terman emphasized the role of the administrator and his judgment in scoring the test. Many of Binet's original items were discarded and new ones added. Terman was also among the first to realize the importance of representative sampling of subjects for use in standardization. Twenty-one years later the 1916 Revision underwent major changes and the 1937 Revision came out.

Immediately after the declaration of war in 1917, a special committee of the American Psychological Association was organized to answer the question: What can psychologists do to help win the war? The committee decided its main contribution was the development of some devices for aiding in the classification of men as to general ability and specific talents. At this time, Arthur S. Otis had been working on a group intelligence test, and he turned over all of his materials to a committee of psychologists. Upon this model the famous Army Alpha, a verbal test, was constructed. Following this came the Army Beta test for illiterates, a non verbal test. In a short time other non verbal standardized tests came into being such as the Arthur Point Scale.

36 Ibid., p. 63.
37 Ibid., p. 64.
38 Ibid., p. 66.
39 Ibid., p. 67.
40 Ibid., p. 67.
the Pintner Paterson and the Cornell Coxe series.\textsuperscript{41}

Galton had emphasized the need for mathematical treatment of psychological data from the beginning, and with later influences from Karl Pearson, G. U. Yule, E. L. Thorndike, W. Brown, C. Spearman and R. A. Fisher, among others, the idea became firmly rooted as a device for determining test construction and test useability.\textsuperscript{42}

**Present Status**

Many tests developed during the periods before and after World War I are still in use today. An example is the Rorschach Ink Blot Test, first published in Bern, Switzerland in 1921 under the title, *Psychodiagnostica: A Diagnostie Test Based on Perception*.\textsuperscript{43} This test is used extensively and gives results which substantiate clinical predictions about fifty percent of the time.\textsuperscript{44}

Probably one of the greatest forces which perpetuated the testing movement was World War II. With millions of men and women being inducted into the military services during the years 1941-1945, the job of classifying and assigning became a major undertaking.\textsuperscript{45} As a result, the field of measurement continued to develop and expand.

A relatively new approach in the field of intelligence and aptitude testing is the use of factory analysis, a statistical method for refining

\textsuperscript{41}\textit{Ibid.}, p. 69.
\textsuperscript{42}\textit{Ibid.}, p. 85.
\textsuperscript{45}Noll, \textit{op. cit.}, p. 29.
tests into basic elements or factors. These are factors which the test author finds most adequate for measuring what the test purports to measure.\(^46\)

The Primary Mental Abilities Test had its origin in factor analysis. Louis L. Thurstone, during test construction, identified the following factors of intelligence: verbal, word fluency, number, spatial, memory, reasoning and perceptual speed.\(^47\) In general, with the emphasis upon measurement of narrower special abilities, factor analysis has become helpful. It gives information about the nature and organization of individual characteristics. It tries to determine how many distinct abilities are being measured more accurately through modification of a few questions. It tends to unify identical traits being measured by several tests but given different names by different authors.\(^48\)

Three types of factors are commonly distinguished, general, group and specific. A specific factor is present in one test but not in another. A group factor is present in more than one test. A general factor is one found in all tests.\(^49\) If a single factor is found predominantly in a test, it is said to have a high factor loading for that trait. If intercorrelations between tests or test items are low, it strengthens the theory that a single factor may be more prevalent.\(^50\)

One necessary step in factor analysis is "rotation". This is a procedure for placing the factors so the results will be more meaningful. Rotation

\(^{46}\)Cronbach, \textit{op. cit.}, p. 247.


\(^{48}\)Cronbach, \textit{op. cit.}, p. 247.

\(^{49}\)\textit{ibid.}, p. 250.

\(^{50}\)\textit{ibid.}, p. 259.
will facilitate identification of any single trait, or groups of factors. One charge against factor analysis is it has not given enough attention to the processes which must be utilized for any given test item as to whether it deals with words, numbers, forms, or other symbols.

Significant in the field of factor analysis and in the study of mental processes is the classification of intellectual tasks into a three way organization by J. P. Guilford. He distinguishes five types of mental operations, memory, cognition, convergent thinking, divergent thinking and evaluation. Tasks within each of these categories can be classified as "content" or "product". The content categories are figural, symbolic, semantic and behavioral. The six kinds of products as represented by Guilford's theory are units of information, classes of units, relation between units, systems of information, transformations and implications. These five operations, four content categories and six products produce 123 different combinations, each representing a type of intellectual task. He portrays this theory with a cubical model separated into 123 cells, each with the potential components of one of the above mentioned processes.

The present decade may be characterized as a time of adjustment and refinement in the field of measurement. Certainly refinement of existing

---

51 Ibid., p. 255.
52 Ibid., p. 263.
54 Ibid., p. 470.
tests and development of new ones is occurring constantly, and the possibility of new and original developments exists. Many unsolved problems persist, only a few which have been discussed here.

Cronbach summarizes the present status of testing as follows:

Measure a lot of things somewhat inaccurately, know your risk and check up.... Interviews, essay tests and projective tests are not rifles aimed at a narrow target; rather they are sawed off shotguns spraying rather wildly but frequently hitting the mark, while at the same time nicking innocent bystanders.

Aspects of the Future

Advancement in knowledge of test construction and improvement of testing techniques will probably encompass the majority of research to be completed in the next few decades. The changes which are sure to occur may be more spectacular than those which previously have taken place. The present emphasis upon providing better facilities for the training of people associated with all forms of testing, plus the movement toward certification of those who enter the field are steps in the proper direction. The future is largely dependent upon the selection, training and research orientation of those people who enter the field. Efforts to get statistical analysis in a more favorable light is perhaps the first step in orientation designed toward the future use of computers, and other automation devices. The influence of the punch card, the printed circuit and the "memory" of these machines will invariably influence

56 Ibid., p. 31.
future direction of test construction and utilization.\textsuperscript{59} The rising tide of school enrollments will no doubt be one of the forces which, in the light of increasing automation, could precipitate a mass testing movement that could make present efforts look meager.\textsuperscript{60} Mention should be made of Robert L. Thorndike's work, being completed at the present time, on the aptitude census of the American people.\textsuperscript{61}

Most indications point to a greater sophistication level in measurement. Increasing consumer discrimination is a powerful force in the development of tests, and in their success.\textsuperscript{62} These demands for higher standards by test users should result in the production of better tests than are now available. Basic to this is a more satisfactory understanding of the nature of human behavior, and the development of clear theories of intelligence.\textsuperscript{63} No serious efforts can be made using techniques invented 30 years ago. It is reasonable to hope that some future worker will start from a theory of mental processes, choose or design tests to measure those particular processes, and so arrive at a superior diagnostic device.\textsuperscript{64} Also of importance is the agreement among test developers as to the meaning of terms and definitions.\textsuperscript{65} Research in the future must, out of necessity, insist upon the sound training

\textsuperscript{59} \textit{Ibid.}, p. 325.

\textsuperscript{60} \textit{Ibid.}, p. 324.


\textsuperscript{62} Rothney, Danielson and Heimann, \textit{op. cit.}, p. 329.

\textsuperscript{63} \textit{Ibid.}, p. 330-331.


\textsuperscript{65} Rothney, Danielson and Heimann, \textit{op. cit.}, p. 332.
and understanding which statistical analysis of data can provide.

Four current activities indicative of future improvement in the field of testing are:

1. Publication of new tests that indicate more careful attempts at standardization and validation.
2. Publication of standards for the sale and distribution of tests.
3. Provision of methods by which critical reviews of tests may reach the potential customer.
4. Raising of standards for training and employment of workers in therapeutic and counseling fields.66

Summary

The 20th century has witnessed the beginning of the measurement movement to its present status. It is firmly entrenched in educational, industrial, social, medical and military structures, and the future indicates it will become more firmly embedded. Test users have come to realize some of the errors and limitations which exist in the testing industry. Test consumers are demanding better and better testing tools, so one can assume these to be forthcoming.67 Adaptations are now occurring for the use of automation in testing. This will speed scoring, diagnosis and decision making. It may produce increased dependency upon tests.

66 Ibid., p. 332.
67 Ibid., p. 330.
CHAPTER III
RESULTS OF THE SURVEY QUESTIONNAIRE

A total of one hundred fifty seven questionnaires were sent to public school guidance counselors throughout the state, of which one hundred twenty eight replied, for an eighty one and five tenths percent return.

The compiled data will give the reader a picture of what tests are commonly used, and will acquaint counselors with what seem to be the more popular tests used.

In the category of achievement testing, it was found that the Iowa Test of Educational Development was used in preference to other achievement series and the American College Testing Program was second in common usage. The National Merit Scholarship Qualifying Test was reported the third most commonly used achievement battery, with the Iowa Achievement Test Series and the Stanford Achievement Test occupying fourth and fifth rank respectively. These were followed by seventeen additional achievement type tests in less common usage than the five previously mentioned. The complete list in rank order can be found on the next page.

Of the representative types of aptitude batteries, the Differential Aptitude Test was found to hold undisputed precedence over all others, having been reported nearly twice as often as the Scholastic Aptitude Test series which ranked second. The General Aptitude Test Battery was found to be third in common usage, followed by eight others less often reported. A list of these can be found in Table II, on page 22.

The Otis Mental Ability Test was checked as the most commonly used intelligence test by Montana public school counselors. The Wechsler Intelligence Scales were reported next in common usage, closely followed by the
<table>
<thead>
<tr>
<th>Name of Test</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Iowa Test of Educational Development</td>
<td>58</td>
</tr>
<tr>
<td>2. The American College Testing Program</td>
<td>35</td>
</tr>
<tr>
<td>3. National Merit Scholarship Qualifying Test</td>
<td>35</td>
</tr>
<tr>
<td>4. Iowa Achievement Tests</td>
<td>33</td>
</tr>
<tr>
<td>5. Stanford Achievement Test</td>
<td>32</td>
</tr>
<tr>
<td>6. California Achievement Test</td>
<td>20</td>
</tr>
<tr>
<td>7. National Educational Development Test</td>
<td>14</td>
</tr>
<tr>
<td>8. Metropolitan Achievement Test</td>
<td>11</td>
</tr>
<tr>
<td>9. Cooperative General Achievement Test</td>
<td>10</td>
</tr>
<tr>
<td>11. Iowa Silent Reading Test</td>
<td>8</td>
</tr>
<tr>
<td>12. Iowa Test of Basic Skills</td>
<td>6</td>
</tr>
<tr>
<td>13. Sequential Test of Educational Progress</td>
<td>2</td>
</tr>
<tr>
<td>14. Coordinated Scales of Attainment</td>
<td>1</td>
</tr>
<tr>
<td>15. Essential High School Content Battery</td>
<td>1</td>
</tr>
<tr>
<td>16. Harison’s High School English Test</td>
<td>1</td>
</tr>
<tr>
<td>17. The Pribble-Ballman Diagnostic Tests in Elementary Language Skills</td>
<td>1</td>
</tr>
<tr>
<td>18. Minnesota Achievement Test</td>
<td>1</td>
</tr>
<tr>
<td>19. Academic Precise Test</td>
<td>1</td>
</tr>
<tr>
<td>20. Wagner Reading Readiness Test</td>
<td>1</td>
</tr>
<tr>
<td>21. Midwest High School Achievement Examination</td>
<td>1</td>
</tr>
<tr>
<td>22. Harry-Durant Essential High School Battery</td>
<td>1</td>
</tr>
</tbody>
</table>
TABLE II, APTITUDE BATTERIES.

<table>
<thead>
<tr>
<th>Name of Test</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Differential Aptitude Test</td>
<td>92</td>
</tr>
<tr>
<td>2. Scholastic Aptitude Test</td>
<td>57</td>
</tr>
<tr>
<td>3. General Aptitude Test Battery</td>
<td>43</td>
</tr>
<tr>
<td>4. California Algebra Aptitude Test</td>
<td>11</td>
</tr>
<tr>
<td>5. Montana State College Nursing Aptitude Test</td>
<td>8</td>
</tr>
<tr>
<td>6. Air Force Preference Inventory</td>
<td>6</td>
</tr>
<tr>
<td>7. Iowa Algebra Aptitude Test</td>
<td>2</td>
</tr>
<tr>
<td>8. Lee Test of Geometric Aptitude</td>
<td>2</td>
</tr>
<tr>
<td>9. Orleans Algebra Test</td>
<td>2</td>
</tr>
<tr>
<td>10. Managan Aptitude Classification Test</td>
<td>1</td>
</tr>
<tr>
<td>11. The Jastak Test of Potential Ability and Stability</td>
<td>1</td>
</tr>
</tbody>
</table>

Preliminary Scholastic Aptitude Test, The Kohlman-Anderson Intelligence Test, the Stanford-Binet, the Kemnon Haison Test of Mental Ability and the California Mental Maturity Test, in that order. Ten others were reported on the questionnaire, a complete list which may be found in Table III, on the following page.

In the field of personality testing, the Hovensy Problem Checklist and the Science Research Associates Youth Inventory Checklist appeared to be most popularly used, followed by three others of somewhat less incidence. These are reported in Table IV, page 23.

Two tests of Vocational Interest have dominated this field of testing for many years, so it is not strange to find the Kuder Preference Record reported on ninety questionnaires and the Strong Vocational Interest Blank
### TABLE III. INTELLIGENCE TESTS.

<table>
<thead>
<tr>
<th>Name of Test</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Otis Mental Ability Test</td>
<td>73</td>
</tr>
<tr>
<td>2. Wechsler Intelligence Scales</td>
<td>36</td>
</tr>
<tr>
<td>3. Preliminary Scholastic Aptitude Test</td>
<td>35</td>
</tr>
<tr>
<td>4. Kuhlman Anderson Intelligence Test</td>
<td>33</td>
</tr>
<tr>
<td>5. Stanford Binet</td>
<td>31</td>
</tr>
<tr>
<td>6. Henmon Nelson Tests of Mental Ability</td>
<td>26</td>
</tr>
<tr>
<td>7. California Mental Maturity Test</td>
<td>25</td>
</tr>
<tr>
<td>8. Kuhlman Finch</td>
<td>7</td>
</tr>
<tr>
<td>9. Cooperative School and College Ability Test</td>
<td>7</td>
</tr>
<tr>
<td>10. Primary Mental Ability Test</td>
<td>6</td>
</tr>
<tr>
<td>11. College Qualifications Test</td>
<td>5</td>
</tr>
<tr>
<td>12. College Entrance Examination Board Scholastic Aptitude Test</td>
<td>5</td>
</tr>
<tr>
<td>13. Lorge Thorndike Intelligence Test</td>
<td>5</td>
</tr>
<tr>
<td>14. American Council on Education Psychological Examination</td>
<td>4</td>
</tr>
<tr>
<td>15. Ohio State University Psychological Test</td>
<td>3</td>
</tr>
<tr>
<td>16. Army General Classification Test</td>
<td>1</td>
</tr>
<tr>
<td>17. California Capacity Questionnaire</td>
<td>1</td>
</tr>
</tbody>
</table>

### TABLE IV. PERSONALITY TESTS.

<table>
<thead>
<tr>
<th>Name of Test</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Meehey Problem Checklist</td>
<td>15</td>
</tr>
<tr>
<td>2. S.R.A. Youth Inventory Checklist</td>
<td>9</td>
</tr>
<tr>
<td>3. Minnesota Counseling Inventory</td>
<td>2</td>
</tr>
<tr>
<td>4. Tarmin Personality Test</td>
<td>1</td>
</tr>
<tr>
<td>5. Evaluation and Adjustment Scales</td>
<td>1</td>
</tr>
</tbody>
</table>
checked twenty times. The scoring ease of the Kuder as compared to the scoring difficulty of the Strong perhaps accounts for much of the seemingly greater use of the Kuder. A total of six vocational tests were reported, and can be found in Table V below.

### TABLE V. VOCATIONAL TESTS.

<table>
<thead>
<tr>
<th>Name of Test</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Kuder Preference Record</td>
<td>90</td>
</tr>
<tr>
<td>2. Strong Vocational Interest Blank</td>
<td>20</td>
</tr>
<tr>
<td>3. Minnesota Clerical Test</td>
<td>6</td>
</tr>
<tr>
<td>4. Montana Employment Aptitude Examination</td>
<td>4</td>
</tr>
<tr>
<td>5. Bennett Test of Mechanical Comprehension</td>
<td>2</td>
</tr>
<tr>
<td>6. Crawford Small Parts Dexterity Test</td>
<td>1</td>
</tr>
</tbody>
</table>

Table VI, on the next page, contains a list of questions which were asked Montana public school counselors concerning uses of test results. They are arranged in rank order with a percentage of the times they were reported.

Indications are the greatest use of test results involves curriculum planning for students. Seventy and three tenths percent of the counselors replied in the affirmative to the question: A student's curriculum is partly determined by test scores. In 69.4% of the cases, test results were interpreted to students and on 53.9% of the returns, parents received explanations of test scores in most or all cases. This would substantiate the fact that 54.7% of the counselors use test results as the basis for counseling interviews.

In 63 of 123 cases, or 49.2%, school teachers were allowed to examine test results and form their own conclusions. The use of test scores as a
<table>
<thead>
<tr>
<th>Uses</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. A student's curriculum is partly determined by test scores.</td>
<td>90</td>
<td>79.3</td>
</tr>
<tr>
<td>2. All test results are interpreted to students.</td>
<td>76</td>
<td>69.4</td>
</tr>
<tr>
<td>3. Intelligence tests are used for predicting future academic success.</td>
<td>70</td>
<td>54.7</td>
</tr>
<tr>
<td>4. Test results form the framework for the counseling interview.</td>
<td>70</td>
<td>54.7</td>
</tr>
<tr>
<td>5. Test results are interpreted to parents in most or all cases.</td>
<td>69</td>
<td>52.9</td>
</tr>
<tr>
<td>6. Teachers are allowed to examine test results and form their own conclusions.</td>
<td>63</td>
<td>49.2</td>
</tr>
<tr>
<td>7. Test scores are used to identify potential dropouts.</td>
<td>57</td>
<td>44.5</td>
</tr>
<tr>
<td>8. Ability grouping is done on the basis of test scores.</td>
<td>50</td>
<td>39.0</td>
</tr>
<tr>
<td>9. Advanced courses are determined somewhat by the use of test results.</td>
<td>45</td>
<td>35.2</td>
</tr>
<tr>
<td>10. Test scores are used as a limiting factor in determining that courses should be taken.</td>
<td>42</td>
<td>32.3</td>
</tr>
<tr>
<td>11. Test results are used in decisions of retention or promotion.</td>
<td>36</td>
<td>28.1</td>
</tr>
<tr>
<td>12. Test scores determine college or non college preparatory curriculum.</td>
<td>27</td>
<td>21.1</td>
</tr>
<tr>
<td>13. Test results form the basis for most parent and teacher conferences.</td>
<td>24</td>
<td>18.7</td>
</tr>
<tr>
<td>14. Local norms are established for all tests given.</td>
<td>23</td>
<td>17.9</td>
</tr>
<tr>
<td>15. Vocational choices of students are determined by test results.</td>
<td>22</td>
<td>17.2</td>
</tr>
<tr>
<td>16. Test results determine the selection of higher institutions of learning to be attended.</td>
<td>19</td>
<td>14.8</td>
</tr>
<tr>
<td>17. Entrance requirements in the elementary school are determined by the use of test results.</td>
<td>17</td>
<td>13.3</td>
</tr>
</tbody>
</table>
TABLE VI, CONTINUED

<table>
<thead>
<tr>
<th>Uses</th>
<th>( \bar{x} )</th>
<th>( n )</th>
</tr>
</thead>
<tbody>
<tr>
<td>19. Achievement test results are used as a means of measuring teacher effectiveness.</td>
<td>14</td>
<td>19.9</td>
</tr>
<tr>
<td>20. Test results are used in determining causes of chronic absenteeism.</td>
<td>14</td>
<td>19.9</td>
</tr>
<tr>
<td>21. Students are permitted to select vocational subjects on the basis of test results.</td>
<td>11</td>
<td>5.6</td>
</tr>
<tr>
<td>22. Test results are used for identifying the emotionally disturbed.</td>
<td>11</td>
<td>5.6</td>
</tr>
<tr>
<td>23. Guidance test scores enter in the selection of students for honors.</td>
<td>9</td>
<td>7.0</td>
</tr>
<tr>
<td>24. Test results are a diagnostic tool for determining disciplinary actions.</td>
<td>6</td>
<td>5.7</td>
</tr>
<tr>
<td>25. Test scores are used in a selection of students as leaders.</td>
<td>2</td>
<td>1.6</td>
</tr>
<tr>
<td>26. Students grades are assigned according to guidance test scores.</td>
<td>0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

tool in identifying potential dropouts was reported by 45.4% of the guidance departments, and 28.1% indicated standardized achievement tests were used in decisions of retention or promotion. Thirty-nine percent of the questionnaires indicated test scores were instrumental in ability grouping and in 15.3% of the cases reported, test results determined entrance requirements into school. A consideration in this respect is the variance in size of the guidance departments polled, some which operated only on the secondary level and others which serve grades K-1 through K-12.

Among other results, 10.9% of those reporting stated that standardized achievement test results were used as a means of measuring teacher effectiveness. Intelligence tests were used 54.7% of the time for predicting future
academic success. Forty two questionnaires indicated test scores were used as a limiting factor in determining what courses should be taken. Counselors reported local norms established for all tests given in 17.9% of the cases. Test scores were instrumental in the selection of students as leaders only 1.6% of the time and student grades were assigned according to guidance test scores in no cases. Referral should be made to Table VI for additional ranks and percentages.
CHAPTER IV
SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

Summary

This survey was conducted to determine what standardized tests are most commonly used and for what purposes test results are used. To accomplish this, a list of the guidance departments was obtained from the State Department of Instruction, which listed 157 departments operating in public schools in Montana. A survey questionnaire was sent to these 157 schools, asking them to list the tests they commonly use, and to indicate on the questionnaire the purposes for which they commonly use test results. Of 157 survey questionnaires sent, 123 schools responded for an 81.5% return.

The results of the data taken from the survey questionnaires is recorded in Chapter III. The various tests represented were separated into five categories, achievement, aptitude, interest, personality and vocational on the basis of groupings in Buros' book, Tests in Print. The tests in each category were ranked according to the number of times they were reported on the survey questionnaire. In determining the uses of test results, the statements from the survey questionnaire were ranked in order and a percentage computed on the basis of how many times each had been reported.

Conclusions

Research into the testing programs of Montana public school guidance departments indicates the following:

1. The Iowa Test of Educational Development is the most popular standardized achievement test series.

---

2. More achievement test batteries were reported than any other type of test.

3. The Differential Aptitude Test is more often used than other aptitude type batteries.

4. The most commonly used intelligence test is the Otis Mental Ability Test.

5. The Kuder Preference Record is most frequently used in the field of personality testing.

6. Few personality tests were reported in the study indicating little use of personality type tests in high school counseling.

7. The Kuder Preference Record dominates the field of vocational testing.

8. All tests reported in this study represent a rather small portion of the total number of available tests.

9. Some tests were reported exceptionally often, such as the Differential Aptitude Test, the Kuder Preference Record and the Otis Mental Maturity Test.

10. Many test results are used to plan and select student curriculums.

11. Test results do not form the basis for many parent teacher conferences.

12. Test results are often used as the framework for the counseling interview.

13. The amount of testing varies greatly from school to school.

Recommendations

The following recommendations may be made as a result of this survey:
1. Recognition of errors and other limitations associated with any testing situation should always be considered when using test results.

2. Local norms should be established for some tests, primarily achievement.

3. Under no circumstances should student achievement test results be used as the sole determiner of classroom teacher effectiveness.

4. Testing programs should not be done solely for testing itself but should improve instruction.

5. Interpreting test results to students, parents and school officials should be a goal of all testing programs.

6. Standardized test results should have little effect upon disciplinary cases, other than their diagnostic value.

7. Additional research is needed to determine why certain tests are used in preference to others.

8. The rapidly changing vocational world indicates the need for a modern interest test.

9. Additional research is needed to determine whether the standardized tests commonly used by Montana public school counselors coincide with those in other states.

10. Counselors should actively participate with teachers and administrators during in-service training sessions.

11. Counselors should interpret guidance test results to teachers in most cases rather than allow teachers to form their own conclusions about test scores, when they are unqualified.
BIBLIOGRAPHY
BIBLIOGRAPHY


Starch, Daniel and Millet, Edward C., "Reliability of Grading Work in Mathematics," School Review, April, 1913, Volume XXI.


Thurstone, L. L., "A New Concept of Intelligence," The Educational Record, July, 1925, Volume 17, Number 3.

Thurstone, L. L., "A New Concept of Intelligence and a New Method of Measuring Primary Abilities," The Educational Record, October, 1926, Volume 17, No. 10.

Thurstone, L. L. and Thelma Guinn, "The 1936 Psychological Examination," The Educational Record, April, 1936, Volume 18, Number 2.


APPENDIX A

Sample of Questionnaire

Dear Counselor:

HELP NEEDED! It is my desire as a student at Montana State College in the field of Education to conduct a survey pertaining to the kinds and numbers of tests commonly used in your guidance department. I am also interested in determining the extent to which test results are used. Will you kindly take a minute of your time to check the proper blank for each test used, and at the bottom of the sheet add tests used but not included in the list.

1. Strong Vocational Interest Blank
2. Kuder Preference Record
3. Minnesota Clerical Test
4. Bennett Test of Mechanical Comprehension
5. Differential Aptitude Test
6. General Aptitude Test Battery
7. Scholastic Aptitude Test
8. Otis Mental Ability Test
9. Wechsler Intelligence Scales
10. Kuhlman-Anderson
11. Stanford-Binet
12. California Mental Maturity Test
13. Homan-Hansen
14. Cooperative School and College Ability Tests
15. Iowa Educational Development Test
16. SRA Youth Inventory Test
17. Money Problem Checklist
18. Iowa Achievement Tests
19. California Achievement Tests
20. Stanford Achievement Tests

Other:

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________

________________________________________________________________________
Select from this list of statements ONLY those which fit your usage of test results.

1. ____Intelligence tests are used for predicting future academic success.
2. ____A student's curriculum is partly determined by test scores.
3. ____Ability grouping is done on the basis of test scores.
4. ____Test results form the framework for the counseling interview.
5. ____Test scores determine college or non-college preparatory curriculums.
6. ____Students are permitted to select vocational subjects on the basis of test results.
7. ____Test results are a diagnostic tool for determining disciplinary actions.
8. ____Students grades are assigned according to guidance test scores.
9. ____Test results are used in decisions of promotion or retention.
10. ____Test scores are used to identify potential dropouts.
11. ____Test results are interpreted to parents in most or all cases.
12. ____Test scores are used in a selection of students as leaders.
13. ____All test results are interpreted to students.
14. ____Local norms are established for all tests given.
15. ____Test results determine the selection of higher institutions of learning to be attended.
16. ____Teachers are allowed to examine test results and form their own conclusions.
17. ____Test results are used for identifying the emotionally disturbed.
18. ____Advanced courses are determined somewhat by the use of test results.
19. ____Guidance test scores enter in the selection of students for honors.
20. ____Test results are used in determining causes of chronic absences.
21. ____Entrance requirements in the elementary school are determined by the use of test results.
22. ____Achievement test results are used as a means of measuring teacher effectiveness.
23. ____Test results form the basis for most parent and teacher conferences.
24. ____Vocational choices of students are determined by test results.
25. Test scores are used as a limiting factor in determining what courses should be taken.

Other Uses: ____________________________________________

____________________________________________________________________

____________________________________________________________________

Comments: ____________________________________________

____________________________________________________________________

____________________________________________________________________

____________________________________________________________________