

1983

## Standardized achievement tests and their implications for guidance counselors

John William Hansen  
*University of Northern Iowa*

*Let us know how access to this document benefits you*

Copyright ©1983 John William Hansen

Follow this and additional works at: <https://scholarworks.uni.edu/grp>



Part of the [Education Commons](#)

---

### Recommended Citation

Hansen, John William, "Standardized achievement tests and their implications for guidance counselors" (1983). *Graduate Research Papers*. 2495.  
<https://scholarworks.uni.edu/grp/2495>

This Open Access Graduate Research Paper is brought to you for free and open access by the Student Work at UNI ScholarWorks. It has been accepted for inclusion in Graduate Research Papers by an authorized administrator of UNI ScholarWorks. For more information, please contact [scholarworks@uni.edu](mailto:scholarworks@uni.edu).

---

## Standardized achievement tests and their implications for guidance counselors

### Abstract

In the last twenty years the general public has challenged the judgment of professional educators regarding the quality of education provided in public schools (Feldt and Forsyth, 1973). At present, the public is demonstrating an interest in knowing how well students are mastering important academic content and skills. Reassurances that educational expenditures are worthwhile are no longer being accepted in the absence of factual evidence. As a result, there has been increased attention given to the problem of the evaluation of student performance in educational programs.

STANDARDIZED ACHIEVEMENT TESTS AND THEIR  
IMPLICATIONS FOR GUIDANCE COUNSELORS

---

A Research Paper  
Presented to  
the Department of Educational Administration  
and Counseling  
University of Northern Iowa

---

In Partial Fulfillment  
of the Requirements for the Degree  
Master of Arts in Education

---

by  
John William Hansen  
May 1983

This Research Paper by: John William Hansen  
Entitled: STANDARDIZED ACHIEVEMENT TESTS AND THEIR  
IMPLICATIONS FOR GUIDANCE COUNSELORS

has been approved as meeting the research paper require-  
ment for the Degree of Master of Arts in Education.

William Kline

4/6/82  
Date Approved

\_\_\_\_\_  
Director of Research Paper

J. F. Kimball

4-8-83  
Date Received

\_\_\_\_\_  
Second Reader

Robert T. Lembke

4/8/83  
Date Received

\_\_\_\_\_  
Graduate Faculty Adviser

Robert Krajewski

4/25/83  
Date Received

\_\_\_\_\_  
Head, Department of  
Educational Administration  
and Counseling

## Table of Contents

Chapter	Page
1. Introduction . . . . .	1
Issues . . . . .	2
Definition of Standardized Achievement Test . . . . .	4
Philosophical Foundations . . . . .	5
A History of Standardized Achievement Testing . . . . .	8
Philosophies of Testing . . . . .	9
Current Status of Standardized Achievement Testing . . . . .	12
Introduction . . . . .	12
Uses of standardized achievement test data . . . . .	14
Dangers of standardized achievement test data . . . . .	15
Counselor role in achievement testing . .	16
Summary . . . . .	17
Statement of the Problem . . . . .	18
2. Review of Related Literature . . . . .	20
Uses of Achievement Testing . . . . .	20
Effectiveness of Achievement Tests . . . . .	25
Misuses of Standardized Achievement Tests . . . . .	27
Counselor Responsibility in Testing . . . . .	34
Summary. . . . .	39

Chapter	Page
3. Recommendations and Conclusions . . . . .	41
Implications for Counselors . . . . .	42
Recommendations for Test Use . . . . .	44
Conclusion . . . . .	46
Bibliography . . . . .	48

## CHAPTER ONE

### Introduction

In the last twenty years the general public has challenged the judgment of professional educators regarding the quality of education provided in public schools (Feldt and Forsyth, 1973). At present, the public is demonstrating an interest in knowing how well students are mastering important academic content and skills. Reassurances that educational expenditures are worthwhile are no longer being accepted in the absence of factual evidence. As a result, there has been increased attention given to the problem of the evaluation of student performance in educational programs.

The issue of student evaluation is complex. While there appears to be a consensus that schools need to utilize improved evaluation procedures, there is little agreement on the origin or degree of the issue. What has been agreed upon is that educators need information about students in terms of their measured interests and academic abilities if they are to assist them. There is also the expectation that evaluation data should be efficiently and cheaply acquired. In short, there is a consensus of public opinion that a simple, low cost method of obtaining

information is essential. This information can then serve as a means of determining how well students and schools compare against normative scores.

The implications that evaluative data has on the various groups of people involved in the evaluation process are important. For example, students want to know how well they have met the goals for particular coursework. For some students this kind of information is a prerequisite for admission to study in higher institutions of learning. Teachers use this information to determine relationships among ability, educational practices, and changes in pupil behavior to better predict and control teaching and learning. Counselors engage continuously in individual appraisal in their relationships with individuals. Students are helped by counselors to make decisions on the basis of their values and educational goals. Administrators use evaluative information in determining how well an educational program is meeting its goals. The public is also interested in viewing this information to see how well schools are educating children.

### Issues

There are three basic issues concerning the use of evaluative data in our schools today. The first is



accountability.

Accountability is a set of procedures that collate information about outcomes and costs to facilitate decision making. A major part of educational accountability is defining objectives, developing practices that enable these objectives to be met, assessing how well they are met, and reporting the outcome to decision-makers and the public (Shertzer and Linden, 1979, p. 533).

There is a tendency to equate the whole of schooling with what is currently measurable by student evaluation instruments. This process is likely to continue.

Another issue concerns minimal competencies. Once students graduate from school they should be able to handle everyday necessities of life (Shertzer and Linden, 1979). Schools have the responsibility of determining if students have mastered these competencies. Therefore, some minimal standards must be met before evaluative data can be interpreted. The difficult tasks are deciding what the minimal standards will be and whether they have been met.

A third issue, which is related to minimal competencies, is basic skills. Basic skills, which include reading, writing, and math, must be mastered so that an individual can succeed in today's world. Again evaluative data can be examined to see how well the skills have been mastered.

### Definition of Standardized Achievement Test

The outcome of the increased emphasis in accountability, minimal competencies, and basic skills is a search for ways of obtaining evaluative data on students. Standardized achievement tests are one way educators are working with this gigantic task. Mitchell (1981) defines an achievement test as "a test that measures the extent to which a person has 'achieved' something, acquired certain information, or mastered certain skills--usually as a result of planned instruction or training" (p. 1). He also states that an achievement test becomes standardized when it "provides a systematic sample of individual performance, administered according to prescribed directions, scored in conformance with definite rules, and interpreted in reference to certain normative information" (p. 7). The tests are generally commercially published and usually provide important information on reliability and validity. Mitchell (1981) defines reliability as "the extent to which a test measures what it is supposed to measure" (p. 6). Reliability is the degree of freedom of error in measurement. Mitchell (1981) defines validity to be "the extent to which a test does the job for which it is used" (p. 8). For achievement tests, validity is the extent to which the content of the test represents a

balanced and adequate sampling of the outcomes of the course or instructional program the test is intended to cover.

### Philosophical Foundations

According to Wingo (1974), the philosophy of education is

a process of inquiry into ideas and basic beliefs that will enable us to form reasoned attitudes about the important issues of our time. Education, being the complex kind of social enterprise that it is, has psychological, sociological, and political dimensions that are of great importance (p. 22).

A serious conception of educational philosophy, then, is looking at the aims and goals of education as viewed by noted philosophers of education. One of the functions of educational philosophy is to inquire into ideas and ideals of education, discover what they are, subject them to critical scrutiny, and judge their adequacy. These functions play a role in the investigation into why professional educators obtain evaluative data such as standardized achievement tests.

Although there are numerous traditions involved in educational philosophy, only two will be noted in this study. One of these is known as the conservative tradition in education, or essentialism. The conservative believes that only certain subject matter is worthy to be

taught in the schools. The conservative also believes that the school should cherish and transmit certain traditional values, and make no attempt to maintain neutrality about questions of value. "The purpose of education, according to essentialists, is the transmission of certain elements of the cultural heritage whose importance is so great that they cannot be neglected" (Wingo, p. 53). As far as the nature of the curriculum is concerned, conservatives feel that the curriculum should consist of a common core of subject matter and intellectual skills that are so essential they must be taught to all who come to school. "The progress of the child from one grade to another should depend on his mastery of the 'essentials' for a given grade" (Wingo, p. 55). Therefore, most students in the conservative tradition study subjects that include English, mathematics, history, science, and foreign languages. Achievement testing is one way to evaluate how well this subject matter is being mastered.

Another tradition in education is known as the liberal tradition or progressivism. This view of education protests against the fundamental character of educational conservatism and its domination of practical educational affairs in most schools. Liberals see education as a necessity for a free people. They

envision the school as the vanguard of social progress. According to the liberal point of view, an educational institution should always be close to the people and constantly be in the mainstream of social life. Liberals also feel that the educational process has grown dull and very much mechanical; that individual differences are being ignored and the grading system is creating an emphasis for uniformity. Wingo (1974) states that the liberalists

were protesting the idea of education as primarily transmission of subject matter; the authoritarian character of an institution whose primary role should be the extension of democratic ideals; and the failure of the school to sponsor social betterment for the masses (p. 152).

Curriculum design in a typical progressive school involves a program of active work--physical and intellectual--rather than the passive absorption of subject matter. Children should have experiences in school that are built on the experiences they have in home and community. John Dewey (1916), a noted liberal in education, felt that the curriculum should be conceived as an ordered series of occupations. He also felt that there should be no lessons, no assignments, and no examinations. In this way the child has room and power to grow into the adult mold. Achievement testing would therefore not play an active role in a progressive school curriculum.

A consideration of various schools of educational philosophy in terms of the concept of evaluation brings forth real issues. More specifically, the essentialist and progressivist views provide relevant and significant questions regarding the role of testing. The question becomes one of values concerning the individual or the educational process.

#### A History of Standardized Achievement Testing

While many great philosophers, such as Plato and Aristotle, proposed theories about evaluation and human nature, formal mental measurement is of recent origin. Thorndike and Hagen (1969) have sorted the history of measurement into five parts covering the years 1900 to the present. These periods are:

1. (1900-1915) This period was known as the pioneering phase of mental measurement. It marked the origin, development, and expansion of many measurement methods.

2. (1915-1930) This period was classified as the boom period in test development. Standardized tests of ability, achievement, personality, and interests were expanded and used indiscriminately.

3. (1930-1945) This period was a time of critical appraisal. Attention shifted from measuring a limited range of academic skills to evaluating achievement of the range of educational objectives.

4. (1945-1960) This period saw an increase in testing programs in the schools. Integrated test batteries for educational and personnel use also increased in size and number.

5. (1960-present) This period can be viewed as a time of public controversy. Issues derived from test use have been taken to court for settlement. Professional associations, test publishers, and testing practitioners are engaged in serious public examination and discussion of the merits, limitations, benefits, and adverse consequences of tests and testing.

### Philosophies of Testing

Cronbach (1970) has identified two philosophies of testing. One philosophy, called the psychometric approach, obtains a numerical estimate of a single aspect of performance. The psychometric approach "deals with 'things', e.g., with distinct elements or traits which have a real existence. All people are considered to possess the same traits (e.g., intelligence or mechanical experience), but in different amounts" (Cronbach, p. 29).

This American-based concept looks at the amount of a distinguished trait and the relationship between and among traits. The psychometric approach is definite and structured, tending to follow the conservative point of view. Achievement testing is also based on this approach to testing.

The second approach, called impressionistic, "seeks to develop a comprehensive picture of an individual by examining significant cues by any available means and integrating them into a total impression" (Cronbach, p. 29). Studying one trait or element at a time is no substitute for considering the person as a whole. The impressionist is not satisfied with knowing how much of some ability a person has; instead, he asks how the subject expresses that ability, what kinds of errors they make, and why. The impressionistic approach relies on observation, descriptive data, and self-report, indicating the liberal point of view.

One difference in the two approaches may be found in the definiteness of the tasks employed. The test designer decides precisely how the task is to be explained to test takers. This involves structuring the test so that all interpretations of the task are the same. Structuring the task permits a definite answer on the test, whereas a less structured technique allows greater



variation of response. Achievement tests tend to have a good deal of structure.

Another difference in the two approaches deals with control of response. Tests are categorized as free response tests, that follow the impressionistic approach, or controlled response tests, that follow the psychometric approach. The free response test permits observations which illuminate the scored aspect of performance. The controlled response test can be more objectively scored, may depend less on expressive skill, and be less subject to misinterpretation. Achievement tests are an example of a controlled response test.

A third difference in the two approaches deals with the product of testing. Psychometric testing concerns itself with the tangible product of performance, such as, the concrete answer. The impressionistic approach watches the student at work in order to form a general opinion. Achievement testing involves looking at what answers are given rather than making observations of students taking the test. Interpretation of test results is also different in the two approaches. Since the impressionistic style deals with observation, this approach assigns a greater deal of responsibility to the test interpreter. One of the reasons achievement testing

is popular is simply because interpretation of the results is easier to accomplish.

The issue concerning the best testing approach is hard to determine. Basically, testing varies according to the amount of structure involved in the process. Schools today use a mixture of testing that involves both approaches.

### Current Status of Standardized Achievement Testing

Introduction. Achievement tests are designed to measure the progress students have made as a result of training. According to the Educational Testing Service, achievement tests can be classified into three groups ("Large Scale Programs", 1958). These groups include (1) end of course achievement tests that measure specifically what a student has learned in a particular subject, (2) general achievement tests that cover a student's learning in a broad field of knowledge and can be given to students who have taken quite different courses of study within a field, and (3) tests that measure critical skills a student has learned and his or her ability to use these skills in solving new problems.

Shertzer and Linden (1979) classified achievement tests as formative or summative measures. Formative measures include the use of achievement tests to assess

progress in the development of knowledge and skills before a course or unit of instruction. Summative measures are those given at course or unit termination to assess the outcome of instruction. Achievement tests can also be either criterion-referenced or norm-referenced. Mitchell (1981) states that a criterion-referenced test is designed to assess an examinee's mastery of fundamental skills or knowledge without reference to the performance of others. Norm-referenced tests tell where a person stands when compared to a population who have taken the test. Most standardized achievement tests used today are norm-referenced although there has been a "recent emergence of criterion-referenced tests" (Shertzer and Linden, p. 178).

There are many standardized achievement tests used today. A partial list includes:

1. Iowa Tests of Basic Skills
2. Metropolitan Achievement Tests
3. The Science Research Associates Series
4. Stanford-Achievement Tests
5. Iowa Tests of Educational Development
6. Sequential Tests of Educational Progress
7. Science Research Associates High School Placement
8. Tests of General Educational Development

Many sources of information have been developed for those interested in tests. These include the mental measurement yearbooks developed by Oscar K. Buros, catalogues and manuals provided by test publishers, and the reviews of newly published tests in educational and psychological journals.

Uses of standardized achievement test data. Shertzer and Linden (1981) discuss four primary uses of achievement test data. The first use is to promote client self understanding. The data should be used to enable the students to learn about themselves. The results make it possible for the clients to reflect on their performance in light of their personal characteristics. A second use is to help in decision making. Data about an individual's performance can be used to support or reject alternatives. Encouraging further self-study is a third use of the data. Testing data, coordinated with other information, encourages the clients to learn about themselves. A fourth use involves assistance in diagnosis. Data produced by tests can help diagnose school problems of individual students and can help plan corrective or remedial work. Test information can also help in correcting policies in certain educational programs to benefit the students.

Dangers of standardized achievement test data.

Standardized testing has been a controversial issue from 1960 to present. Debate on the use of tests and their development is widespread. Shertzer and Linden (1979) outline major criticisms of achievement testing and testing in general. One criticism is that tests label and predetermine individuals early in their school career which continues into adulthood. Another criticism is that tests are imperfect in measuring what they are supposed to. Some critics feel that tests are biased towards white middle-class experience, achievement, and personality--including a bias towards sex. Criticisms that tests obscure a child's talents and discourage the student from developing other qualities have also erupted. An invasion of privacy is a criticism of testing due to test results being filed in a cumulative folder that follows an individual through school and into work. Numerous people, therefore, have access to score results. Interpretation of tests can also be a problem. Critics charge that students and parents might misunderstand test interpretations and use the data in harmful ways. Critics also feel that testing provides students with test taking skills rather than developing skills for true learning. These skills stress competition and success rather than cooperation. Decision

making becomes a problem in achievement testing in that test data is used to make decisions without regard to other factors such as grades and conduct reports, etc. A final criticism is that tests control schools. Educators criticize test publishers and companies for dictating what is being taught in the classroom.

Counselor role in achievement testing. Secondary schools now accept the responsibility for providing each student with counseling, educational and career. Achievement testing programs need qualified counselors to administer and interpret achievement tests in order to guarantee appropriate test usage. Well-trained counselors must use tests that are valid for both the ostensible purpose and the person to be evaluated. Counselors must be familiar with research literature on tests and be able to evaluate the test's technical merits. When administering the tests, the counselor must be aware of conditions that might affect a person's performance. A counselor should also make recommendations about a student only after considering the test scores in light of other important information such as grades, teacher recommendations, and work records. Above all, the counselor should have adequate expertise in order to guard against unwarranted interference.

A counselor, well-trained in test theory and practice, can develop a testing program which meets the student's specific needs. Daley (1959) presented four principles that should characterize school testing programs. Firstly, the program should be continuous. Effective use of tests is only possible when they are part of a continuing program that permits measurement of growth and systematic evaluation of change. Secondly, the testing program should be comprehensive. Spot testing is of value but the results of all tests are enhanced when they are a part of a comprehensive evaluation program. Thirdly, is that decisions relevant to selection, scheduling, test result reporting, and other aspects of the program should be made jointly by teachers, counselors, and administrators. Finally, the testing program should be integrated with the total education program. Standardized testing must be consistent with instructional goals and counseling activities.

Summary. Dr. William Glasser (1969) in Schools Without Failure points out that standardized tests are based on what he calls the certainty principle.

According to the certainty principle there is a right and a wrong to every question; the function of education is then to ensure that each student knows the right answers to a series of questions that educators have decided are important. Children who come to school with the idea

that many questions have more than one possible answer soon get over the idea (p. 36).

In 1980, Dr. Leonard S. Feldt, director of the Iowa Tests of Educational Development, spoke to members of the Muscatine-Scott County (Iowa) Advisory Committee on Group Testing. He stated that achievement tests helped teachers and counselors keep themselves more intimately and reliably acquainted with the educational development of each student. The tests also helped administrators evaluate the total educational offering of the school. These two statements represent two different opinions regarding the worth of standardized achievement testing. Guidance counselors use achievement testing to obtain information about students. "As a matter of fact, so strong is the dependence of the guidance program on testing, that in some high schools testing has come to be identified almost exclusively with the guidance function" (Lennon, 1968, p. 2). The issue concerning the value of standardized achievement testing in our schools should therefore be of value to all conscientious guidance counselors.

#### Statement of the Problem

Standardized tests of achievement continue to be used in assessment practices. Accountability, minimal



competencies, and basic skills constitute the primary rationale for the testing process. Public outcry over standardized achievement tests and their use has increased. In sharp contrast, professionals have reacted to public opinion with outspoken criticism of the testing process, citing the dangers and abuses inherent in depending on tests for student assessment. Responses to these criticisms must be found if achievement testing is to continue to be a valid process in schools. Since counselors use tests a great deal they should be concerned with the role that standardized achievement testing will have in the future.

The resolution of the achievement testing controversy necessitates responding to two primary questions. These questions are: (1) What do standardized achievement tests actually do? and (2) What are the advantages and dangers of using standardized achievement tests? The goals of this study are to look at the counselor's role in the achievement testing process and to see what counselors can do to assure that achievement tests are administered and interpreted ethically and appropriately.

## CHAPTER TWO

### Review of Related Literature

The value of standardized achievement testing is directly related to the diligence and prudence with which the results are interpreted and used. Achievement testing is an important factor in American education; no one can afford to overlook testing or be indifferent to test results. Therefore, professional educators need to look at standardized achievement testing in a concise and responsible manner. There is a need to examine the uses of test scores to see if the tests are being used appropriately. Educators need to research the effectiveness of achievement tests to determine whether the tests fit their educational purposes. The dangers and misuse of achievement testing must be confronted to discover what needs improving. Finally, the counseling role in the testing process must be studied to see what possible revisions can be made. The review of literature will involve these important areas of standardized achievement testing.

#### Uses of Achievement Testing

Standardized achievement testing assists educators in the educating of students in many ways. One application is the evaluation of the curricula and efficiency of a

school (Kirkland, 1971; Postel, 1983; Beck, Note 1; Travers, Note 2; Feldt and Forsyth, Note 3). Beck (Note 1) describes two methods achievement tests utilize in the evaluation of schools. The first is to assess individual school performance, thus determining whether the school is providing quality instruction that meets the needs of the students. A second method is to compare performance among schools. Schools are ranked according to achievement test scores so that schools with lower scores can set up programs to improve efficiency.

Another use of achievement testing is to identify students who need special diagnostic study (Kirkland, 1971; Postel, 1983; Beck, Note 1; Travers, Note 2; Feldt and Forsyth, Note 3). Students may show irregularities in achievement in particular courses and even in the same course from time to time. Achievement tests data does not automatically simplify the complex interaction of factors that produce irregular performance. When disparities are found between test performance and behavior, then explanations in a student's response to environment and experience must be examined.

Achievement test data also helps students set educational and vocational goals (Kirkland, 1971; Beck, Note 1; Travers, Note 2). Students need help in selecting courses and subjects to study. The content of

standardized achievement tests helps considerably in this matter. Feldt and Forsyth (Note 3) identify the skills found on the Iowa Tests of Educational Development. These include the ability to recognize the essentials of good writing, resolve quantitative problems, weigh discussions of social issues critically, recognize sound methods of scientific inquiry, perceive the subtle meanings and moods of literary materials, and use sources of information. Once students know their scores on achievement tests, counselors can help them determine strengths and weaknesses in different skill areas. This information can be integrated with the employment or educational goals of the student and help in setting goals for the student's future.

Test data also provides information for parents, community, and outside agencies such as employment agencies (Kirkland, 1971; Postel, 1983; Beck, Note 1; Travers, Note 2). Professional educators and parents at home should have a common understanding of the ability and achievement of students. Most parents are highly interested in learning how their child is doing. Parents can also make significant contributions in helping develop and improve a child's skills. If these contributions are to be positive, parents must be consulted and informed of their child's development. The most common way of dis-

cussing test data with parents is through the parent-teacher-counselor conference. The parent-school partnership should be a positive partnership in the child's school years.

The participation of the community in school matters has also increased tremendously. Testing information is an excellent means of keeping the community informed regarding how well the schools are educating students. When the residents in a school district want to know the status of their schools, one possible method of reporting results is through the media. Many schools also provide newsletters which can be used in supplying test information. The concern of the community over the state of their schools represents positive regard toward any improvements that might be needed. Standardized achievement testing is a viable approach of letting the community know the status of their schools.

Achievement test data also provides a means for grouping students by ability (Kirkland, 1971; Travers, Note 2). A popular method, used by most schools, is assigning students to classrooms that contain students of similar ability. This is especially true of students entering a new school. A counselor can investigate a student's cumulative file, interpret his past test scores, and decide which subjects best fit the student's

ability. Kirkland (1971) states a related use of standardized testing. She finds that the test use most visible to the public is the use of scores to determine admission to higher institutions of learning. "People generally view all achievement tests as accurate and are quite willing to accept test results as lawful for this purpose" (Kirkland, p. 304). In fact, in a 1969 survey of high school students, 52 percent placed considerable importance on standardized achievement tests for success in life after graduation (Brim, 1969).

Dr. Steven Holbrook, director of research, evaluation, and development for the Waterloo Community School District, Waterloo, Iowa, finds that test data can be used to improve teaching, but must be done in an informal manner. Many teachers resist having instructional effectiveness tied to test scores. A greater understanding of testing and test data among teachers, counselors, and administrators has helped alleviate these apprehensions (Postel, 1983). "Test data can help teachers improve the educational process" (Postel, 1983, p. E-8). Teachers can view test data in deciding what content areas to emphasize in their instruction.

In summary, the uses of standardized achievement tests are numerous. These uses include promotion or retention of individual students, classification and

placement of students in different classes, evaluation of student achievement, and diagnosis of achievement for improvement of future schoolwork. "Standardized testing programs will continue to be used to measure the degree of change in student proficiency" (Postel, p. E-8). A summary of the uses of standardized achievement testing finds most authors agreeing with this assertion. However, with the high cost of testing, in both time and money, questions on the value of testing and effectiveness are still raised.

#### Effectiveness of Achievement Tests

A discussion of the values of achievement testing can be divided into three areas. These three areas are test accuracy, test objectivity, and test comparability (Turnbull, 1978). According to Turnbull, test accuracy deals with the precision in testing in regard to the use of tests. More importantly, Turnbull states that a prediction of student achievement based on quantitative data is more likely to be accurate than prediction of achievement based on clinical observation or subjective judgment. "Testing provides a balance against human error" (p. 292). Turnbull concludes that teacher judgment is important, but finds that it should not be the only factor involved in judging student improvement in skills.

A second category deals with the objectivity of standardized achievement tests. Turnbull (1978) states that achievement test scores are considerably less biased than test scores that teachers give. There are no "teacher's pets" involved in achievement test scoring (p. 291). Turnbull concludes that achievement test scores will be highly accurate when compared to test scores graded by teachers. When teachers mark scores, they often have a favorite student whose scores are marked slightly higher. Scores may also be marked lower for some students. Many factors, such as sex, race, and honesty, might play a role in the teacher's decision of a test score for certain individuals. When machines are used to score standardized tests, these factors are minimized and the scores increase in accuracy.

A third category regarding the virtues of standardized achievement test is comparability (Turnbull, 1978). Comparisons are needed in testing to denote deficiencies in students and schools. Comparisons between the educational achievement in states and individual students can be noted in order to see where improvement is needed. Other types of comparison include grade by grade comparisons, subject by subject comparisons, and year to year comparisons. "There is no other method of measurement that is better than the achievement test when comparing test scores



between students, classes, schools and among states" (Turnbull, 1978, p. 296).

In summary, Haag (Note 4) defines six issues where standardized achievement testing proves effective. These are:

1. Participation of students, parents, teachers, school boards, and local community citizens in decisions about what the schools should teach.
2. An increase on individualization of instruction.
3. An increase in cross disciplinary education.
4. Accountability.
5. Long range planning and systematic control of educational development.
6. Effective teacher evaluation.

Haag (Note 4) states that achievement tests are the most consistent measuring device on the market. Encouragement of individual efforts and detection of individual deficiencies would be more difficult without effective standardized achievement tests.

#### Misuses of Standardized Achievement Tests

Standardized tests have been inappropriately used in a variety of ways. A frequently occurring problem in achievement testing is the misinterpretation of test results (Brim, 1965; Hawes, 1973; Kirkland, 1971;

Frechtling, Note 5). One issue involves data from tests designed to compare groups used instead for individual comparison. Questions arise about group data being used for decisions on individuals, and whether it is accurate. A second problem contends that educators and the public tend to overrate the information provided by standardized test scores. The publication of school by school test scores always makes headlines in local newspapers and leads readers to draw rapid and sometimes inaccurate inferences about school and individual performances. The problem, in this regard, is that achievement testing becomes the sole-criterion for decision making (Brim, 1965; Howe, 1980; Beck, Note 1; Frechtling, Note 5). Beck (Note 1) further contends that media, boards of education, legislators, and others should not be criticized for overinterpreting test scores when no other evidence is provided. Beck reports a failure in reporting test results to all concerned groups. A survey, done by Beck in 1979, found that 90 percent of school students would like to know their test scores, but fewer than 40 percent of the teachers and counselors reported the scores directly to the students. Beck (Note 1) concludes that the fact that students gave less than 100 percent on achievement tests was not surprising.

Turnbull (1978) lists three factors that have encouraged the misinterpretation of tests. The first

he calls the micrometer fallacy. Turnbull states that "people have invested test scores as infallible" (p. 292). They feel testing is accurate. The second factor is called the whole-person fallacy. Turnbull defines the fallacy as "the tendency for people to read into achievement test scores much more than the amount a student has learned in a given subject" (p. 293). He concludes that some people read characteristics such as honesty, leadership, and social consciousness into the scores. The third factor is known as the equal preparation fallacy. This fallacy is where "people eye tests to compensate for the differences in academic development of children whose learning capacities have differed dramatically" (p. 293). The test scores tell little about the difficulties a student has overcome to acquire a given level of proficiency. Achievement tests do indicate the student's level of achievement.

Echternacht (Note 6) criticizes the interpretation of summative testing results without any formative testing results being available. He traces this over-reliance on summative evaluation results to two factors. First, with the current emphasis on using evaluation results, people are going to use the information they have available which includes summative data. Secondly, many people only want to look at the bottom line, which includes

summative data. These people feel that formative evaluation is superfluous. Echternacht (Note 6) contends that formal and summative testing data should be combined in improving instruction. Formative evaluation is viewed as a prerequisite for summative evaluation.

Another interpretation problem is an over-reliance on grade equivalent scores (Davis, 1972; Echternacht, Note 6). Grade equivalent scores were intended to help interpret tests by providing a reference that most people could understand. Grade equivalent scores make following the progress through the grade levels easier. One problem with this score, though, is that the dispersion of grade equivalent scores systematically increases from grade to grade (Davis, 1972). As students pass into the higher grades, their levels of achievement and ability are widely spaced. Some students accelerate faster than others.

This fact forces an individual who is below the national median in grade one, but who maintains his relative standing in the national norms group from grade to grade, to lay further behind the national medians in terms of grade equivalent scores (Davis, p. 306).

Another factor concerning misleading grade equivalent scores is the use of means instead of medians in interpreting the scores. "Median scores are not expressed in interval units (e.g., units that represent

equal amounts of competence in a skill), yet this is assumed in the interpolation of mean scores based on them" (Davis, p. 307).

A final criticism, dealing with misinterpretation, is stated by Beck (Note 1). He concludes that people are pretending that norm referenced tests are diagnostic and criterion referenced. Beck states that one test cannot tell all things to all people. He contends that there is a fine line between using a test for as much information as it can reasonably yield and the over-interpretation of that information. Beck remarks that all professional educators are guilty of crossing this fine line.

There are many misuses of achievement tests that do not deal with misinterpretation. For example, Beck (Note 1) states that frequent changes in a test series can cause problems in a particular school district when checking progress across subject matter by grades. If the school uses two or three different test series at different grades, the possibility of making comparisons is eliminated. Beck (Note 1) also states that this happens when changing from an older to a newer edition of the same series. Another problem is that achievement tests are used for mandatory reporting (Echternacht, Note 6). If the evaluation is positive, then the evaluation is offered as evidence of effectiveness in the schools.

If the evaluation is negative, then the tests are filed and forgotten. Echternacht asserts that this is often a result of poor management, but underneath achievement testing in general lies a fundamental belief which asserts that test scores are not valid and should not be used for anything other than filling out required forms and reports. The preference is for observational and judgmental evaluations dealing with testing. Echternacht contends that anti-testing sentiment is characteristic of people who know their programs are good, and by many teachers and counselors who see tests as being a negative evaluation device for students who need positive educational experiences.

Many critics also conclude that achievement testing puts external pressure on students (Darehshori, 1977; Hawes, 1973; Kirkland, 1971; Postel, 1983). Darehshori states that this external pressure is especially evident in the lower grades. She contends that "in giving standardized tests, we place children in positions over which they have no control, then we direct them to perform illogical tasks, and to act as if everything was perfectly logical" (p. 16). She states that this damages the self esteem of children and predetermines their social status. "Some children deal with this by being absent during test time. Other children deal with achievement testing by

not really trying; by just marking answers and going through the motions" (Darehshori, p. 17). Darehshori concludes that children should be excluded from standardized testing until they actually have the skills of test taking. She also maintains that test manufacturers should design tests based on the developmental levels of young children and not adults. Consistent with this argument is another criticism revealed by Badal and Larson (1970). They argue that test scores will suffer because of unusual conditions in the school atmosphere during the administration of tests. Factors such as administering the tests next to a noisy cafeteria or gym, giving the tests on a Friday afternoon, or giving the test immediately before recess or gym class may be reflected by a lower score.

Some criticisms concern the basic function of the schools. One such criticism is that test publishers control the curriculum (Badal and Larson, 1970; Kirkland, 1971). Some tests may reflect what is taught in the school. Teachers might also teach what is on the tests; the resultant high scores serve to help them keep their job. Kirkland (1971) also contends that testing programs may interfere with the functions of the high school. Many achievement tests require numerous hours to complete. In most schools, a whole week of classes can

be devoted to testing. Counselors also spend considerable time in preparing and administering the tests and question if that time is well spent. Others question whether the high cost of testing is worthwhile (Kirkland, 1971).

In summary, Brim (1965) discusses four sources of criticisms towards standardized achievement testing. These include oppositions arising from a general system of values, personality characteristics, an individual's experience with the test, and oppositions arising from the restrictions of testing on life opportunities. Brim (1965) states that "testing doesn't occur in isolation. There is definitely a social context involved" (p. 125).

#### Counselor Responsibility in Testing

Counselors have been under increasing pressure to define their role in the use of standardized tests. The American Personnel and Guidance Association (APGA) developed a policy statement concerning the issues ("Responsibilities of Users", 1978). This statement was built on the assumption that "test data is neutral and that guidelines are needed to promote constructive use of the tests" (p. 5). The policy statement consisted of six sections dealing with the counselor's responsibilities for test use.



The first section dealt with defining the purposes for testing. The four purposes listed were:

1. Placement - If the purpose is selection or placement, the counselor must know about the programs or institutions in which the student may be placed and be able to judge the consequences of such placement or exclusion for the students.

2. Prediction - If the purpose is prediction, counselors deciding to test and interpret test results must understand the pitfalls of labeling, stereotyping, and prejudging people.

3. Diagnosis - If the purpose is diagnosis, the counselor should understand enough about the general domain being measured to be able to identify those aspects adequately measured and those that are not.

4. Growth and change - If the purpose is to examine growth, counselors need to understand the problems of measurement such as the unreliability of change measures, the pitfalls of using norm references, the problems of comparability, and the limitations of scoring scales such as grade equivalents.

The second section of the policy statement discusses test selection. Tests should be selected for a specific measurement purpose, use, and interpretation. The selection of tests should be guided by such considerations

as the characteristics of the population to be tested, the knowledge and skills to be assessed, the purposes for testing, and how the test scores will be obtained and interpreted. The responsible counselor will determine reliability and validity for a particular test. The counselor will also select tests to satisfy local use. Special attention on test design and the variation of motivation among students should be considered. Finally, the counselor needs to know the technical characteristics of tests, such as the ease and accuracy of scoring tests and communicating test results. Test cost and timing tests should also be considered, but not used as the primary criteria for test selection.

The third section in the policy statement sets forth the qualifications of test users. This section concludes that all professional personnel and guidance workers should have formal training in psychological and educational measurement and testing. Lack of proper qualifications leads to misuse, error, and damage to the students. The qualifications for test use by counselors depends on the particular role of the counselor, the setting in which the test use takes place, the nature of the test, and the purpose of testing ("Responsibilities of Users", 1978). The counselor's responsibility is to be knowledgeable about any test he plans to use.

Counselors are also responsible for the administration of the test. This responsibility includes presenting the test in the manner specified by the test developers, and conducting the test under conditions that maximize opportunity for optimum performance. The counselor should inform students that they will be tested. This orientation should describe the purposes and contents sampled by the test, how the test is administered, and how the scores will be reported and used. The counselor should be sure to present all tests in an identical manner to insure that the test is a fair and comparable demonstration of the performance of each individual taking the test. The counselor should also determine whether the testing environment is conducive to the best possible performance of the students. Factors such as seating, work surfaces, lighting, heating, and freedom from distractions should be considered. If the counselor demonstrates clear verbal articulation, calmness, positive anticipation, and an impartial treatment of students, then testing will be more efficient ("Responsibilities of Users", 1978).

The fifth section of the policy statement deals with test interpretation. "Test interpretation encompasses all the ways we assign value to the scores. For adequate test interpretation, a counselor requires knowledge in administration and scoring procedures, scores and norms,

reliability, and validity" (p. 5). Counselors should study the test manual to make interpretation easier. Related to interpretation of test results is the communication of those results. Communication consists of reporting data in a comprehensible and informative way. The counselor should report test data with a concern for clarity, objectivity, and accuracy. There must also be a concern for the right of the tested individual to be informed about how the results will be used for his or her benefit, who will have access to the results, and what safeguards exist to prevent misuse.

In summary, Shertzer and Linden (1979) list five guidelines for the counselor's use of achievement test data. They state that "achievement test data should be used cautiously" (p. 190). Counselors, and others, need to realize that achievement measures are not foolproof, and they do not always give dependable results even in the hands of experts. Counselors should also "supplement achievement test data with grades" (p. 190). People assessing individuals or interpreting achievement test scores cannot ignore teacher's marks, even though they must be aware of their limitations. A third guideline is to "search for patterns in achievement test performance" (p. 190). *A counselor should examine achievement test data and other performance indicators for relation-*

ships and trends in order to improve their interpretations. "By noting subscores, counselors can arrive at a conception of a pattern; having grasped a pattern, they can understand and reinterpret the meaning of the parts as well as the whole" (p. 191). A counselor must also remember that "test data represents individuals" (p. 191). The counselor's main goal is understanding the individual. "When interpreting scores to a student, a counselor must be concerned not only with a score on an achievement test, but more specifically, with the reasons why the score is high, average, or low" (p. 191). Finally, counselors must remember that "achievement test scores are no guarantee that the measured performance is a typical performance" (p. 192). A student who has put forth less than maximum effort or who has been under emotional strain, may produce a below-average performance. Records of achievement test performance, school marks, and other indicators of attainment, over a period of years, provide better analysis for counseling purposes and give a true indication of the level of achievement of individual students. Without a variety of information concerning students, weaknesses would be difficult to ascertain in student achievement.

### Summary

There is a considerable amount of literature concerning the use and misuse of standardized achievement tests.

Different opinions are that tests are good, bad, immoral, unfair, un-American, useless, and infallible. Tests, of course, are none of these things. A test is merely a sample of behavior taken under standardized conditions from which other behavior is inferred (Shertzer and Linden, 1979, p. 506).

Tests are devices for making observations. A problem develops when the users of test information make inferences from these observations. To allow test scores to outweigh other data or to ignore test scores in favor of other data is to misuse the scores. Counselors must apply stringent guidelines in using information gained from tests. Hill (1969) suggests that counselors ask questions when confronted with testing data. Such questions would include: "Is this test information true? Is this test information fresh? Is this test information developmental? Is this test information complete?" (p. 141). Test information never constitutes an end in itself, but is merely a tool to be used to attain important counseling goals. Counselors should therefore be looking for ways to improve the uses of testing. This improvement will have to come primarily from the ways in which counselors, teachers, and others use tests rather than from changes in the tests themselves.

## CHAPTER THREE

### Recommendations and Conclusions

An overview of achievement literature describes numerous issues in American education. "Standardized achievement test data are used toward such purposes in schools and other educational institutions as promotion or retention, classification and placement, or evaluation and diagnosis" (Shertzer and Linden, 1979, p. 188). Achievement tests provide teachers and counselors with a periodic objective description of student achievement that is economically feasible and convenient. Achievement tests also provide counselors with a dependable basis for judging the relative strengths and weaknesses of the academic portions of the educational offering. Finally, achievement tests aid students in educational placement and career planning and assist the students in understanding themselves.

Cronbach (1970) states that achievement tests are designed to measure the progress students have made as a result of training. He concludes that:

One significant contribution of standardized tests has been to break down the 'time serving' concept of education. A person's standing in school is frequently judged by the number of years he has put in, or the number of courses he has passed through. Time spent is no index of education received. In one study, where thousands of college students took standardized

tests of knowledge in various fields, many college seniors know less than the average high school senior. Since number of units accumulated tells little about proficiency, tests are being given increasing weight as evidence of educational development (p. 363).

Achievement tests have not changed the time spent aspect in education, but they definitely give an indication of the level of achievement the student has obtained.

The literature has also shown that standardized achievement testing has certain disadvantages. Test experts conclude that interpretation and communication of test results cause a majority of these problems. There are two primary reasons for not routinely interpreting test results to students and parents. Counselors are wary of reporting low scores to students who would be disturbed by such results. A child could also be labeled inferior, both in their own mind or their parents'. Counselors must therefore minimize the possibility of such problems.

#### Implications for Counselors

If counselors are to improve the interpretation and communication of test results, they must follow a definite set of guidelines. One guideline is to be sure that there is a clear and immediate goal in mind which serves as the reason for test interpretation. A second guideline stipulates counselors should never discuss the implication



of scores in terms of absolute answers to questions from students and parents. A counselor should speak in general terms and avoid the use of specific scores, whenever possible. A third guideline indicates counselors should never pose as experts. They should concentrate on increasing the understanding of scores and should not compare one student to another. A final guideline is that discussion of test interpretation should include other factors concerning a student's progress in school. Such factors could be absenteeism, physical well-being, pupil motivation, pupil mobility, and a complete description of school marks and efforts. Standardized achievement tests should never be the sole criterion for student progress.

An important issue to consider is what counselors can do to make test use more informative, reduce misuse, and improve attitudes towards achievement testing.

Neulinger (1966) states:

Attitudes toward the use of tests vary as a function of the context of test administration and furthermore, that within any given context we find differences related to social background characteristics of the respondents. Tests are used by society as a tool to differentiate among people in many ways that have very real consequences. Only to the degree that society is fair and just in making these discriminations will people agree that it is fair and just to use tests (p. 341).

Ebel (1972) specifies four solutions for counselors in avoiding harmful consequences of test usage. One

solution is to emphasize the use of tests to improve the status of students, and then to deemphasize test use to determine a child's status. A test should never be used as the sole criteria for determining levels of achievement. A second solution is to broaden the base of achievements tested in order to recognize and develop the wide variety of talents needed in our society. For example, achievement tests often lack testing in such areas as the fine arts, industrial arts, business arts, and vocational agriculture. These areas are important in the educational development of certain students. A third solution is to share openly with students the information that a test reveals about their abilities and prospects. Finally, there must be a decrease in the use of tests to impose decisions on others, and an increase in their use as a basis for better decision making. The emphasis must be on making educational gains for more than one student.

#### Recommendations for Test Use

The following is a list of suggestions found to be important in using standardized achievement tests.

1. Counselors must be educated in principles of educational measurement. Programs must be developed by educational institutions and state and local agencies in

helping counselors interpret and communicate test results. A measurement specialist must be considered as a viable member of the school's professional staff.

2. Graduate and undergraduate institutions that offer educational programs must emphasize the notions of measurement error, test validation, and score interpretation, as well as how to use these notions in decision making.

3. As counselors, we must bear some of the responsibilities for past misuse and must continue to look for improvement. A counselor should never ignore past abuses, but try and correct them.

4. Scoring reports of standardized achievement tests must be made clear to students and parents.

5. Test presentations and conferences must become geared to the interests and expectations of those concerned. New concepts and ideas must be examined in concluding whether they are useful in achievement testing.

6. Test results should be presented in an effective manner. The decision maker must look at all test results in making educational decisions and not just a sampling of the results.

7. There must be continued research on accountability procedures, competency testing, and the use and misuse of standardized achievement testing.

8. Counselors must contribute substantially to the total mission of the school if they are to maintain their hard-earned and necessary reputation as an essential component of the educational process. This definitely includes the counselor's work with standardized achievement tests.

### Conclusion

In summary, it seems obvious that standardized achievement testing is going to continue. Goslin (1963) states that at least 75 percent of the public school systems in the United States, as well as a large proportion of independent schools, have regular testing programs and that all school systems use achievement tests to some extent. No longer can achievement tests be used without carefully considering their effects on the individuals and institutions involved. Counselors should frequently review their goals and priorities when considering the possible effects of achievement tests on their students.

There is concern over the social implications of standardized achievement testing. These concerns should instigate further study on the use and misuse of standardized tests and their limitations. The studies should be objective and made by people who are concerned about testing and aware of the social context in which tests

are used. Only then will standardized achievement testing be a vital component in the education of students.

## BIBLIOGRAPHY

## BIBLIOGRAPHY

### Reference Notes

1. Beck, M. D. Uses and Misuses of Standardized Test Scores on a Local Level-A Test Developers Perspective. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, April 1981.
2. Travers, E. J. Some Uses of Standardized Tests-A Director's Handbook on Topics in Measurement and Guidance. Trenton, New Jersey: New Jersey Department of Education, 1979.
3. Feldt, L. S. and Forsyth, R. A. On the Validity of the Iowa Tests of Educational Development as an Aid in Program Evaluation. Iowa Testing Programs Papers, No. 3, June 1973.
4. Haag, C. H. The Uses and Abuses of Standardized Testing. Paper presented at the summer convention of the American Association of School Administrators, Minneapolis, Minnesota, July 1978.
5. Frechtling, J. A. Use and Misuse of Norm Referenced Test Scores for Decision Making at the Local School Level-A School Districts Perspective.

Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, April 1981.

6. Echternacht, G. Uses and Misuses of Test Scores: Technical Assistance Perspective. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, April 1981.



- Anderson, S. B. Parents and standard tests. Peabody Journal of Education, January 1981, 58, 96-107.
- Badal, A. W., and Larson, E. P. On reporting test results to community groups. Measurement in Education, May 1970, 1, 1-4.
- Brim, O. G. American attitudes towards achievement tests. American Psychologist, 1965, 20, 125-130.
- Cronbach, L. J. Essentials of Psychological Testing (3rd ed.). New York: Harper and Row, 1970.
- Cronbach, L. J. Five decades of public controversy over mental testing. American Psychologist, 1975, 30, 1-14.
- Daley, W. C. Test scores: Fragment of a picture. Test Service Notebook No. 24. New York: Harcourt, Brace & World, 1959.
- Darehshori, C. Standardized Testing Issues-A Teacher's Perspective. Washington, D. C.: National Education Association, 1977.
- Davis, F. B. Reporting test data in the media: Two case studies. Reading Teacher, December 1972, 26, 305-310.
- Dewey, J. Democracy and Education. New York: Macmillan, 1916.
- Ebel, R. L. Perspectives in Educational and Psychological Measurement, Englewood Cliffs, New Jersey: Prentice Hall, 1972.

- Glasser, W. Schools Without Failure. New York: Harper & Row, 1969.
- Goslin, D. A. The Search For Ability. New York: Russell Sage Foundation, 1963.
- Hawes, G. R. Releasing test scores: Urgent or unthinkable. Nations Schools, April 1973, 89, 41-55.
- Hill, G. E. Standards for test users. Measurement and Evaluation in Guidance, Fall 1969, 2, 141.
- Howe, H. Tests and schooling. The College Board Review. Fall 1980, 117, A2-A7.
- Kirkland, M. C. The effects of tests on students and schools. Review of Educational Research, 1971, 41, 303-350.
- Large Scale Programs of Testing for Guidance. Princeton, New Jersey: Educational Testing Service, 1958.
- Mitchell, B. C. A glossary of measurement terms. Test Service Notebook No. 13. New York: Harcourt, Brace & Jovanovich, 1981.
- Neulinger, J. Attitudes of American school students toward the use of intelligence tests. Personnel and Guidance Journal, 1966, 45, 337-341.
- Postal, R. Student progress judged closely by test scores. Waterloo (Ia.) Courier, January 23, 1983, p. E8.

- Responsibilities of users of standardized tests: An American personnel and guidance policy statement. American Personnel and Guidance Association Guidepost, October 5, 1978, pp. 5-9.
- Shertzer, B., and Linden, J. D. Fundamentals of Individual Appraisal. Boston: Houghton Mifflin, 1979.
- Thorndike, R. L., and Hagen, E. Measurement and Evaluation in Psychology and Education (3rd ed.). New York: John Wiley & Sons, 1969.
- Turnbull, W. W. Testing scores in perspective. Educational Record, 1978, 59, 291-296.
- Wingo, G. M. Philosophies of Education: An Introduction. Lexington, Massachusetts: D. C. Heath & Company, 1974.