DOCUMENT RESUME

ED 110 497 TM 004 770

AUTHOR Ferguson, Richard L.
TITLE The Use of Assessment in the Selection of College Students in the United States.
PUB DATE [Apr 75]
NOTE 15p.; Paper presented at the Annual Meeting of the American Educational Research Association (Washington, D.C., March 30-April 3, 1975)

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS Academic Aptitude; *Achievement Tests; *Aptitude Tests; *College Admission; Competitive Selection; Higher Education; High School Students; History; Predictive Ability (Testing); Test Construction; *Testing Programs; Test Reliability; Test Validity
IDENTIFIERS ACT; American College Test; SAT; Scholastic Aptitude Test

ABSTRACT

Assessment programs like the American College Test (ACT) and the Scholastic Aptitude Test (SAT) continue to play an important role in the admission decisions of many colleges. In many cases, they also serve a meaningful counseling and guidance function for high schools and colleges. The data reported in this paper suggest that the two programs can be used effectively for both of these purposes. Moreover, the procedures used to develop and administer the programs are highly refined and generally contribute to the overall quality of the data which result from use of the programs. (Author/DEP)

THE USE OF ASSESSMENT IN THE SELECTION OF COLLEGE STUDENTS IN THE UNITED STATES

Richard L. Ferguson
American College Testing Program

Introduction

According to the most recent data available from the U.S. Bureau of the Census (1974), approximately 2.3 million individuals 14 to 34 years old were enrolled in October 1973 as first-year students in public and private colleges in the United States. Of that number, an estimated 1.75 million were full-time students.

When selecting a college, most of these first-year students had a wide range of options available to them including junior colleges, technical schools, and institutions offering programs leading to undergraduate and graduate degrees. About 57 percent of the students enrolled in 4-year colleges. The remaining 43 percent enrolled in 2-year institutions.

Large numbers of these first-year students, though certainly not all, were required to complete either the ACT Assessment Program (ACT) or the Scholastic Aptitude Test (SAT) as a prerequisite to college admission or as one of the requirements of all enrolled first-year students at individual colleges. These two assessment programs are administered by The American College Testing Program and the College Entrance Examination Board, respectively. The two programs and some of the uses which colleges make of the data resulting from them are the focus of this paper.

Historical Background on the SAT and the ACT

Tests have long been used by colleges in the United States as one means for obtaining information useful for informed decision making about

---

Paper presented at the Annual Meeting of the American Educational Research Association, Washington, D.C., 1975.

student admissions. The first testing program to be administered on a large scale for that purpose was the College Board Scholastic Aptitude Test (Angoff, 1971). The SAT was initially administered in 1926 as a battery of nine subtests. Subsequent years brought changes in the SAT including a division of the SAT into two sections, one measuring verbal aptitude and the other measuring mathematical aptitude. The verbal section includes items from several domains including the social, political, scientific, artistic, philosophic, and literary. The mathematics section measures skills drawn from arithmetic, algebra and geometry (Boros, 1972). Recently, a section measuring standard written English has been added.

In 1959, some 33 years after the SAT was first administered, the American College Testing Program was founded and its admission program, the ACT Assessment, became the second of two national testing programs available for use by postsecondary institutions. Initially, the ACT Program included four tests of students' general educational development, an English Usage Test, a Mathematics Usage Test, a Social Studies Reading Test, and a Natural Sciences Reading Test. In 1965, a Student Profile Section (SPS) was added to the ACT Program. The SPS was included to provide a broad information base for student and institutional decision making and for counseling and guidance purposes. More recently, an interest inventory has been added to enhance the effectiveness of the ACT as a counseling tool (ACT, 1973).

Together, the ACT and the SAT are completed by over two million examinees each year. Because many students take the SAT both as juniors and seniors (Angoff, 1971) and because not all students participating in the two programs enroll in a college, it is inaccurate to conclude that nearly all of the approximately 2.3 million first-year students enrolled in 1973 completed either the ACT or the SAT. However, it is clear that a very substantial proportion of those students did participate in one of the two programs.

Of the first-year college students who were not required to take either the ACT or the SAT as a prerequisite to college admission, many were required to complete the assessment programs once enrolled in college. This procedure frequently occurs when the colleges are conducting institutional research and need complete data on all enrolled students. Still other students enrolled in colleges which had no testing requirement at all or in state-related institutions in states that have their own testing programs. Examples of the latter institutions include state affiliated colleges and universities in Florida, Minnesota, New York, and Washington. State-related institutions in New York, for example, use the results of the high school Regents examinations as partial input into their admission decisions. Similar types of senior-level testing are also used in the other three states.

Having established the fact that large numbers of prospective college students participate in the ACT or the SAT programs either before or shortly after they are admitted to college, the next two sections of this paper consider the procedures used to develop the assessment instruments which comprise the two programs and the uses of the data which result from those programs.

Development and Administration of the Admission Assessment Programs

Both the ACT and the SAT are objective-type instruments which are administered in hundreds of test centers throughout the United States several times each year. The ACT is administered five times a year and the SAT six times. Each is administered on Saturdays, thus enabling the use of school facilities (high school and college) as test centers, and each is administered by qualified school personnel who are provided specific test administration instructions to insure relatively standardized testing conditions. Non-Saturday testing is provided by both programs for students whose religious faith prohibits taking the examination on a Saturday.

Because of the uses which colleges make of the two programs and because they are administered so frequently, considerable effort is expended to insure that the tests and test results are not compromised. One step taken to assure the integrity of the programs involves the introduction of several new test forms each year. Since it would be redundant to discuss in detail the developmental procedures used with both the ACT and the SAT, only the procedures used with the ACT Assessment Program are discussed in this paper. A detailed description of the procedures used with the SAT are printed in a book authored by Angoff (1971).

Development of the ACT Assessment

Three new forms of the ACT tests are begun each year. The development of those forms involves a complex sequence of activities which takes approximately two and one-half years to complete. During the first stage of this developmental process, the specifications for the tests are designed (or modified). At this stage, input reflecting the concerns and advice of content experts in each of the curriculum areas included in the ACT Program and input from ACT's own staff of experts are incorporated into the specifications. Concurrently, item writers from high schools and colleges all over the country are identified to write test items and stimulus materials. Once written, those materials proceed through extensive editorial cycles during which they are reviewed for technical quality, appropriateness of content, and clear and effective expression.

To provide the data needed for technical evaluation of the test materials, the materials are field tested with samples of students who have characteristics similar to the population of students who will eventually take the ACT tests. In the past, ACT has field tested its items in high schools in the state of Iowa

in conjunction with the State's fall testing program. More recently, however, "new" items are field tested by including them as extra items in the tests administered on the national test dates. These items neither contribute to nor detract from students scores on the tests. The latter approach is descriptive of the procedures employed for the SAT as well.

Once the new test materials have been field tested and item analysis data obtained, the tests are assembled according to precise statistical and content specifications. Then, the tests are reviewed by members of minority groups and by measurement experts external to ACT, as well as by the ACT testing staff. Once the latter input has been evaluated and, where appropriate, incorporated in the tests, the tests are printed and readied for distribution.

A final step in the test development process occurs when the new forms of the tests are equated to forms administered on earlier test dates. In the case of the ACT, equating takes place in conjunction with the October national test date each year. At that time, three new forms of the tests are equated to old tests by administering new and old forms of tests to parallel samples of students and then performing an equipercentile equating.

New forms of the SAT are equated to two previous forms and the results averaged. This is accomplished by including items on the new forms which are common to items on previously used forms. The two old forms to which each new form is equated are determined in a systematic braiding fashion which aims to shorten the equating distance between each new form and all other forms.

College Uses of Admissions Assessment Data

Student Selection Procedures

It is no small task to attempt to describe the various student selection procedures used by colleges and the roles that testing plays in those procedures.

Although a survey of those procedures is presently underway at ACT, the most recent survey completed by ACT on that subject was conducted in 1969 (Crockett, 1969). That survey involved 1,091 colleges and universities that participated in ACT programs and services. Of these colleges, 855 or 78 percent, responded to a questionnaire addressed to, among other things, testing requirements and admission practices.

Of the institutions responding, 55 percent indicated that test data from the ACT Assessment were used for deciding student admission to certain curricula, 60 percent used the data for probationary or conditional admissions, and 52 percent indicated they used the data as a basis for rejecting an applicant. Only 17 percent of these institutions reported the use of non-test data (data from the Student Profile Section) in making admission decisions. Collectively, these data indicated that most admission decisions involving ACT data (approximately 50 percent) were based on test scores and high school grades.

A more recent survey (Sedlacek et. al., 1973) was completed involving a national sample of 110 large, primarily white, public and private schools throughout the United States. That survey indicated that high school grades, high school rank, and tests like the ACT and SAT remain the most commonly employed admissions criteria used by schools. Specifically, 84 percent of the schools sampled reported using as admissions criteria either an average of self-reported grades in 4 subject areas combined with ACT scores or high school rank combined with SAT scores.

The data reported to colleges by ACT also include percentile ranks of predicted grade point averages of all students taking the ACT and applying to the college. The student who ranks at the 62nd percentile in the all-freshman group, for example, could be expected to do better than 62 percent of the students on which the data are based. Those percentile ranks are based

on an analysis of how students in the past with similar characteristics have achieved at the college. ACT also provides the probability of a student obtaining a "C" average of higher, another way of interpreting the same data. All of these types of predictive data can also be provided for specific subgroups of students. For example, some colleges opt to obtain predictions on subgroups of their freshman class such as students enrolling in engineering, pre-medicine, liberal arts, etc., or for other groups which can be identified by the colleges.

It is important to note that both ACT and CEEB advise caution in the use of test results from their respective programs. For example, ACT has long taken the responsibility for providing a broad spectrum of student data to colleges, data which reflect student attributes other than academic ability. The SPS section discussed earlier reflects a serious commitment to that responsibility. It includes information on educational-vocational plans, out-of-class high school accomplishments, educational needs, extracurricular plans, factors influencing college choice, and student biographical data. The interest inventory recently added to the test also contributes to this objective. These diverse data are collected and are provided to colleges because ACT believes that the types of information available affect the nature and the quality of the decisions which colleges and students make.

Other Uses

The ACT test scores are used by many colleges for purposes other than the selection of students. Some of those uses include sectioning of classes, placement, and, in some instances, the award of credit. The ACT survey cited earlier in the paper revealed that of the 855 ACT participating colleges responding, 61 percent indicated use of ACT test scores for freshman course sectioning. Approximately one-third of the colleges reported the use

of ACT data both for curriculum placement and as a criterion for limiting the number of hours a student may carry. Fifteen percent of the colleges used the test scores for granting advanced placement.

Many of these same functions are satisfied by the College Board Achievement Tests which are distinct from the SAT. Those tests, are administered several times each year in conjunction with the SAT and include tests in English Composition, Literature, American History and Social Studies, European History and World Cultures, Chemistry, and several languages, to name a few. The CEEB College Level Examination Program (CLEP), the CEEB Advanced Placement Program, and the CEEB Comparative Guidance Program have all been used by colleges for similar purposes.

Research on the Effectiveness of the Uses of Admission Assessment Data

In this section, important statistical characteristics of the two programs are discussed. First, reliability studies on the ACT and the SAT are summarized. Then, the validity of the programs for predicting academic success in college is considered.

Test Reliability

Both the ACT and the SAT are evaluated annually to ascertain their reliability. Thus, a wealth of reliability data for each program are available, some of which are reported next.

ACT Assessment Reliability

Median odd-even reliabilities for the four ACT tests range from .85 for the Natural Sciences Reading Test (52 items) to .90 for the longer English Usage Test (75 items). For the ACT composite, the median odd-even reliability is

.96 with a standard error of measurement of 1.04. The standard error of measurement is based on scaled test scores which range from 1 to 36 with a national mean of approximately 20. Although the ACT tests are designed to be power tests, the time limits do impose a moderate degree of speededness. This might lead to somewhat inflated odd-even estimates or reliability.

KR-20 reliability estimates for the four tests are .89, .89, .85 and .84. For the composite, the KR-20 estimate is .91 with standard error of measurement of 1.44. Again, these estimates may be somewhat inflated due to the test speededness.

As for the stability of the ACT Assessment measurements over short periods of time, several studies designed to produce parallel form reliabilities for the four test resulted in estimates in the range from .73 to .85, with correlations generally in the high 70s and low 80s. For the ACT composite, the correlation was approximately .90. These data were based on parallel forms of the tests administered from 2 to 6 weeks apart.

Longer term estimates of stability have also been computed. Parallel form reliability of the ACT composite over a 3-month interval, a 5-month interval and 2-year interval are reported at .90, .88, and .88 respectively.

SAT Reliability

Internal reliability estimates for the SAT over the three year period from 1966 to 1969 based on Dressel's adaptation of KR-20 are about .90 to .91 for the Verbal Section and .89 to .91 for the Mathematical Section. The standard errors of measurement for the two sections range from 30.9 to 33.3 for the Verbal and from 32.4 to 36.3 for the Mathematical.

Estimates of parallel form reliability based on students who took the SAT on two occasions somewhat less than a year apart are about .89 for the

Verbal Section and .88 for the Mathematical Section (Angoff, 1971).

Test Validity

Over the years, literally hundreds of validity studies have been conducted on the ACT and the SAT. Many of these studies have been conducted by staff at the colleges which use the two programs and involve specific uses of programs at those colleges. Many of these studies are documented in the vast literature on both the ACT and the SAT.

In addition, both ACT and CEEB have collected data on the validity of their respective assessment programs for a variety of uses. To facilitate discussion, but at the risk of over simplifying the uses made of admission assessment programs, the uses of those programs are divided into two categories in this paper, one emphasizing the selection of students, the other emphasizing decisions about students once they are admitted to college.

Validity of the Programs for Selecting Students

Many institutions use predictive data like those discussed earlier as indices in the student selection process. Consequently, ACT and CEEB routinely examine the predictive validity of their instruments for this purpose, specifically for predicting first semester and/or first-year college grade point average. Data available on predicting college grades using the ACT and the SAT indicate median multiple correlations in the range of .55 to .60 for both programs. In both cases, an index of high school academic accomplishment is used along with test scores in the prediction system.

The multiple R's for predicting college success have remained relatively stable over the past few years for both the ACT and the SAT. This stability has been maintained at the same time significant changes appear to have been occurring in the makeup of the college student population. For example, both ACT and CEEB report a trend toward decreases in the range

of test scores of students taking the ACT and the SAT. ACT composite scores have declined by approximately one standard score, or about 1/5 of a standard deviation over the past decade. CEEB has reported a decline of similar magnitude.

An interesting phenomenon, however, is that over the same period of time that test scores were on the decline, high school and college grades were on the increase. ACT reports that the high school average in four subjects estimated by students each year when they take the ACT Assessment rose .06 from 1966-67 to 1969-70 and another .14 from 1969-70 to 1972-73. Those figures were based on enrolled students at over 400 colleges. Similar trends were observed for first term or first year college grades. These trends were all accompanied by larger standard deviations in test scores and high school and college grades, thus suggesting a trend toward greater heterogeneity in the enrolling classes at U.S. colleges.

In spite of this apparent change in the abilities of students applying to and enrolling in college, the accuracy of predictions of students' college grade point average based on the ACT tests and self-reported grades has remained rather stable. Ferguson and Maxey (1975) determined the multiple correlations for a representative sample of colleges in which such changes have taken place over the last six years. They found slight decreases in a few colleges and the same or higher correlations in most of the colleges. National norms for several hundred colleges over the same six year period result in similar findings.

Validity of the Programs for Decisions about Enrolled Students

Many institutions using admission tests are less concerned with being selective and instead are more concerned with identifying students who have

adequate academic preparation for the courses and programs they offer This is one of the most typical uses of ACT test scores. In such cases, ACT recommends that colleges look closely at the chances of "C or better" in specific courses. For example, a particular college might opt to have ACT indicate the chances for students to obtain a "C or better" in a specific course in freshman English, or an entering level mathematics course, etc. Colleges may request ACT to provide such estimates for up to different courses. Cross validation studies indicate that the ACT prediction equations based on the four ACT tests and constructed on first year's entering class are generally as predictive of first semester GPA for the subsequent year's entering class as of the original class. As noted earlier, these probabilities are provided on the college's copy of students' score reports for those colleges participating in ACT research services.

Validity of the Programs for Predicting Career Accomplishment

Although a wealth of data exist which establish a strong relationship between test scores from the two programs and first-term or first-year college grade point average, there is less evidence available on the relationship of ACT and SAT test scores to accomplishment or job success after college.

Munday and Davis (1974) have investigated the extent to which different types of accomplishment after college are related to academic accomplishment before or in college. Their research involved young adults who had taken the ACT Assessment as high school seniors 6 years earlier and who were 2 years out of undergraduate school. These young people were asked to respond to an Alumni Survey which sought information on adult accomplishments as defined by six scales, one each on Leadership, Music, Literary, Art, Drama, and Science. Scores on these scales were correlated with

college admission data. The adult accomplishments were found to be uncorrelated with academic talent, including test scores, high school grades, and college grades. Adult accomplishments were, however, related to comparable high school non-academic accomplishments.

findings should not, however, obscure the fact that a college degree is a prerequisite to entering many professions. And when test scores from the ACT and the SAT are combined with some index of high school schievement (e.g., class rank or grade point average) they are highly predictive of student success in college.

Summary

Assessment programs like the ACT and the SAT continue to play an important role in the admission decisions of many colleges. In many cases, they also serve a meaningful counseling and guidance function for high schools and colleges. The data reported in this paper suggest that the two programs can be used effectively for both of these purposes. Moreover, the procedures used to develop and administer the programs are highly refined and generally contribute to the overall quality of the data which result from use of the programs.

References

Angoff, W. The College Board Admissions Testing Program: A technical report on research and development activities relating to the scholastic aptitude test and achievement tests. New York: College Entrance Examination Board, 1971.

Boros, O. Editor, The Seventh Mental Measurement Yearbook. Vol. I, 1972.

Crockett, D. The 1969 survey of ACT participating colleges and universities. Report No. 2, Iowa City, Iowa: The American College Testing Program, 1969.

Editors. Technical report for the ACT Assessment Program: Assessing students on the way to college. Iowa City, Iowa: The American College Testing Program, 1973.

Ferguson, R. and Maxey, J. Trends in the academic performance of high school and college students. Iowa City, Iowa: The American College Testing Program, 1975.

Munday, L. and Davis, J. Varieties of accomplishment after college: Perspectives on the meaning of academic talent. ACT Research Report No. 62, Iowa City, Iowa: The American College Testing Program, 1974.

U.S. Bureau of the Census, Current Population Reports, Series P-20, No. 272, "Social and Economic Characteristics of Students; October 1973." U.S. Government Printing Office, Washington, D.C., 1974.