DOCUMENT RESUME

ED 300 461 TM 012 514

AUTHOR Sawyer, Richard; Brounstein, Paul
TITLE The Relationship between ACT and SAT Scores among Academically Talented Seventh Grade Students.
PUB DATE Apr 88
NOTE 25p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (New Orleans, LA, April 6-8, 1988).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academically Gifted; College Bound Students; *College Entrance Examinations; Comparative Analysis; English; *Grade 7; Junior High Schools; *Junior High School Students; Language Tests; Mathematics Tests; *Scores; Talent; Verbal Tests
IDENTIFIERS *ACT Assessment; Concordance (Data); *Scholastic Aptitude Test; Talent Identification Program NC

ABSTRACT

This paper describes the efforts of the Talent Identification Program (TIP) of North Carolina to relate American College Testing (ACT) Assessment and Scholastic Aptitude Test (SAT) scores of academically talented seventh graders. The study was also designed to interpret the results in the context of "concordance" studies that have been done with other groups of students. The target population for the central study consisted of students who applied to the TIP during the 1986-87 talent search by taking the SAT in December 1986 or January 1987. A stratified random sample of 2,042 applicants was invited to take the ACT Assessment on the April 1987 test date. Results from the 611 students who responded to the invitation indicate that the ACT Composite and SAT Total scores are nearly as strongly related among academically talented seventh graders as they are among college-bound high school students. However, the English Usage/Verbal and Mathematics Usage/Mathematics components of the two test batteries measure different characteristics of academically talented young students, and these characteristics are developed to relatively different extents than they are among high school students. The results illustrate the tendency of conversion tables for non-parallel tests to differ among various subgroups. (TJH)

The Relationship Between ACT and SAT Scores Among Academically Talented Seventh Grade Students

Richard Sawyer

American College Testing Program

Paul Brounstein

Talent Identification Program

A paper presented at the annual meeting of the

National Council on Measurement in Education

April, 1988

THE RELATIONSHIP BETWEEN ACT AND SAT SCORES AMONG ACADEMICALLY TALENTED SEVENTH GRADE STUDENTS

The Talent Identification Program (TIP), one of four university-based talent searches in the United States, has been administered by Duke University since 1981. The major objectives of TIP are to identify precocious students of junior high school age and to assist them in developing their abilities. About 46,000 seventh-grade students apply for admission to the program each year.

Since it was established, TIP has used the Scholastic Aptitude Test (SAT) to determine the eligibility of academically talented youngsters for advanced courses and other special educational opportunities available through TIP (Duke University, 1987). In 1987, TIP and the American College Testing Program (ACT) undertook research to determine whether the ACT Assessment could also be used for this purpose. The educational considerations that led to this research are described in the paper by Stanley and York (1988) presented at this symposium. An essential statistical task in using alternative tests for selection is to determine the comparable scores at which selection will be made. This paper describes our efforts to relate the ACT and SAT scores of academically talented seventh grade age students, and to interpret the results in the context of "concordance" studies that have been done on other groups of students. Given that the two tests have different content, it is possible that some students will tend to have higher standing on one test than on the other; a further goal of this paper was, therefore, to compare the demographic and background characteristics of such students.

Methodological Considerations

E. F. Lindquist (1964) described the limitations of conversion tables for scores on non-parallel tests such as the ACT and the SAT. I will briefly review the methodological problems Lindquist identified and discuss their implications for this particular application.

The fundamental limitation in using conversion tables for non-parallel tests is a direct consequence of their non-parallelness: Depending on their abilities, some students will score systematically higher on one test than the other. For this reason, scores on the two tests can never be used interchangeably without restriction.

In the situation being considered here, both the ACT Assessment and the SAT are designed to measure abilities considered to be strongly related to success in college. Beyond this superficial global similarity, however, the two tests are in many respects quite different. A primary difference is that the content of the ACT Assessment is achievement-oriented; the items in the ACT Assessment tests are representative of knowledge and skills taught in the typical high school curriculum. The SAT, on the other hand, is designed to measure more abstractly defined educational aptitudes. For example, the ACT English Usage test primarily measures students' knowledge and skills in reading and writing standard English, while the SAT Verbal test includes some items (such as verbal analogies) that measure more abstractly defined verbal reasoning abilities. The ACT Mathematics Usage test measures students' skills in solving mathematical problems typically encountered in mathematics courses through Grade 12, while the SAT Mathematics test includes some items that can be stated in very basic terms, but that require considerable intellectual agility to solve. These comparisons are not intended to be complete or precise descriptions of either test battery; for this, the documentation

prepared by the two test publishers should be consulted (American College Testing Program, 1987; College Entrance Examination Board, 1987). These comparisons do, however, give an indication of how the general characteristics of the ACT Assessment differ from those of the SAT.

One must expect, given the differences in the content of the two test batteries, that for any given ACT Assessment score, there will be a range of comparable SAT scores, and that these SAT scores will vary systematically with the particular educational abilities of students taking the tests. This practically guarantees that the relationship between ACT and SAT scores will vary among student subgroups defined by demographic or other background variables. This also affects the use of the tests for selection: To the extent that a student possesses abilities measured by one test and not by the other, that student can improve his or her chances of being selected merely by taking the former test. Of course, the greater the extent that the distinct abilities measured by the two batteries are statistically related among students, the smaller the range of comparable SAT scores will be, and the less systematic variation in the ACT/SAT relationship there will be among student subgroups.

In identifying students for TIP, the concern is not so much to select students with the particular academic abilities measured by either the ACT or the SAT, as it is to identify those students who will have a high probability of successfully completing advanced coursework. This consideration would suggest that comparable selection scores be statistically associated with a comparable predicted level of performance in the program. Setting the selection scores this way would not only allow students more flexibility in presenting their strengths, but would also be directly related to the intended goals of the program. Unfortunately, the test score and course grade data

needed to make this comparison were not available, given practical constraints on this study. During the time available for doing the study (January, 1987 to June, 1987), it was feasible only to collect and study the ACT Assessment scores for a sample of TIP applicants who had previously taken the SAT. Some other method of establishing comparable scores was therefore needed.

(In July, 1987, TIP administered the ACT Assessment to a random sample of students enrolled in the Duke summer program, some of whom had been admitted on the basis of their ACT test scores. We matched the summer ACT test scores of the enrolled students with their summer course grades and with their original SAT scores. From the matched data set, Maxey and Dreyden (1988) studied the criterion-related validities of the ACT Assessment and the SAT, but not enough data were available to study the association between ACT and SAT implied by common predicted grades.)

One conceivable method of relating SAT scores to ACT scores is to use the regression of one on the other. I believe, though, that this method is always inappropriate when the tests are used for selection, as they are in this application. To see why this is so, consider the following example: Suppose that the ACT Composite score can be predicted from the SAT Total score by some regression function:

$$\text{predicted SAT Total} = f(\text{ACT Composite})$$

The function f can be any monotone increasing function (typically, it is linear); moreover, the predictor variables could be the four ACT subtest scores instead of the ACT Composite. In a situation like this, there will always be a "regression effect": For an ACT Composite score above the mean, the predicted SAT Total score will be at a lower percentile rank than the ACT Composite score, and for an ACT Composite score below the mean, the predicted SAT Total score will be at a higher percentile rank than the ACT Composite

score. This means that students who took the ACT Assessment and who had above average scores would be penalized if their predicted SAT Total scores were judged against the SAT Total scores of students who actually took the SAT; and the opposite would be true of students with below average ACT scores. The reason this occurs is that a student with an above average ACT Composite score has that score partly because of true abilities and partly because of positive random measurement errors. In a repeated measurement, either on the ACT Assessment or on a statistically related test like the SAT, these random errors would on average not be positive. A consequence of the regression effect is that the regression of SAT scores on ACT scores yields different comparable scores than the regression of ACT scores on SAT scores. The inappropriateness of using regression-based conversion tables for selection can be illustrated by the hypothetical situation where the two tests are strictly parallel: In this situation we would be selecting some students on the basis of their observed scores and other students on the basis of regression estimates of their true scores! Thus, if predicted SAT scores were to be used for selection, then they should not be used with observed SAT scores (though perhaps they could be used with regression estimates of SAT true scores).

In this study we used the common "equipercentile" procedure to establish comparable scores. In the equipercentile procedure, scores with equal percentiles on the two tests are designated as comparable. This method is inefficient in that it does not utilize the bivariate structure of the data; on the other hand, it (by definition) preserves percentile ranks.

Converting scores using the equipercentile method (or other method) is also subject to two other limitations described by Lindquist:

1. Students who choose to take both tests may not be representative of the population for which the tables will be used. By the previous remarks, we should not be surprised if these students were also non-representative with respect to the relationship between the two tests.
2. Typically, there is no control either on the order in which students take the two tests or on the time that elapses between the two administrations. In an ideal design, students would be administered the two tests in random order within a short time of each other. Departures from this ideal raise the possibility that observed relationships have been contaminated by practice and growth effects.

In the present study, both limitations apply, but the second is the more serious, as all students took the SAT in either December, 1986 or January, 1987, and took the ACT Assessment in April, 1987. It is therefore likely that the comparable ACT score for a given SAT score is higher than it would have been had we been able to administer the tests in random order at nearly the same time. Mindful of these limitations, we decided to develop simple equipercentile tables that could be used to establish interim selection scores for the ACT Assessment until data became available that would permit making more refined estimates.

METHOD

In fall, 1987, TIP encouraged academically able students to apply for admission to the 1986-87 talent search. The applicants sought were students of seventh grade age who attended school in one of the 16 states served by TIP and whose scores on any of several nationally standardized academic ability tests were in the upper 3 percent. The target population for this study

consisted of students who applied to TIP during the 1986-87 talent search by taking the SAT in December, 1986 or January, 1987, and by sending the scores to TIP. There were 32,901 students in this applicant population.

We selected a stratified random sample of 2,042 applicants and invited them to take the ACT Assessment on the April, 1987 test date. The stratification variables were: racial/ethnic background (black, white, Hispanic, native American, oriental, other); school location (large city, medium city, small city, suburb, small town, rural); and sex. Sample sizes within racial/ethnic strata were made as nearly equal as possible; this was done to allow the possibility of studying the ACT scores of the separate racial/ethnic groups.

Two incentives were offered students to take the ACT Assessment: The test fee of $10.50 was waived, and students were informed that they could qualify for admission to Duke's summer program on the basis of either their ACT or their SAT scores.

Of the 2,042 students invited to take the ACT Assessment, 611 (or 30%) did so. We were able to match the ACT and SAT records of 558 of the 611 ACT-tested students either by telephone number, SSN, or name. One of the matched ACT/SAT records contained incomplete data, and was dropped. The analysis was therefore based on the matched ACT/SAT records of 557 students.

Weighted frequency distributions and means were computed for the ACT English Usage, ACT Mathematics Usage, ACT Composite, SAT Verbal, SAT Mathematics, and SAT Total test scores. Weighted correlation coefficients were also computed between the ACT and SAT test score pairs to provide a conventional index of the strength of their relationships. The weight for each case was determined by the value x of the racial/ethnic variable RACE:

$$\text{WGT} = \frac{\text{Number of records in applicant file for which RACE} = x}{\text{Number of matched ACT/SAT records for which RACE} = x}$$

The weights, therefore, inflated the matched sample to the SAT-tested TIP applicant population on the basis of the stratification variable RACE. Because school location and sex categories were proportionately represented in the sample, the weighting was not based on these variables.

The weighted frequency distributions were used to create equipercentile ACT/SAT conversion tables. Raw comparable scores were developed from the weighted frequencies; the raw comparable scores were then smoothed with cubic splines (Kolen, 1984). The resulting smoothed ACT/SAT conversion tables were then compared with other ACT/SAT conversion tables developed at ACT or published elsewhere.

We next examined the characteristics of students who would be selected using the ACT Assessment, but not the SAT, or vice versa. The "selection scores" in this analysis were the median and third quartile of each test scores. (No analysis could be done with respect to the actual selection scores used by TIP for its summer program, as the sample was too small.) The student characteristics examined were sex, racial/ethnic background, and school location.

Results and Discussion

Table 1 summarizes the background characteristics and test scores of the TIP applicant population and the matched ACT/SAT sample. The two sets of proportions for the racial/ethnic categories are, of course, the same. The corresponding proportions for the other background characteristics are also very similar, and are well within sampling error of each other. The two SAT Total mean scores are also very close (725 in the population and 729 in the matched sample), given the SAT Total standard deviation of 126 and the sample size of 557. Some of the subgroups defined by race and school location (particularly blacks, American Indians, large city, and rural area) have

larger differences in their SAT Total score means. These differences cancel out, though, as far as the total group is concerned.

Results concerning the distributions of ACT and SAT test scores and their relationships with demographic characteristics are given in the paper by Dreyden and Stanley (1988) presented in this symposium.

Table 2 contains weighted correlations between selected pairs of ACT Assessment and SAT test scores. The ACT Composite/SAT Total correlation of .83 is only slightly smaller than the correlations of .85-.90 typically observed among 12th grade college-bound students. According to the ACT Assessment Technical Manual (ACT, 1987), the typical test-retest reliability of the ACT Composite score is also in this range.

Tables 3-5 contain the smoothed ACT/SAT conversions; the score ranges in these tables correspond to the ranges observed in the sample. The relationships in Tables 3-5 are depicted in Figures 1-3, along with relationships found from other ACT/SAT studies. The range in ACT scores of 20 to 25 in the graphs is typical of the scores of students selected for the summer program at Duke.

The graphs illustrate the point made earlier that the relationship between the ACT and SAT differs among different groups of students. For a particular ACT Composite score, for example, the comparable SAT Total score can differ by as much as 75 SAT score points, depending on the study.

The graphs also show that the ACT English Usage/SAT Verbal and ACT Mathematics Usage/SAT Mathematics relationships observed among TIP applicants are not typical of those observed among college-bound high school students. For a given SAT Mathematics score, the ACT Mathematics Usage scores of TIP applicants tend to be lower than the ACT Mathematics Usage scores of high school students; for a given SAT Verbal score, the opposite tends to be

true of ACT English Usage scores. Furthermore, these two opposing trends tend to cancel in the relationship between ACT Composite and SAT Total; there, TIP applicants appear to be typical of college-bound high school students.

The differences in the ACT/SAT relationships for English and mathematics are doubtless related to the differences in content between the ACT and SAT. Using the broad characterization of the ACT and SAT given earlier, one could hypothesize that TIP applicants with SAT Mathematics scores above 500 already have highly developed mathematical aptitudes, though they may not have been exposed to the contents of typical high school mathematics courses. These students' scores on the ACT Mathematics Usage test may be driven by their exceptional abilities to solve complex mathematical problems, and their capacity to learn to do so quickly, rather than by their prior exposure to the particular kinds of problems in the ACT Mathematics Usage test. Apparently, the opposite is true of seventh graders' achievements and aptitudes with respect to their ACT English Usage and SAT Verbal scores, though it is difficult to find a theoretical explanation for this.

Table 6 shows the demographic characteristics of students who would be "selected" at the median and third quartiles of the weighted test score distributions. The table contains the proportions of students in various "selected" groups who are female, minority, or who attended school in a small town or rural area. The sample sizes listed under the column headed by "N" vary because the quartiles were selected from weighted statistics and because the frequency distributions did not contain cumulative relative frequencies corresponding exactly to .50 and .75.

At the median, the ACT English Usage test tends to favor females, and the SAT Verbal test tends to favor males: about 56% of the students "selected" by the ACT English Usage test are female, as compared to 45% of the students "selected" by the SAT Verbal test. Moreover, 67% of students who would be "selected" by the ACT English Usage test, but not by the SAT Verbal test, are female.

A different result occurs at the third quartile. About 48% of the students "selected" by the ACT English Usage test are female, which corresponds very closely to the 47% proportion in the total group; on the other hand, about 54% of the students "selected" by the SAT Verbal test are female. Furthermore, about 94% of the students selected by the SAT Verbal test, but not by the ACT English Usage test, are female.

The differences between students "selected" at the median and third quartiles suggest that there is an interaction between skills in English usage, English verbal aptitude, and sex. Explaining the educational and social meaning of this phenomenon is beyond the scope of this paper. No differences of the magnitudes observed between females and males occurred in the proportion of students selected who were minority or who attended school in a small town or rural area.

Conclusions

The ACT Composite and SAT Total scores are nearly as strongly related among academically talented seventh grade age students as they are among college-bound high school students. The English Usage/Verbal and Mathematics Usage/Mathematics components of the two test batteries, however, measure different characteristics of academically talented young students, and these characteristics are developed to relatively different extents than they are

among high school students. These results illustrate the tendency of conversion tables for non-parallel tests to differ among various subgroups.

REFERENCES

American College Testing Program (1987). ACT Assessment Program Technical Manual. Iowa City, Iowa: Author.

American College Testing Program (1987). Content of the tests in the ACT Assessment. Iowa City, Iowa: Author.

Astin, A. W. (1971). Predicting academic performance in college: Selectivity data for 2,300 American colleges. New York: The Free Press.

College Entrance Examination Board (1987). Taking the SAT. Princeton, NJ: Author.

Dreyden, J. & Stanley, G. (1988, April). College entrance test score profiles and demographic information for talented seventh grade youth. Paper presented at the meeting of the National Council on Measurement in Education, New Orleans.

Duke University (1987). 1986-87 talent search results summary. Durham, North Carolina: Author.

Kolen, M. (1984). Effectiveness of analytic smoothing in equipercentile equating. Journal of Educational Statistics, 9, 25-44.

Langston, I. (1987). SAT-ACT Equivalents (Research Memorandum 87-4). Champaign, Illinois: University Office of School and College Relations, University of Illinois.

Lindquist, E. F. (1964). Equating scores on non-parallel tests. Paper presented at the meeting of the National Council on Measurement in Education, Chicago.

Maxey, J. & Dreyden, J. (1988, April). Measures of validity between the ACT Assessment and other achievement variables for precocious seventh grade youth. Paper presented at the meeting of the National Council on Measurement in Education, New Orleans.

Stanley, G. & York, A. (1988, April). The ACT Assessment as a measure for identifying talented seventh grade youth. Paper presented at the meeting of the National Council on Measurement in Education, New Orleans.

Table 1

Summary Statistics for TIP Applicant Population and Matched ACT/SAT Sample

Subgroup variable	Category	TIP applicant population[a]			Matched ACT/SAT sample[b]			
			SAT Total				SAT Total	ACT Composite
		Proportion	Mean	S.D.	N	Proportion	Mean	Mean
Race	Black	.04	665	110	105	.04	699	13.3
	White	.90	726	125	96	.90	728	14.8
	Hispanic	.02	677	120	106	.02	692	13.4
	American Indian	.00	681	104	24	.00	718	15.2
	Oriental	.03	804	152	152	.03	817	17.1
	Other	.01	754	133	74	.01	782	16.0
Sex	Male	.49	737	119	264	.47	741	15.4
	Female	.51	712	133	293	.53	716	14.0
School location	Large City	.12	753	130	90	.11	788	17.3
	Medium City	.10	738	127	64	.10	711	14.4
	Small city	.20	731	130	130	.19	743	14.6
	Suburb of city	.12	751	131	60	.12	727	14.7
	Small town	.34	709	120	174	.36	733	15.1
	Rural area	.10	689	116	39	.12	659	12.3
Total Group		1.00	725	126	557	1.00	729	14.8

Notes: a. N = 32,901
b. N = 557; weighted by Race

Table 2

Weighted Correlations Between ACT Assessment and SAT Test Scores*

ACT Assessment test score	SAT test score	Weighted correlation
ACT English Usage	SAT Verbal	.68
ACT English Usage	TSWE	.71
ACT Mathematics Usage	SAT Mathematics	.77
ACT Composite	SAT Total	.83

*N = 557

Table 3

Comparable ACT English Usage
and SAT Verbal Scores*

ACT English score	SAT Verbal score
10	240–250
11	260–270
12	280
13	290
14	300
15	310
16	320–330
17	340
18	350
19	360–370
20	380–400
21	410
22	420–430
23	440–450
24	460–490
25	500–510
26	520–540

*N = 557

Table 4

Comparable ACT Mathematics Usage and SAT Mathematics Scores*

ACT Mathematics score	SAT Mathematics score
6	290-300
7	310-320
8	330
9	340
10	350
11	360
12	370-380
13	390
14	400-410
15	420
16	430-450
17	460-470
18	480
19	490
20	500
21	510
22	520-530
23	540-550
24	560-570
25	580-590
26	600-610

*N = 557

Table 5

Comparable ACT Composite and SAT Total Scores*

ACT Composite score	SAT Total score
7	520
8	530-540
9	550-570
10	580-610
11	620-640
12	650-670
13	680-690
14	700-710
15	720-740
16	750-760
17	770-800
18	810-820
19	830-850
20	860-880
21	890-910
22	920-950
23	960-980
24	990-1000
25	1010

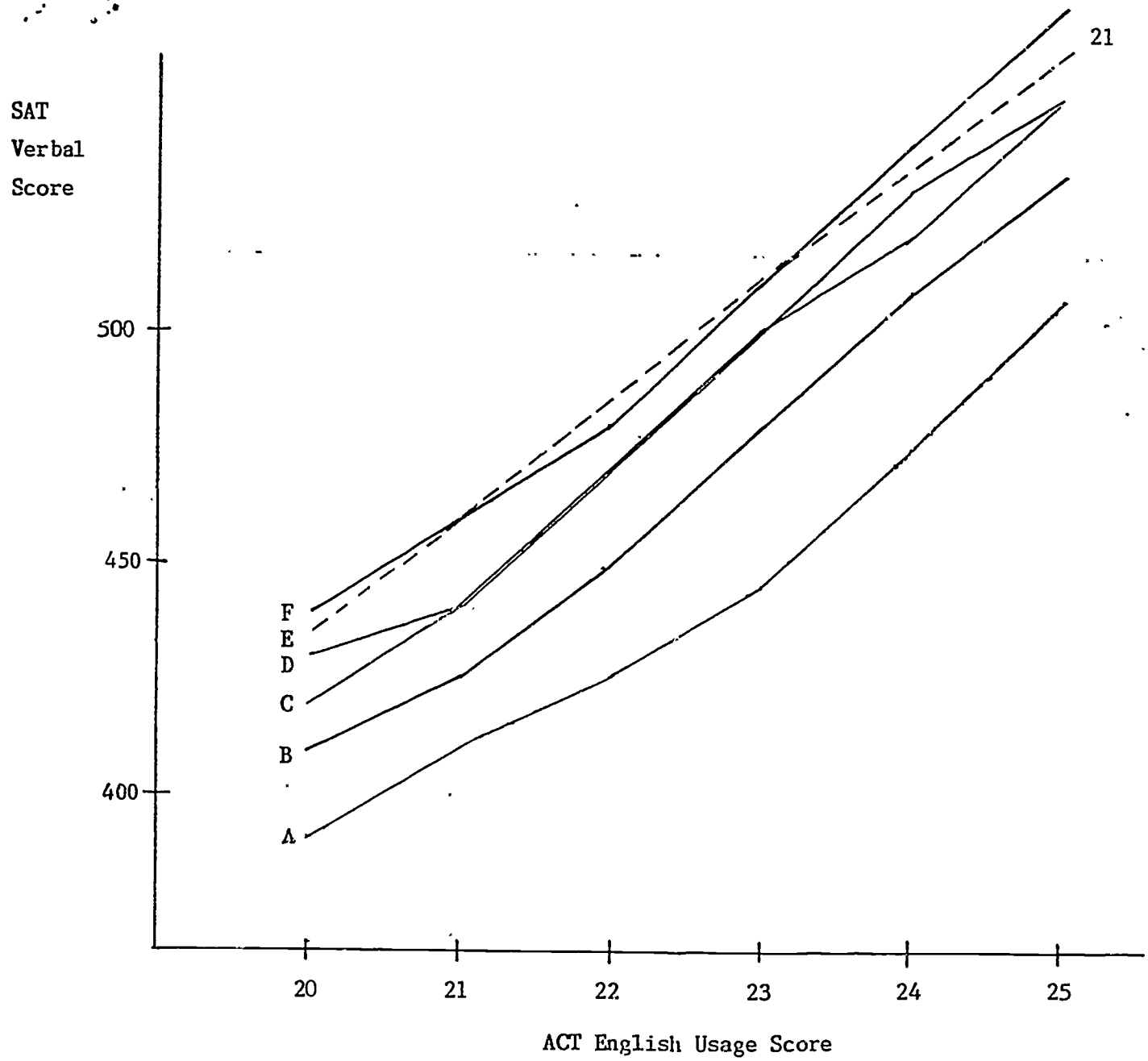
*N = 557

Table 6

Demographic Characteristics of Selected and Rejected Applicant Groups

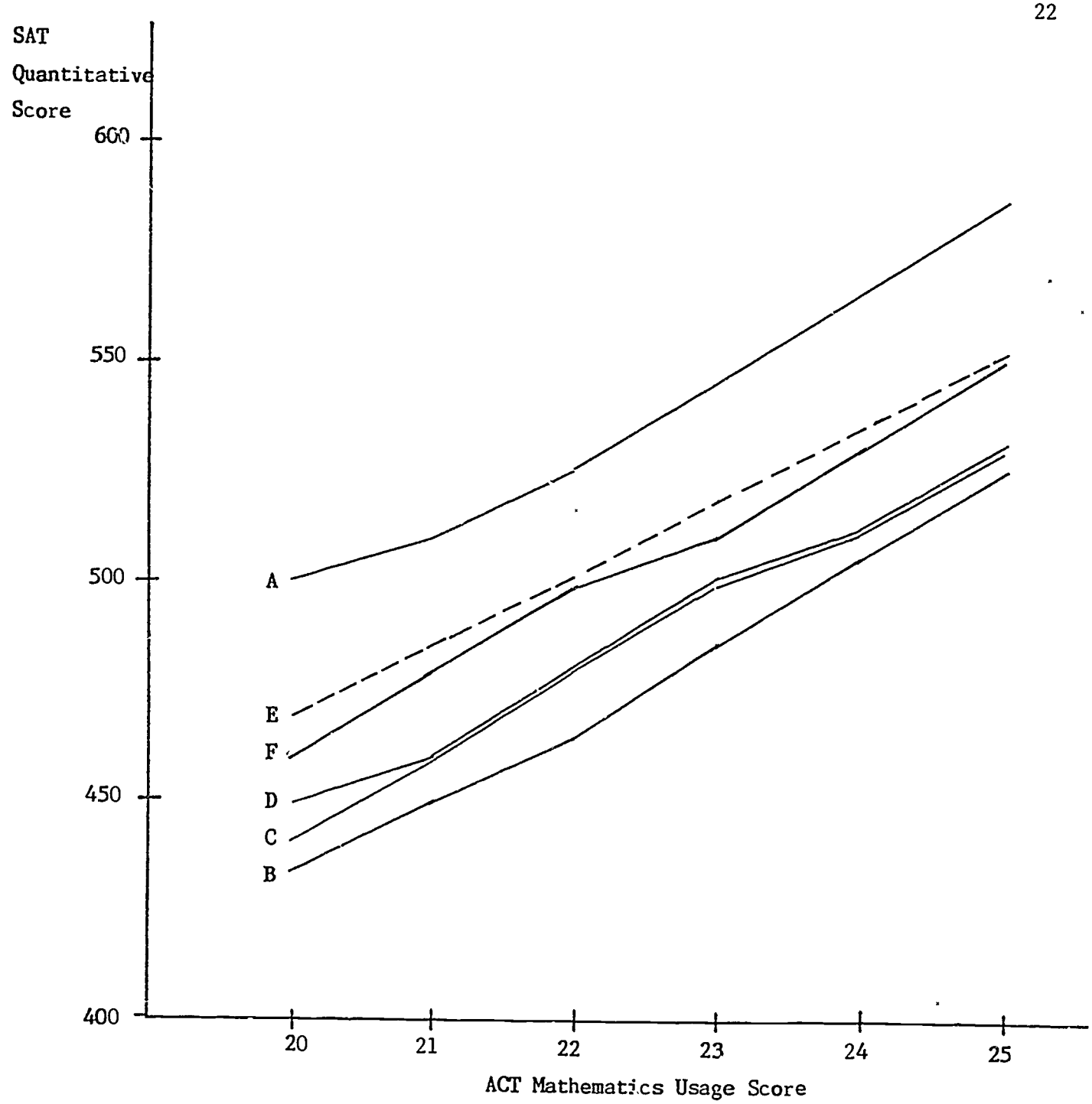
Percentile cutoff	Group	N*	Proportion in group who are... Females	Minority	From small town or rural h.s.
50	Selected by ACT English Usage	359	.56	.09	.44
	Selected by SAT Verbal	334	.45	.10	.48
	Selected by ACT English Usage, not by SAT Verbal	75	.67	.09	.46
	Selected by SAT Verbal, not by ACT English Usage	50	.19	.08	.64
	Selected by ACT Mathematics Usage	356	.41	.10	.43
	Selected by SAT Mathematics	342	.43	.10	.39
	Selected by ACT Mathematics Usage, not by SAT Mathematics	74	.45	.08	.43
	Selected by SAT Mathematics, not by ACT Mathematics Usage	60	.59	.08	.26
75	Selected by ACT English Usage	175	.48	.11	.40
	Selected by SAT Verbal	161	.54	.10	.49
	Selected by ACT English Usage, not by SAT Verbal	60	.50	.12	.37
	Selected by SAT Verbal, not by ACT English Usage	46	.94	.10	.54
	Selected by ACT Mathematics Usage	242	.36	.12	.42
	Selected by SAT Mathematics	234	.42	.13	.50
	Selected by ACT Mathematics Usage, not by SAT Mathematics	62	.38	.10	.53
	Selected by SAT Mathematics, not by ACT Mathematics Usage	54	.64	.15	.62
--	Total group	557	.47	.10	.47

*The counts for the various groups selected vary because the percentiles were selected from weighted relative frequency distributions and because the frequency distributions did not contain cumulative relative frequencies that coincided exactly with .50 and .75.



Key: A = Talent Identification Program (1987)
B = Langston (1987)
C = Mountains region public university (1982)
D = Midwest private university (1982)
E = West Coast public university (1980)
F = Southern public university (1982)

Figure 1: Comparable ACT English Usage and SAT Verbal Scores



Key: A = Talent Identification Program (1987)
B = Langston (1987)
C = Mountains region public university (1982)
D = Midwest private university (1982)
E = West Coast public university (1980)
F = Southern public university (1982)

Figure 2: Comparable ACT Mathematics Usage and SAT Mathematics Scores

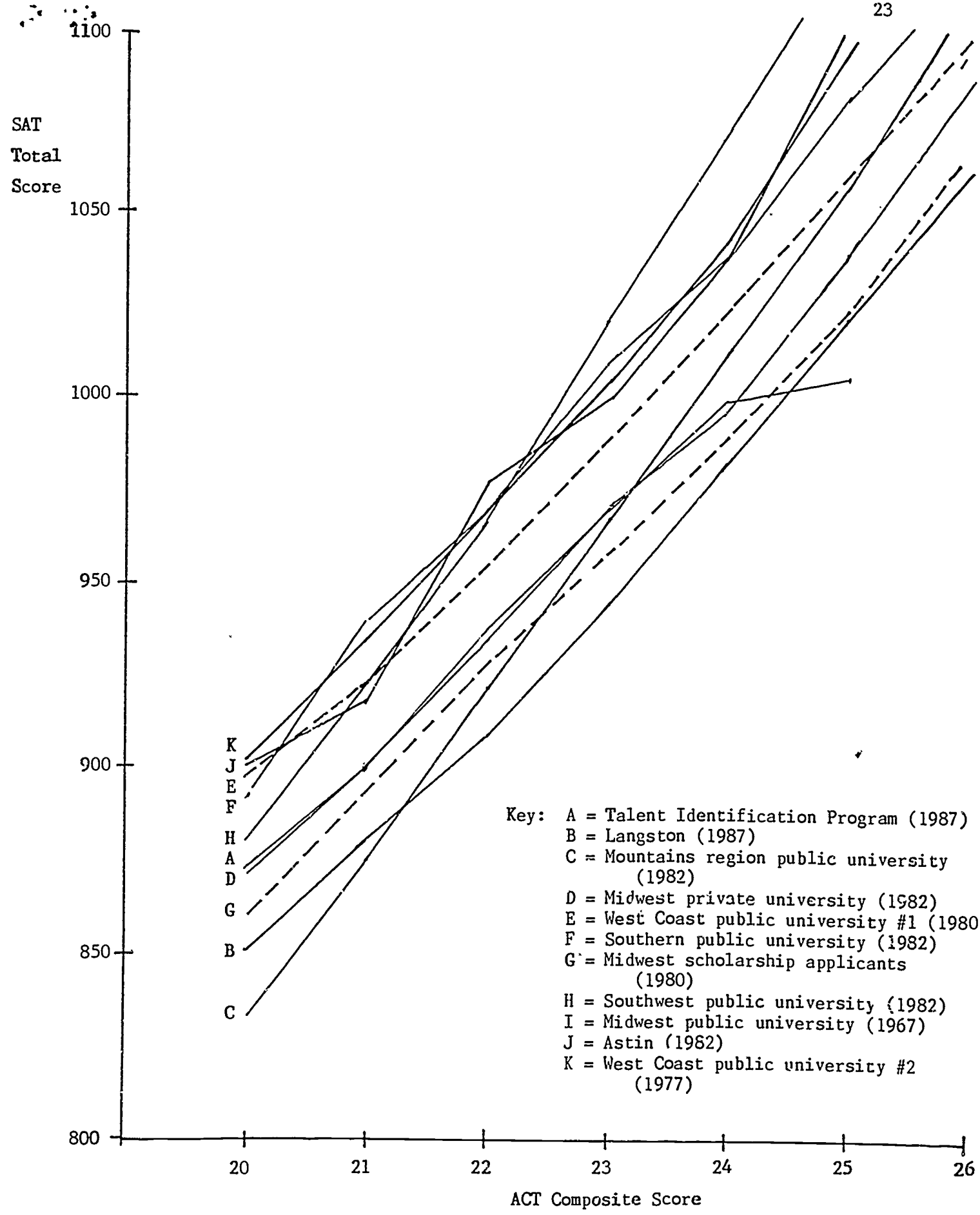


Figure 3: Comparable ACT Composite and SAT Total Scores