DOCUMENT RESUME

ED 242 794 TM 840 224

AUTHOR	Baird, Leonard L.
TITLE	Predicting Predictability: The Influence of Student and Institutional Characteristics on the Prediction of Grades. College Board Report No. 83-5
INSTITUTION	Educational Testing Service, Princeton, N.J.
SPONS AGENCY	College Entrance Examination Board, New York, N.Y.
REPORT NO	CB-R-83-5; ETS-RR-83-30
PUB DATE	83
NOTE	16p.
AVAILABLE FROM	College Board Publications, Box 886, New York, New York 10101 ($5.00).
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	Academic Aptitude; Aptitude Tests; College Entrance Examinations; College Freshmen; *Correlation; *Grade Prediction; Grades (Scholastic); Higher Education; *Institutional Characteristics; Predictive Measurement; *Predictor Variables; *Statistical Analysis; *Student Characteristics; Test Validity
IDENTIFIERS	Scholastic Aptitude Test

ABSTRACT

This study examined the statistical and institutional influences on the prediction of first-year college grades. The sources of information were the Validity Study Services file which summarizes the results of College Board validity studies and the College Handbook file which includes data about college characteristics. The criterion was the size of the multiple correlation between academic predictors and first-year college grades. The independent variables were the statistical data of the validity study and college characteristics. In general it was found, as expected, that the extent of the variation of the academic ability of the students was positively related to the size of the multiple correlation. In addition, several variables suggested the interpretation that the heterogeneity of the programs and experience of college were negatively related to the size of the multiple correlation. Further analyses provided evidence for the influence of institutional characteristics on the prediction of grades, an influence that needs to be taken into accout when interpreting the "validity" of tests like the Scholastic Aptitude Test. (Author/PN)

College Board Report

No. 83-5

Predicting Predictability: The Influence of Student and Institutional Characteristics on the Prediction of Grades

Leonard L. Baird

Predicting Predictability: The Influence of Student and Institutional Characteristics on the Prediction of Grades

Leonard L. Baird

College Board Report No. 83-5

ETS RR No. 83-30

College Entrance Examination Board, New York, 1983

To Jerilee Grandy for preparing the analyses; to Marian Helms for typing the manuscript; to Robert Boldt and Hunter Breland for critical reviews of earlier versions of the manuscript; these contributions are acknowledged with appreciation.

Researchers are encouraged to express freely their professional judgment. Therefore, points of view or opinions stated in College Board Reports do not necessarily represent official College Board position or policy.

The College Board is a nonprofit membership organization that provides tests and other educational services for students, schools, and colleges. The membership is composed of more than 2,600 colleges, schools, school systems, and education associations. Representatives of the members serve on the Board of Trustees and advisory councils and committees that consider the programs of the College Board and participate in the determination of its policies and activities.

Additional copies of this report may be obtained from College Board Publications, Box 886, New York, New York 10101. The price is $5.

CONTENTS

Tables

ABSTRACT

This is a report of a study that examined the statistical and institutional influences on the prediction of first-year college grades. The basic sources of information were the Validity Study Service file which summarizes the results of College Board validity studies and the College Handbook file which includes data about college characteristics. The criterion was the size of the multiple correlation between academic predictors and first-year college grades. The independent variables were the statistical data of the validity study and college characteristics. In general it was found, as expected, that the extent of the variation of the academic ability of the students was positively related to the size of the multiple correlation. In addition, several variables suggested the interpretation that the heterogeneity of the programs and experience of college were negatively related to the size of the multiple correlation.

Further analyses investigated the characteristics that were associated with the greater or lesser efficiency of SAT-verbal scores and SAT-mathematical scores in the prediction of grades. Similar analyses examined the characteristics associated with greater or lesser efficiency of SAT scores and the high school record. Finally, the prediction of grades at colleges of different selectivity was examined. In all of these analyses evidence was provided for the influence of institutional characteristics on the prediction of grades, an influence that needs to be taken into account when interpreting the "validity" of tests like the SAT.

INTRODUCTION

The validity of the Scholastic Aptitude Test (SAT) and high school grades—the correlation between these variables and college grades—has long been a concern of users of the SAT, the College Board and Educational Testing Service, and critics of standardized testing. The college user of the SAT and grade records is chiefly concerned with the validity of the test and grades at his or her own campus. The College Board and Educational Testing Service look at the validity of the SAT as one of the criteria of the quality of the instrument. Finally, one of the criticisms of test critics is that the SAT does not adequately predict grades.

Each of these groups generally thinks in terms of the prediction of *individual* grades, and a great deal of research has been done on the predictability of an individual's grades and the personal characteristics that are related to prediction. This research has included over- and under-prediction, under- and over-achievement, moderator variables, and so on. This large literature has also focused on the validity of tests among different populations (Breland 1978), the difficulty of long-term prediction (Humphreys 1968; Wilson 1983), the reasons, such as the diversity of grades, that grades are difficult to predict (Goldman and Slaughter 1976), and most recently, the question of validity generalization. However, most of the research on the prediction of college grades has neglected the *institutional* characteristics that influence the predictability of college grades. For example, on both statistical and logical grounds, the range of SAT scores and high school grades in a class should have a marked effect on the predictability of college grades. Linn (1980) has demonstrated that using data from Schrader (1977), the size of the correlation between LSAT scores and first-year law school grades is appreciably influenced by the size of the standard deviation of the LSAT scores of a school's students.

The purpose of the research reported here is to examine the influence of a variety of institutional and student body characteristics on the prediction of college grades. Some of these influences are quite understandable. For example, Ramist (1980) has shown that the size of the SAT-freshman grade correlation is affected by the size of the SAT-standard deviation. Among colleges with standard deviations of 100 or higher, the mean correlation is .50; among colleges with standard deviations below 80, it is .36. This result would be expected on statistical grounds. It also would be expected on logical grounds: the greater the range of academic talent among the students entering an institution, the greater the range of grades that might be expected to be awarded, and the greater the potential role of academic ability in determining grades. Thus, whatever increases the heterogeneity of academic input is likely to increase the predictability of grades. Similar arguments could be made about the *relative* power of predictors to predict college grades, as Dawes (1975) has argued. For example, if a college is very selective on SAT-mathematical scores, the verbal scores may be the only predictor with sufficient variance to correlate with college grades.

Other research has shown that the predictability of grades is also affected by the heterogeneity of the academic programs of colleges. That is, the more diverse the curricula in terms of major fields, divisions, colleges, and so forth, the less predictable are grades (Munday 1970). In addition, the less homogeneous the college *experience*, the less predictable the grades. For example, grades at urban colleges enrolling many part-time and working students might be less predictable simply because these students' academic performance can be affected by many influences other than the institution's program. Munday (1970) found that grades at colleges with students living in college-controlled residences were more predictable than grades at colleges with commuting students, and Ramist (1980), after controlling for the size of the test standard deviation, found that grades in small colleges were more predictable than those in large colleges. It might also be expected that colleges with core curricula would have higher correlations than colleges with completely elective systems. It might further be expected, for example, that at colleges emphasizing mathematics and science in their curricula, the SAT-mathematical score might be a better predictor than the SAT-verbal.

Some characteristics of colleges that have been found to be related to the prediction of grades do not have very clear-cut explanations. For example, some studies have shown that the grades of four-year college students are more predictable than those of community college students, the grades of students in private colleges more predictable than those in public colleges, and the grades of students in high-cost institutions more predictable than those in low-cost institutions. Although it is possible to speculate as to the reasons for these differences, the differences will simply be noted as factors related to predictability.

Finally, and again expectedly on theoretical and logical grounds, the range of the criteria is related to predictability. Not only is the standard deviation of the grades awarded related, but also such factors as the percentage of pass/fail courses allowed students, policies on dropping courses, whether low grades can be expunged from the academic record by repeating a course, and so on. That is, the smaller the range of grades awarded and the less homogeneous and reliable the basis upon which grade averages are based, the lower their predictability. (Furthermore, Breland [1979] and Werts, Linn, and Joreskog [1978] report single year GPA reliability coefficients of about .60, which clearly limit the size of any correlation.)

In sum, the size of the correlation between admissions tests such as the SAT and grades is influenced by many factors that have nothing to do with the intrinsic validity of the tests themselves. Some of these factors, such as the range of scores, have clear statistical explanations; some, such as the percentage of students in college-controlled housing, have sociological explanations; and some, such as the percentage of women in the class, do not have clear explanations although they can have a marked effect on predictability. Whatever their underlying mechanism, these factors can appreciably increase or reduce the size of the correlation between test scores and grades.

In the current discussion about the validity of tests, it would be useful to demonstrate how the validity of the tests—the correlation between test scores and freshman grades—can be influenced by various factors. The point is not to argue that the correlation in the most predictable colleges is the "real" one, but that correlations must be interpreted within the context of the characteristics of the institution. These characteristics put limits on the size of the correlations that are possible and expectable. There are conditions in which the correlation could not be expected to be high. Obviously, each correlation must be evaluated in terms of local conditions. Also, it is clear that it does not make much sense to use an "average" correlation coefficient as an indicator of the expected validity of a test, particularly since some of the very characteristics that are associated with participating in the validity service, such as selectivity, also may be associated with lower validity coefficients. Thus, this study attempted to demonstrate the influence of institutional characteristics on the obtained and expectable correlations between SAT scores and grades, and to examine the pattern of validity results as they relate to these characteristics.

METHODS

The basic data for this study were obtained from the College Board's Validity Study Service (VSS) files for 1981 and the files for the *The College Handbook* for 1981.

The VSS files provide information about the following:

1. Number of students in the study
2. Mean SAT-verbal and SAT-mathematical scores
3. High school record, reported in four different ways: (a) college reported high school grades; (b) college reported high school rank; (c) student reported high school average; and (d) student reported class rank
4. Standard deviations for SAT-verbal and SAT-mathematical scores
5. Standard deviations for the four high-school-grade-performance measures
6. The mean college GPA (CGPA) earned by the students in each study
7. The standard deviation of the CGPA in each study
8. The multiple correlations of the two SAT scores with CGPA
9. The multiple correlations of the two SAT scores and high school grades with CGPA

The College Handbook file includes information about the following:

10. Full-time undergraduate enrollment
11. Part-time undergraduate enrollment
12. Total undergraduate enrollment
13. Total graduate enrollment
14. Minimum GPA for good standing
15. Percentage of freshmen in good standing
16. Percentage of transfers in good standing
17. Number of men who applied to the freshman class
18. Number of men who were accepted as freshmen
19. Number of men who enrolled as freshmen
20. Number of women who applied to the freshman class
21. Number of women accepted as freshmen
22. Number of women who enrolled as freshmen
23. Percentage of students from within the state
24. Percentage of students who live in college housing
25. Percentage of students who live at home and commute
26. Percentage who live in private housing
27. Percentage of all students who commute
28. Percentage of students who are minority students
29. Percentage of students who are foreign
30. Percentage of students majoring in architecture
31. Percentage of students majoring in liberal arts
32. Percentage of students majoring in business

33. Percentage of students majoring in education
34. Percentage of students majoring in engineering
35. Percentage of students majoring in arts
36. Percentage of students in health sciences
37. Percentage of students in home economics
38. Percentage of students in religion
39. Percentage of students in technology
40. Percentage of students who transferred from two-year colleges
41. Percentage of students who transferred from four-year colleges
42. Percentage of graduates of two-year programs who enter four-year programs
43. Percentage of graduates of four-year programs who enter graduate programs
44. Number of communication and publication activities available
45. Number of performing arts activities available
46. Number of sports activities available
47. Tuition and fees
48. Tuition and fees and room and board
49. Number of freshmen receiving financial aid
50. Percentage of freshmen who are offered full financial aid

The previous variables are all continuous variables. In addition, a number of categorical variables from *The College Handbook* file were included in the analyses:

51. College level: two-year or four-year
52. Public or private college
53. Liberal arts college or university
54. Catholic or Protestant college
56. Type of academic calendar: semester, quarter, trimester, or 4-1-4
57. Location: metropolis, large city, small city, suburb, large town, small town, rural
58. Offers AA degree or does not
59. Offers BA degree or does not
60. Offers academic acceleration or does not
61. Offers honors or does not
62. Allows student to design major or does not
63. Allows double major or does not
64. Offers independent study or does not
65. Provides internships or does not
66. Provides cooperative education or does not
67. Offers study abroad or does not
68. Offers combination program of liberal arts and career program or does not
69. Provides remedial courses or does not
70. Provides tutoring or no tutoring
71. Provides counseling or no counseling
72. Has a learning center or does not
73. Gives credit to AP scores greater than 3, versus giving credit to lower scores
74. Allows CLEP credits or does not
75. Requires a minimum GPA versus no minimum
76. Requires interview for admission or no interview
77. Open admission policy or no open admission
78. Application deadline: January-April, May-September, no closing date
79. Has fraternities and/or sororities, or has neither

Statistics

Because some institutions conduct multiple validity studies, whereas others conduct only one, there are two ways to examine this data. The first is to use the validity study as the unit of analysis. This approach concentrates on the variables that affect the results of the study. The second is to use the *institution* as the unit of analysis. This approach concentrates on identifying the types of institutions that tend to have higher or lower validity coefficients. Both types of analyses were conducted in this study. The institutional sample was constructed by identifying the most representative study conducted by each institution and using it with other data about the institution.

Because the results of multiple regression analyses are affected by the number of predictors, studies were also chosen for the analyses that had only the three basic predictors, SAT-verbal scores, SAT-mathematical scores, and high school grades. In addition, only those studies that included all freshmen rather than subgroups were included. Finally, the studies that appeared to be the most representative of the total freshman classes were included.

Variables 1 through 9 were correlated, using the validity study as the unit of analysis for examining their relationships. The two multiple correlations (SAT-verbal and SAT-mathematical with grades, and SAT-verbal and SAT-mathematical plus the high school record) were considered the dependent variables; the other variables were considered independent variables. These analyses were repeated for variables 1 through 50 using the institution as the unit of analysis. The mean multiple correlations for each categorical variable (51-79) were also calculated using the validity study as the unit of analysis. A difference of .03 in the mean multiple correlation was considered important enough to note.

Further analyses were conducted using the multiple correlations as the dependent variables. Multiple regression was used to study the relationship between the other variables and the size of the multiple correlations to predict freshmen grades.

Additional analyses examined the question of the relative effectiveness of the various predictors in predicting college grades. This was done by creating a variable which was the difference between the correlation of SAT-verbal scores with college grades ($r_{V.CGPA}$) and the correlation of SAT-mathematical scores with college grades ($r_{M.CGPA}$), or ($r_{V.CGPA} - r_{M.CGPA}$). Then this variable was correlated with the other statistical and institutional variables. That is, the analyses were designed to address the question as to why *V* is a relatively better predictor in some colleges, whereas *M* is a better predictor in others.

Similar analyses were conducted to examine the possible reasons for the relative effectiveness of SAT scores and the high school record in predicting college grades. Two variables were created, subtracting the correlation of verbal scores from the correlation of the high school record ($r_{HSGPA\ CGPA} - r_{V\ CGPA}$) and subtracting the correlation of mathematical scores from the correlation of the high school record ($r_{HSGPA\ CGPA} - r_{M\ CGPA}$). These variables were then correlated with the statistical and institutional variables.

Finally, because of the widely held belief that the prediction of grades is inherently more difficult at selective institutions, analyses of the patterns of validity results at colleges with different mean combined verbal and mathematical scores were conducted.

RESULTS

Table 1 shows the correlations between the statistical characteristics of the students at each college and the size of the two multiple correlations: (1) SAT-verbal scores and SAT-mathematical scores with college grades (R_2); and (2) SAT-verbal scores, SAT-mathematical scores, and high school grades with college grades (R_3). The first two variables, reflecting the average level of ability of the students in the college—mean SAT-verbal and SAT-mathematical scores—were not meaningfully related to the prediction of grades. This result does not confirm Munday's (1970) finding that the level of ability was positively related to the prediction of grades. However, the range of academic ability as reflected in the standard deviation of the SAT scores of the students in the college was related to both multiple correlations.

Although the standard deviation of the college grades obtained in each study was expected to have a substantial influence on the multiple correlations, it did not. The mean college grade awarded across the studies was also unrelated to the size of the multiple correlations.

Table 1. Correlations with Multiple Correlations Predicting College Grades

	Correlation with:			
	R_2†	N	R_3††	N
SAT-V mean	−.08	155	−.03	177
SAT-M mean	−.13	155	−.07	177
SAT-V standard deviation	.31**	155	.25**	177
SAT-M standard deviation	.30**	155	.29**	177
College GPA mean	.10	150	.13	172
College GPA standard deviation	−.10	150	−.04	172
HS rank mean	−.21*	85	−.18	97
HS GPA mean	.16	70	.11	80
HS rank standard deviation	.14	85	.20*	97
HS GPA standard deviation	−.13	70	−.13	80

Note: Only studies that used high school GPA and college grades on a 0 to 4 scale were used.

†R_2 is the multiple correlation of SAT-verbal and SAT-mathematical scores with college grades.

††R_3 is the multiple correlation of SAT-verbal scores, SAT-mathematical scores, and high school grades with college grades.

*This correlation is significant at the .05 level.

**This correlation is significant at the .01 level.

High school grades can be reported as ranks or average grades in the Validity Study Service, which may help to explain the disparity in results. The correlation between mean high school GPA and the multiple Rs are significant for R_2. The standard deviation of high school rank is significantly related to R_3 but not to R_2. In contrast, the standard deviation of high school grades is negatively related to both R_2 and R_3, but not significantly.

The significant correlations between the multiple coefficients predicting grades and *institutional characteristics*, using the institution as the unit of analysis, are shown in Table 2. A number of variables suggest that the more diverse the student body in nonacademic terms, the lower the multiple correlation. For example, undergraduate enroll-

Table 2. Correlations Between Institutional Characteristics and Validity Coefficients

	N	R_2	N	R_3
Enrollment Variables				
Full time undergraduate enrollment	152	−.13	174	−.16*
Total undergraduate enrollment	155	−.16	177	−.19**
Total enrollment	97	−.13	115	−.18*
Residence Variables				
Percentage who live at home and commute	92	−.07	106	−.15
Percentage living in college housing	146	.17*	167	.20**
Minority Students				
Percentage of enrollment that is minority	150	−.03	170	−.14*
Activity Variables				
Number of communication and publishing activities available	155	−.11	177	−.08
Number of sports activities available	146	−.27**	177	−.16*
Aid Offered				
Percentage of freshmen in financial need	116	−.15*	133	−.18*
Percentage of freshmen offered full financial aid	78	−.20*	90	−.17*

*This correlation is significant at the .05 level.

**This correlation is significant at the .01 level.

Table 3. Variables Associated with a Difference in Multiple Correlations of .03 or More

Variable	*N*	R_2	*Difference*	*N*	R_3	*Difference*
College Type						
Liberal Arts College	285	.40		87	.56	
University	648	.41	.00	288	.53	.03
Women's College	43	.41		20	.51	
Coed Institution	964	.40	.01	366	.54	.03
Catholic College	146	.42		64	.56	
Protestant College	115	.43	.01	31	.61	.05
Location						
Metropolitan	157	.36		51	.48	
Large City	142	.43		59	.56	
Suburb	231	.42		64	.54	
Small City	288	.41	.08	135	.53	.18
Large Town	62	.39		23	.48	
Small Town	114	.40		50	.61	
Rural	13	.44		5	.66	
Academic Programs						
Acceleration Offered	570	.41		222	.55	
Acceleration Not Offered	438	.39	.02	165	.52	.03
Honors Offered	750	.40		326	.53	
Honors Not Offered	258	.42	.02	61	.58	.05
Student-Designed Major Allowed	624	.40		234	.52	
Not Allowed	384	.42	.02	153	.56	.04
Double Major Allowed	852	.40		368	.53	
Not Allowed	156	.42	.02	19	.59	.06
Individualized Study	907	.40		375	.54	
No Individualized Study	101	.46	.06	12	.52	.02
Internships Available	866	.41		359	.55	
Not Available	142	.37	.04	28	.42	.03
Year Abroad Available	871	.40		366	.54	
Not Available	137	.45	.05	21	.54	.00
Tutoring Offered	818	.39		321	.54	
Not Offered	190	.44	.05	66	.58	.05
Allows CLEP Credit	191	.40		57	.61	
Does Not Allow Credit	817	.40	.00	330	.52	.09
Admissions Procedures						
Open Admissions Policy	59	.48		37	.59	
Not Open Policy	949	.40	.08	350	.53	.06
Applications Accepted October-May	196	.35		49	.50	
Applications Accepted June-September	332	.43	.09	118	.54	.04
No Closing Date	472	.40		211	.55	.05
Requires Interview	50	.40		22	.63	
Does Not Require, but Recommends Interview	500	.40	.00	228	.52	.09

ment is negatively related as are the percentage of students who are minority students (for R_3), the percentage of freshmen who are in financial need, and the percentage offered financial aid. These variables all suggest human diversity.

Similarly, the greater the variety of activities available, especially sports activities, the lower the multiple correlation. This may be related to the overall diversity of the institutions' offerings. It may also reflect the number of distractions from academic studies.

In contrast, the percentage of students who live in college housing is positively related to both R_2 and R_3. This may also represent a lower number of distractions from college life.

The average R_2 and R_3 figures for *studies* conducted in colleges of different types were also calculated. The categorical variables associated with differences in validity study multiple correlations of .03 or more are shown in Table 3. The variables are grouped into clusters of related variables. The results suggest that grades were more predictable in studies conducted at liberal arts colleges than at universities, at coeducational institutions than at women's colleges, and at Protestant colleges than at Catholic colleges. It is easy to understand that universities have much more diverse programs than liberal arts colleges, and that grades therefore are based on more kinds of courses and are therefore less predictable. The other differences are more difficult to understand. Catholic colleges include a number of large, urban institutions, whereas there are few Protestant institutions of that type, which may account for the difference.

There is a large difference between the average *R*s for studies conducted in colleges in metropolitan areas and those in small towns and rural areas. There seems to be a general, but not perfect, negative association between the size of the Rs and the complexity of the communities surrounding the colleges. Presumably there are fewer commuting and part-time students in colleges in smaller settings. There are also probably fewer distractions from studies.

A considerable number of variables involving the academic program were related to differences in the multiple correlations. In general, the more flexible an institution, the lower the correlation. For example, studies conducted in colleges that offer honors programs, that allow the students to design their own majors, that allow double majors, and that offer tutoring resulted in lower Rs on the average. However, studies in colleges that offer acceleration and internships tend to obtain higher Rs. Studies in colleges that allow CLEP credits also have higher Rs.

Admissions procedures are also related to the magnitude of the multiple correlation. Studies in colleges with open admissions policies tend to have higher Rs than do those in colleges that accept applications up to September or have no closing date. Additional analyses indicated that these college have lower average SAT scores and admit a broader range of students, as indicated by SAT standard deviations. Studies in colleges that require interviews obtain higher Rs than studies in colleges that do not require interviews. These results suggest that either heterogeneity of input or homogeneity of program are positively associated with the predictability of grades.

One of the interesting aspects of these results is the variety of variables that were *not* related to predictability, although they might be expected to be on statistical or logical grounds. These included the standard deviation of college grades, the minimum college GPA required for good standing, the percentage of freshmen in good standing, the percentage of out-of-state students, the percentage of foreign students, and the percentage of graduates who go on to graduate or professional school. This may be a credit to the robustness of the validity coefficient.

Obviously, there is probably collinearity among these variables. For example, the finding that the validity coefficients are lower in studies in urban institutions than in institutions in small towns or rural areas may be due to a wider range of academic ability in the urban institutions. For these reasons, the multiple regression analyses described earlier were conducted.

Multiple Regression Analyses

In order to examine the relative contribution of the variables we have studied, we conducted multiple regression analyses

Table 4. Prediction of R_2 and R_3 from Validity Study Data Using Four Grade Reporting Systems

R_2 *Results*	β	R_3 *Results*	β
Studies Using Student Reported Rank			
SAT-M—SD	.31	SAT-M—SD	.32
College GPA SD	.44	College GPA SD	.44
College GPA Mean	.42	College GPA Mean	.31
SAT-M Mean	−.31	HS GPA SD	.25
N = 152, R = .62		N = 152, R = .67	
Studies Using Student Reported HS GPA			
HS GPA SD	.44	College GPA SD	−.48
College GPA SD	−.46	HS GPA SD	.33
SAT-M SD	.14	HS GPA Mean	.45
SAT-V Mean	−.37	SAT-V Mean	−.42
HS GPA Mean	.25	SAT-M SD	.07
N = 146, R = .49		N = 146, R = .44	
Studies Using College Reported Rank			
SAT-V SD	.34	SAT-M SD	.40
SAT-M Mean	−.34	College GPA SD	−.29
College GPA SD	−.19	SAT-M Mean	−.22
SAT-M SD	.25	SAT-V SD	.12
HS GPA SD	−.14	HS GPA SD	−.10
HS GPA Mean	.14	HS GPA Mean	.08
N = 379, R = .53		N = 379, R = .68	
Studies Using College Reported HS GPA			
SAT-V SD	.44	SAT-V SD	.42
HS GPA SD	−.22	HS GPA SD	−.36
SAT-V Mean	−.26	SAT-V Mean	−.27
College GPA	.20		
N = 186, R = .57		N = 186, R = .57	

Table 5. Multiple Regression Analyses for Institutional Characteristics

	R_2 Results	β	R_3 Results	β
All Studies and All Groups	N = 513, R = .43		N = 513, R = .46	
	SAT Math Standard Deviation	.39***	SAT Math Standard Deviation	.41***
	Number of Sports Activities	−.06	Enrollment	−.13**
	Enrollment	−.07	Number of Sports Activities	.07
	Majors in Education	.07	Majors in Education	.11**
	Majors in Engineering	−.07	Majors in Engineering	−.04
	Location	.11**	Location	.15***
Institutional Sample only	N = 36, R = .70		N = 35, R = .82	
	SAT Math Standard Deviation	.46**	SAT Math Standard Deviation	.46**
	Number of Sports Activities	−.35*	Enrollment	.01
	Enrollment	.01	Number of Sports Activities	−.27*
	Number of Female Applicants	−.28	Number of Female Applicants	−.65*
	Percentage in College Housing	.17	Percentage in College Housing	.16
	Number in Financial Need	.12	Number in Financial Need	.13
			Location	.24*

*This correlation is significant at the .05 level.
**This correlation is significant at the .01 level.
***This correlation is significant at the .001 level.

using R_2 and R_3 as the criteria. In the first set of analyses, the relative influence of the basic statistical prediction information was examined. In the second set, the role of institutional characteristics was also examined.

Statistical data

First, we examined the relationships between the standard statistical data and the size of the multiple correlations, using R_2 and R_3 as the dependent variables. The VSS allows four high school grade systems to be used: (1) student-reported high school rank, (2) student-reported high school GPA, (3) college-reported high school rank, and (4) college-reported high school GPA. Therefore, eight separate regression analyses were conducted. The results of the stepwise multiple regression analyses are shown in Table 4. Variables were selected until the multiple correlation no longer increased by at least .01. The positive predictor appearing most consistently across all analyses was the SAT-mathematical standard deviation, which occurred in six of the analyses; the SAT-verbal standard deviation, which appeared in four of the analyses, was the next most consistent positive predictor. However, it is clear that there is so much collinearity in the data that some other variables have erratic and surprising values in the equations. For example, the standard deviations of college grades had a positive value in two equations and a negative value in four. The standard deviation of high school grades had a positive value in three equations and a negative value in four. These results are almost certainly due to the fact that several variables reflect the same underlying variable, the range of academic ability in the freshman class, and the method of analysis used—stepwise multiple regression.

Analyses of institutional characteristics and statistical data

Because there was so much collinearity among the measures of the variability of the academic ability of the students, only one was used in these analyses. This was the standard deviation of SAT-mathematical scores, which was chosen because it had had the most consistent relationship to the size of the multiple correlation across the studies. The remaining institutional variables were chosen because they had the strongest relationships with the multiple correlations, and also seemed to most adequately represent a domain. For example, total enrollment was chosen rather than undergraduate enrollment since it seemed more adequately to represent the diversity of programs and people in an institution. The remaining variables chosen were: the number of sports activities, the percentage of students enrolled in education, the percentage of students enrolled in engineering, and the location of the institution (metropolis was high, rural location was low).

The results for the institutional sample for all studies are shown in Table 5. In this case, the full multiple regression results are shown. (The interested reader can find the zero order correlations in other tables.) In the case of the institutional sample, the standard deviation of the SAT-mathematical scores and the number of sports activities had significant beta weights with the size of both multiple correlations predicting grades. (The *N*s for the institutional sample were reduced considerably because *The College Handbook* file does not clearly distinguish between zeros and no data reported. Therefore, the number of cases that clearly had complete data was fairly small.)

In the case of R_2 across all the studies, only SAT-

Table 6. Correlations Between Statistical and Institutional Characteristics and the Difference Between the Correlations of Verbal and Mathematical Scores with First-Year Grades ($r_V - r_M$)

Variable	N	Correlation
SAT-Verbal Mean	172	−.02
SAT-Mathematical Mean	172	.15*
SAT-Verbal Standard Deviation	172	.08
SAT-Mathematical Standard Deviation	172	−.04
High School Grades	80	.22*
Percent Minority Students	165	−.17*
Percent Enrolled in Engineering	44	−.33*

*p<.05

mathematical standard deviations and location ("large city" was low, "rural" was high) were significantly related. For R, across all studies, these same variables were significantly related to the size of the multiple correlation, as well as to total enrollment, which had a negative weight, and to the percentage of students majoring in education, which had a positive weight.

Because of the variations allowed in the reporting of high school performance, multiple regressions were also prepared for each of the four grading systems across all studies. The results, shown in the Appendix, demonstrate that the variables most consistently related to the size of the multiple correlation coefficient were the standard deviation of the SAT-mathematical scores, total enrollment, the number of sports activities, and location. One interpretation of these results is that they show the influence of the variability of the academic ability of the incoming students (SAT-mathematical standard deviation), the variability of programs and thereby the meaning of grades (enrollment), the distractions from academic subjects, and the possible lower emphasis on scholarship (number of sports activities), and the distractions within the community, as well as the number of commuting and part-time students (location).

Analyses of the relative predictive effectiveness of *V* and *M*

Why does the SAT-verbal score predict college grades better in some colleges and the SAT-mathematical score predict better in others? To answer this question, we examined the differences in the correlations of SAT-verbal and SAT-mathematical scores with college grades. That is, we calculated the difference between $r_{V\,CGPA}$ and $r_{M\,CGPA}$. Then we correlated this figure with the statistical and institutional characteristics of the colleges conducting the study. As shown in Table 6, SAT-verbal scores tended to predict better than SAT-mathematical scores in colleges that had *higher* mean mathematical scores and students with higher high school grades. The SAT-mathematical scores tended to be better predictors in colleges with higher proportions of minority students, and with large percentages of students enrolled in engineering. Interestingly, the differences in the power of *V* and *M* to predict grades were not related to the standard deviation of these scores. This is probably owing to the fact that most colleges tend to select students on both scores; that is, the diversity of students would be about the same on both measures.

The tendency for verbal scores to be a better predictor at colleges with higher mathematical scores is hard to interpret. It may be, as Dawes (1975) has suggested, there is

Table 7. Correlations Between Statistical and Institutional Characteristics and the Differences Between the Correlations of High School Grades and SAT Scores with First-Year College Grades ($r_H - r_V$, $r_H - r_M$)

Variable	N	Correlations with: $(r_H - r_V)$	Correlations with: $(r_H - r_M)$
Statistical Variables			
N in Study	172	−.21**	−.21**
SAT-V Mean	172	.06	.11
SAT-M Mean	172	.08	.11
SAT-V Standard Deviation	172	−.09	−.10
SAT-M Standard Deviation	172	−.08	−.07
High School Rank Mean	92	.35**	.36**
High School GPA Mean	80	.03	.16
High School Rank Standard Deviation	92	−.21*	−.16
High School GPA Standard Deviation	80	−.13	−.11
Institutional Variables			
Total Graduate Enrollment	112	−.22*	−.20*
Percentage of Commuters	68	−.20	−.23*
Percentage of Minority Students	165	−.11	−.17*
Percentage Enrolled in Engineering	44	.21	.04
Percentage Enrolled in Art	59	−.14	.17
Percentage Going on to Graduate School	103	.16	.20*
Number of Students Offered Full Financial Aid	118	−.19	−.22*

*p<.05
**p<.01

Table 8. Mean Statistical Characteristics of Institutions of Different Selectivity

	Selectivity by V and M			
Variable	*<800*	*801-900*	*901-1,000*	*>1,000*
SAT-V Standard Deviation	78	87	90	89
SAT-M Standard Deviation	79	91	94	87
High School Rank Mean	55	56	56	63
High School GPA Mean	2.8	2.9	3.1	3.2
High School Rank Standard Deviation	8.6	7.3	8.9	6.4
High School GPA Standard Deviation	.52	.56	.51	.51
College GPA Mean	2.2	2.5	2.6	2.7
College GPA Standard Deviation	.8	.8	.7	.6
R, V and M	.41	.45	.42	.40
R, V and M and H	.50	.56	.57	.51
r_V	.34	.39	.37	.33
r_M	.37	.39	.35	.32
r_H	.38	.47	.48	.42
N	19	55	57	41

Note: High School Rank is placed on a standard scale ranging from 20 to 80; High School GPA is on a scale from 0.0 to 4.0.

compensatory selection which results in a restriction in range on mathematical scores, although the correlation with the SAT-mathematical standard deviation was not significant.

There is a suggestion of an explanation in the related finding that studies conducted at colleges where higher high school grades are required also show verbal scores to be better predictors. It is possible that colleges that have traditional letters and science curricula tend to emphasize reading and writing skills that draw on skills reflected in the SAT verbal scores. These colleges also tend to have higher academic standards.

A similar explanation may apply to the results showing that colleges with high minority enrollments and with large proportions of engineering students tended to find SAT-mathematical scores to be relatively better predictors. As documented elsewhere (Baird 1981), minority students tend to enroll in technical fields, and clearly, if curricula such as engineering and technical programs emphasize mathematical skills, SAT-mathematical scores should be better predictors.

Analyses of the relative predictive effectiveness of SAT scores and the high school record

Why are high school grades better predictors in some colleges and SAT scores better predictors in others? To answer this question, we first calculated the difference between the correlation of high school grades or rank with college grades and the correlations of SAT-verbal and mathematical scores with college grades ($r_{HSGPA.CGPA} - r_{V.CGPA}$) and ($r_{HSGPA.CGPA} - r_{M.CGPA}$). We then correlated these differences with the statistical and institutional characteristics of the colleges in the validity studies. The results are shown in Table 7. Oddly, the idea that selection on one variable leads to a restriction of range on that variable and to a lower correlation did not seem to work. Although the mean high school rank is positively related to the correlations, that is, the high school rank is a better predictor in colleges with students ranking high in high school, the high school rank is a better predictor at colleges with a *narrower* range of high school ranks.

The other variables provide some suggestions that the SAT is a relatively better predictor at colleges with large graduate programs, relatively fewer commuters, minority students, and students supported by financial aid. All this suggests that high school grades are better predictors at diverse colleges or universities that do not emphasize graduate study. However, high school grades tended to be better predictors at colleges that sent large proportions of students on to graduate school.

Analyses of the influence of selectivity upon the pattern of validity results

Colleges were grouped by their combined mean SAT-verbal and SAT-mathematical scores. There were four groups: less than 800, 800 to 900, 901 to 1,000, and greater than 1,000. The general results, shown in Table 8, indicated that the multiple correlations predicting college grades tended to be higher in the middle two categories and lower in the least and most selective institutions. Although it is sometimes contended that the range of ability should be smallest in the most selective institutions, this was not the case in this sample; in fact, SAT scores were most truncated in the least selective institutions. Thus, the lower validity coefficients in the most selective institutions are probably due to other factors. One possibility is the relatively narrower range of college grades awarded to students in the most selective institutions, and the relatively narrower range of high school ranks represented among the freshmen.

DISCUSSION

These analyses show how the "validity" and relative predictive efficiency of SAT scores are affected by many conditions—some statistical, some social, and some educational. These conditions limit the size of the possible correlations and need to be considered when any correlation between the SAT and grades is considered. Thus, the analyses indicate how the size of the correlation can be influenced by many factors other than intrinsic validity of the scores as a measure of academic aptitude. The analyses of the correlates of the relative efficiency of SAT-verbal and SAT-mathematical scores, these scores and the high school record, and the analyses by college selectivity support this finding. This study provides strong indications that grades are predicted more and less efficiently for understandable reasons related to the statistical and institutional characteristics of the colleges.

The results of this study suggest that much could be learned by examining the basic validity data that affect the use of the SAT. The present analyses are fairly simple. Probably more could be understood by using analyses that are more statistically sophisticated. A current College Board project by Robert Boldt is, in fact, examining the question of validity generalization using very advanced methods. This project should lead to further insights.

The results of the current study also suggest that to understand the prediction of grades we need to pay attention to some fundamental considerations about the use of admissions tests, the characteristics of the students in any validity study, and the character of the educational experience in any specific college. Even a theoretically perfectly valid measure of academic aptitude could not be expected to provide perfect or even very good prediction in certain circumstances. As argued by Weitzman (1982), the SAT obviously provides highly useful information in situations where the prediction of academic performance is very difficult. In fact it may be those very situations where its contribution to the admissions prediction-decision situation is greatest. As the Taylor-Russell tables (1939) demonstrate, and as Cronbach and Glaser (1965) have pointed out ". . . a test validity of .20 in one situation may be more beneficial than a test validity of .60 in another."

REFERENCES

Baird, L. L. 1981. *Career Decisions of Minority Students and Entrance to Graduate Management Programs* (RR-81-43). Princeton, NJ: Educational Testing Service.

Breland, H. M. 1978. *Population Validity and College Entrance Measures*. New York: College Entrance Examination Board.

Cronbach, L. J., and Glaser, G. C. 1965. *Psychological Tests and Personnel Decisions*. 2d ed. Urbana: University of Illinois Press.

Dawes, R. M. 1975. Graduate Admission Variables and Future Success. *Science* 187, 721–723.

Goldman, R. D., and Slaughter, R. E. 1976. Why College Grade-Point Average Is Hard to Predict. *Journal of Educational Psychology* 60, 9–14.

Humphreys, L. B. 1968. The Fleeting Nature of the Prediction of College Academic Success. *Journal of Educational Psychology.* 59(5), 375–380.

Linn, R. L. 1980. Admission Testing on Trial. Paper presented at the Annual Meeting of the American Psychological Association.

Munday, L. A. 1970. Factors Influencing the Predictability of College Grades. *American Educational Research Journal* 7, 99–108.

Nairn, A., and Associates. 1980. *The Reign of ETS: The Corporation That Makes Up Minds*. Washington, D.C.: Ralph Nader.

Schrader, W. B. 1977. Summary of Law School Validity Studies, 1948–1975. In *Reports of LSAC Sponsored Research: 1975-1977*. Vol. III by Law School Admissions Council. Princeton, NJ: Law School Admission Council, 519–550.

Taylor, H. C., and Russell, J. T. 1939. The Relationship of Validity Coefficients to the Practical Effectiveness of Tests in Selection. *Journal of Applied Psychology* 23, 565–578.

Weitzman, R. A. 1982. The Prediction of College Achievement by the Scholastic Aptitude Test and the High School Record. *Journal of Educational Measurement* 19, 179–191.

Werts, C. E., Linn, R. L., and Joreskog, K. G. 1978. Reliability of College Grades from Longitudinal Data. *Educational and Psychological Measurement* 38, 89–145.

Wilson, K. M. 1983. *A Review of Research on the Prediction of Academic Performance After the Freshman Year.* (College Report No. 83-2). New York: College Board.

Appendix: Regression Results for Institutional and Statistical Studies for Variables Using Different Grading Systems

	R_2 Results	β	R_j Results	β
	N = 235, R = .58		N = 235, R = .62	
Studies Using College-Reported Rank	SAT Math standard deviation	.49***	SAT Math standard deviation	.51***
	Number of Sports Activities	.11	Enrollment	−.28***
	Enrollment	−.21**	Number of Sports Activities	.13*
	Majors in Education	−.09	Majors in Education	−.01
	Majors in Engineering	−.14**	Majors in Engineering	−.06
	Location	.09	Location	.13*
	N = 63, R = .43		N = 63, R = .61	
Studies Using Student-Reported Rank	SAT Math standard deviation	.29**	SAT Math standard deviation	.47***
	Number of Sports Activities	.13	Enrollment	−.39**
	Enrollment	−.32**	Number of Sports Activities	.32*
	Majors in Education	.07	Majors in Education	−.04
	Majors in Engineering	−.15	Majors in Engineering	−.06
	Location	−.05	Location	.07
	N = 114, R = .61		N = 114, R = .56	
Studies Using College-Reported GPAs	SAT Math standard deviation	.41***	SAT Math standard deviation	.34***
	Number of Sports Activities	−.51***	Enrollment	.26
	Enrollment	.35**	Number of Sports Activities	−.48***
	Majors in Education	.18*	Majors in Education	.06
	Majors in Engineering	−.15	Majors in Engineering	−.16
	Location	.19*	Location	.22**
	N = 101, R = .31		N = 101, R = .30	
Studies Using Student-Reported GPAs	SAT Math standard deviation	.05	SAT Math standard deviation	.02
	Number of Sports Activities	−.24	Enrollment	−.26
	Enrollment	−.01	Number of Sports Activities	.09
	Majors in Education	.10	Majors in Education	.08
	Majors in Engineering	.26	Majors in Engineering	.24
	Location	−.04	Location	.01

*This correlation is significant at the .05 level.
**This correlation is significant at the .01 level.
***This correlation is significant at the .001 level.