College Board Report No. 93-1
ETS RR No. 94-27

Student Group Differences in Predicting College Grades: Sex, Language, and Ethnic Groups

LEONARD RAMIST, CHARLES LEWIS,
and LAURA McCAMLEY-JENKINS

College Entrance Examination Board, New York, 1994

Leonard Ramist is research consultant/project director and former program director at ETS.
Charles Lewis is principal research scientist at ETS.
Laura McCamley-Jenkins is principal research data analyst at ETS.

Additional copies of this report may be obtained from College Board Publications, Box 886, New York, New York 10101-0886. The price is $12.

Printed in the United States of America.

Contents

Abstract 1

Part 1: Course Grade Study 1
- Investigation 1
- Course Selection and Grading 2
  - Grade Comparability 2
  - Grading Difficulty 2
- FGPA as the Criterion 3
  - By Academic Level 3
  - Predicting Validity for FGPA from Grade Comparability 3
- Course Grade as the Criterion 3
  - Compared to FGPA 3
  - SAT Predictive Validity by Type of Course 3

Part 2: Extending the Analysis 4
- To Additional Colleges 4
- By Student Group 4
- Course Selection and Grading 5
- Predictive Effectiveness 5
- Over- and Underpredictions 6
- Differences Among Colleges 6
- Statistical Precision of Results 6

Variations by Academic Composite Group 7
- Student Characteristics 7
- Course Selection and Grading 7
- Predictive Effectiveness 10
- Over- and Underpredictions 14
- Differences Among Colleges 16

Differences by Sex 18
- Student Characteristics 18
- Course Selection and Grading 18
- Predictive Effectiveness 19
- Over- and Underpredictions 23
- Differences Among Colleges 26

Differences by English as Best or Not Best Language 27
- Student Characteristics 27
- Course Selection and Grading 27
- Predictive Effectiveness 28
- Over- and Underpredictions 28
- Differences Among Colleges 28

Differences by Ethnic Group 28
- Student Characteristics 28
- Course Selection and Grading 28
- Predictive Effectiveness 31
- Over- and Underpredictions 32
- Differences Among Colleges 35

Summary of Student Group Differences 35
- Course Selection and Grading 35
  - Academic Composite 35
  - Sex 35
  - English as Best or Not Best Language 35
  - Ethnic Group 35
- Predictive Effectiveness 36
  - Academic Composite 36
  - Sex 36
  - English as Best or Not Best Language 36
  - Ethnic Group 36
- Over- and Underpredictions 37
  - Academic Composite 37
  - Sex 37
  - English as Best or Not Best Language 37
  - Ethnic Group 37

References 37

Appendix A: Colleges Participating in the Study 40

Appendix B: Course Categories 41

Tables

1. Student Composition of Each Academic Group 7
2. Courses Selected by Student Group 8
3. Course Selection and Grading Characteristics by Student Group 9
4. Predictive Effectiveness by Student Group 11
5. Average Correlations for All Students 13
6. SAT Increment by Type of Course 13
7. Comparison of the Effectiveness of Verbal and Mathematical Scores in Predicting Course Grade, by Type of Course 14
8. Average Over- (–) and Underpredictions (+) (Actual – Prediction) by Student Group, Using Prediction Equations for All Students 15
9. Largest Average Over- (–) and Underpredictions (+)

for High and Low Academic Composite Groups, by Type of Course, Using HSGPA and SAT Scores to Predict Course Grade *15*

10. Student Groups at More Selective and Less Selective Colleges and at Large and Small Colleges, by Academic Composite Group *16*

11. Characteristics of More Selective and Less Selective Colleges and of Large and Small Colleges, by Academic Composite Group *17*

12. Student Characteristics by Sex *18*

13. Comparisons of Courses Selected for High and Low Academic Composite Groups by Sex and for White Students Whose Best Language Is English *20*

14. Comparisons of Course Selection and Grading Characteristics for High and Low Academic Composite Groups by Sex and for White Students Whose Best Language Is English *21*

15. Comparisons of Predictive Effectivness for High and Low Academic Composite Groups by Sex and for White Students Whose Best Language Is English *22*

16. Comparisons of Over- (–) and Underpredictions (+) by Sex (Actual – Prediction) for High and Low Academic Composite Groups and for White Students Whose Best Language Is English *23*

17. Average Over- (–) and Underpredictions (+) of Course Grades for Females by Type of Course *24*

18. Characteristics of More Selective and Less Selective Colleges and of Large and Small Colleges by Sex *25*

19. Summary of Over- (–) and Underpredictions (+) of Course Grades for Females, and of FGPA, Average Grade Mean Residual, and SAT Mean, by Sex and Selectivity of College *26*

20. Student Characteristics by English as Best or Not Best Language *27*

21. Course Categories with the Largest Over- (–) and Underpredictions (+) Using SAT Scores and HSR for Students Whose Best Language Is Not English *28*

22. Characteristics of More Selective and Less Selective Colleges and of Large and Small Colleges, by English as Best or Not Best Language *29*

23. Student Characteristics by Ethnic Group *30*

24. Course Categories with Over-(–) and Underpredictions (+) of Course Grade Using HSGPA and SAT Scores, by Ethnic Group *32*

25. Characteristics of More Selective and Less Selective Colleges by Ethnic Group *34*

Abstract

Part 1 of this study investigated possible causes of the observed decline in correlations between SAT scores and freshman grade-point average (FGPA). The results were described in Chapter 12, "Implications of Using Freshman GPA as the Criterion for the Predictive Validity of the SAT," and were the basis for much of Chapters 2 and 3 of the monograph *Predicting College Grades: An Analysis of Institutional Trends Over Two Decades* (Willingham, Lewis, Morgan, and Ramist 1990). Working with a data base of 38 colleges, the study found that the comparability of course grades received by entering freshmen declined in the 1980s. Three new measures of grade comparability—variety of courses taken, variation in average student aptitude among courses, and appropriateness of average course grade in relation to student aptitude level—proved to be excellent indicators of both the level of and the change in SAT validity for predicting FGPA among the 38 colleges. Using course grade as the criterion instead of FGPA reduced the decline in both SAT and high school GPA (HSGPA) validity for predicting course grades by 40 percent. Contrary to the assumption that high school record (HSR) is a better predictor than the SAT, compared with HSR the SAT had higher or equal average validities for predicting course grade in almost all categories of courses. (Each course was placed into one of 37 categories based on subject, skills required, and level.)

Part 2 of this project examines course selection, grading patterns, grade comparability, SAT predictive effectiveness, and average over- and underpredictions in each type of course for groups defined by an academic composite index, sex, English as best or not best language, and ethnic group. SAT predictive effectiveness is determined with and without HSR on the basis of correlations that are corrected for restriction of range. Over- and underpredictions are determined by residuals from predictions. All results are analyzed by college selectivity level and size.

On average, males took more rigorously graded courses and females obtained a higher FGPA: two-thirds of the .09 difference by sex in FGPA related to course selection. Predictions of course grades based on the SAT were better for females, on average, than for males, and the SAT added more incremental information over HSR for females. Underprediction of FGPA for females, using the SAT and HSR, averaged .06. Underprediction of course grade for females, using the SAT and HSR, averaged .03, but was reduced to .02 using the Test of Standard Written English (TSWE) as an additional predictor, and was eliminated entirely at more selective colleges.

Although on average the SAT predicted FGPA and course grades better for students whose best language was English, it added more incremental information over HSR for students whose best language was not English. Asian American students took, on average, very strictly graded courses, but obtained a high average FGPA. The SAT predicted FGPA and course grades better for them than for any other ethnic group. On average, the SAT added more incremental information over HSR in predicting FGPA and course grades for black students than for any other ethnic group. Course grades were the least comparable for Hispanic and black students. They were so lacking in comparability for Hispanic students that, on average, there was better prediction of one course grade, as long as the course was identified, than of FGPA, even though the latter was typically based on eight or nine courses (for black students, course grade and FGPA predictions were equally good).

On average, substantial improvement in the prediction of FGPA could be obtained by using, as an additional predictor, the average grading difficulty of courses selected by students. The improvement was greatest at less selective colleges and for students of lower academic levels.

The highest average correlations in predicting FGPA were obtained by predicting each of a student's course grades separately and averaging the predictions to obtain a predicted FGPA. After correcting for predictor restriction of range and criterion unreliability, the average correlations were .64 for the SAT, .67 for HSR, and .75 for the multiple correlation of the SAT and HSR. These correlations may be the best estimates ever made of the effectiveness of the SAT and HSR for predicting FGPA, because they were based on a large cross section of colleges and on comparable grades for all courses taken by a student, and because the effects of both predictor restriction of range and criterion unreliability were removed.

Part 1: Course Grade Study

Investigation

Ramist, Lewis, and McCamley (1990) analyzed a data base of course grades to gain a greater understanding of the FGPA criterion. The 38 participating colleges (identified in Appendix A) varied widely in terms of selectivity, size, and control. The colleges supplied student identifications, courses taken and grades received, and HSGPA or HSR (27 of the 38 colleges supplied a mea-

sure of high school record) for enrolled freshmen in 1982 and 1985. Matching the student identifications against the files of the Admissions Testing Program provided SAT scores and Student Descriptive Questionnaire (SDQ) responses, including HSGPA, whether English was the student's best language, and ethnic group.

All courses taken by freshmen were assigned one of 37 categories based on subject, skills required, and level. Among the categories were five for mathematics (based on level), nine for English (based on level as well as whether the emphasis was on reading/literature, writing/composition, or both), three for biological sciences and three for physical sciences (based on level, for majors or nonmajors, and laboratory or nonlaboratory), two for foreign languages (based on level), and two for art/music/theater (studio and nonstudio). The other categories were: history, social sciences/humanities, economics, business/communications, computers, health/nursing, education, physical education, military science, home economics, architecture/environmental design, technical/vocational, and other. The full list is provided in Appendix B. The remainder of this section is a summary of the results of Part 1 of the course grade study.

Course Selection and Grading

Grade Comparability

A good criterion for predictive validity studies should provide comparable measurements, using approximately the same standard of performance, from student to student. Three concepts and related measures were associated with grade comparability.

If all students took the same courses, then grades probably would be as comparable as they could be—although there would still be differences among instructors. But variation in course taking influences comparability. Counting from higher to lower volume, the number of courses that account for half of all the credits taken by freshmen provided a measure of grade comparability. A high number showed course variety; a low number showed course concentration. The average across all the colleges was 16 courses, but one college averaged as few as 5 courses. Colleges with large declines in SAT predictive validity from 1982 to 1985 tended to show increases in the average number of courses.

Variation in course taking may not necessarily have an adverse effect on grade comparability if average student aptitude levels are about the same from course to course, but it may have a large effect if students with greater aptitude take certain courses and students with lesser aptitude take other courses. To measure variation in student aptitude levels among courses, the standard deviation of course SAT means was used. A high standard deviation shows great variation among courses; a low standard deviation shows similarity in aptitude among courses. Colleges with high SAT predictive validity tended to have lower standard deviations.

Even variation in student aptitude among courses may not have an adverse effect on grade comparability if grades correspond with aptitude levels. If average course grades reflect student aptitude levels, so that courses with higher student aptitude levels have higher grades and courses with lower aptitude levels have lower grades, there would be little adverse effect on grade comparability. The measure used, the correlation between course grade mean and course SAT mean, averaged only .10. Colleges with large declines in SAT predictive validity from 1982 to 1985 tended to show substantial declines in the correlation of course SAT mean and course grade mean, approaching .00.

Grading Difficulty

For each course, the *grade mean residual* is defined as the difference between the average course grade and the predicted GPA of the students in the course based on their SAT scores and HSR. A positive grade mean residual indicates higher grades than would be expected given student academic credentials (lenient grading). A negative grade mean residual indicates lower grades than would be expected from student academic credentials (strict grading).

Students with high SAT scores (compared to other students at the college) tended to select more strictly graded courses, which, on average, were science or quantitative and reduced their FGPAs. Similarly, students with low SAT scores (compared to other students at the college) tended to select more leniently graded courses, which on average were nonquantitative and increased their FGPAs.[1] These patterns became clearer during the period from 1982 to 1985, especially for less selective colleges. As colleges increasingly allowed and encouraged students to take courses most appropriate to their aptitude levels, course variety increased and grades became less comparable.[2]

[1]The division of academic departments into those that grade strictly, usually those with a scientific and quantitative focus and those that grade more leniently, usually those with a nonquantitative focus, has long been recognized. See Goldman, Schmidt, Hewitt, and Fisher (1974); Goldman and Hewitt (1975); Goldman and Widawski (1976); Goldman and Slaughter (1976); Ramist (1984); Willingham (1985); Milton, Pollio, and Eison (1986); Strenta and Elliott (1987); Elliott and Strenta (1988); and Sabot and Wakeman-Linn (1991).

[2]Reacting to this lack of comparability of grades, Milton, Pollio, and Eison (1986), p. 218, go so far as to suggest abolishing the computation of GPA because it is becoming meaningless.

FGPA as the Criterion

By Academic Level

Dividing the freshman class at each college into equal thirds based on a composite of SAT scores and HSR, over all colleges both the SAT and HSR predicted best for the top third. For the bottom third, the SAT predicted slightly better than HSR and the SAT had its highest incremental correlation over HSR. The SAT is most useful in predicting FGPA for the bottom of the class, where the most difficult selection decisions are made. This is especially true for more selective colleges: in predicting FGPA, the SAT incremental correlation over HSR averaged .06 to .07 for all students at all colleges, .09 for the lowest third of the class at all colleges, and .12 for the lowest third of the class at more selective colleges.

Predicting Validity for FGPA from Grade Comparability

The three measures of grade comparability—the number of courses that account for half of all the credits taken by freshmen, the standard deviation of course SAT means, and the correlation between course grade mean and course SAT mean—were used to predict SAT validity for FGPA (the multiple correlation between the two SAT scores and FGPA) and HSR validity for FGPA (the correlation between HSR and FGPA). The average multiple correlation for 1982 and 1985 was close to .70 in predicting SAT validity for FGPA and close to .60 in predicting HSR validity for FGPA. Essentially, the validity of both the SAT and HSR for predicting FGPA depended on how comparable the grades were: high comparability leading to high validity for FGPA and low comparability leading to low validity for FGPA.

Using changes in the three measures from 1982 to 1985 to predict changes in the levels of validity for FGPA, the prediction was again very good for the SAT and moderately good for HSR. The multiple correlation was about .60 for predicting the change in SAT validity for FGPA and about .30 for predicting the change in HSR validity for FGPA. Especially considering the difficulty of predicting change, the predictability of the change in validity for FGPA was high. The change in the SAT and (to a moderate extent) HSR validity for predicting FGPA depended on the change in the comparability of grades: increased comparability leading to higher validity for FGPA and decreased comparability leading to lower validity for FGPA. The typical pattern of change was for grade comparability and validity for FGPA to decline at less selective colleges, where course variety increased (especially in mathematics courses) and the correlation of course SAT mean and course grade mean declined.

Course Grade as the Criterion

Compared to FGPA

The College Board has always recommended FGPA as the criterion for determining the predictive validity of the SAT in admission.[3] But there may be another possible criterion: the single course grade.

To evaluate this new criterion, separate correlations were computed for predicting course grade for each of 4,680 courses (with at least 5 freshmen in each of the years) and corrected for restriction of range. For each college, the course correlations for each year were averaged, weighted by the number of freshmen in the courses, and then the college means were averaged. The average correlation of one course grade with the SAT was an unexpectedly high .49. Although the correlations for all the courses were averaged, the average correlation represented the average predictability of only one course grade, with all the subjectivity that may be inherent in one grade. When FGPA was computed, with a mean course load of 8.6, and used as the criterion, the increase in average SAT predictive validity was only .05, increasing from .49 to .54. If grades were comparable, based on the Spearman-Brown formula, it was shown that 8.6 grades, each having a correlation with the SAT of .49, would increase the predictive validity of the SAT for FGPA to .75, not .54, increasing the predictive validity of the SAT by +.26, instead of only +.05, from predicting course grade to predicting FGPA.[4] The lack of comparability of grades substantially offset the benefits of increased sample size from one to eight or nine grades, eliminating 80 percent of the expected increase in predictive validity.[5] The effect of lack of comparability of grades was less on HSR than on the SAT, eliminating only a little more than half of the expected increase in predictive validity.

SAT Predictive Validity by Type of Course

The courses with the highest average correlations between the SAT and course grade were science or quantitative—especially biological sciences (after correction for restriction of range, nonmajors .61, advanced .60, and lab/majors .58), physical sciences (nonmajors .54, advanced .61, and lab/majors .53), economics (.59), advanced mathematics (.57), and calculus (.55)—with negative average grade residuals indicating strict grading. The

[3]See College Board (1988), p. 9.

[4]Willingham, Lewis, Morgan, and Ramist (1990), p. 329.

[5]Goldman and Widawski (1976) and Elliott and Strenta (1988) recognized that the expected improvement in validity from increased sample size by adding courses is diminished by differential grading standards.

courses with low average correlations between the SAT and course grade were all nonquantitative and included physical education (.21), remedial English (.25), technical/vocational (.27), studio art/music/theater (.32), remedial reading/literature (.33), and military science (.34), with substantial positive average grade residuals, from +.28 to +.78, indicating lenient grading.

Part 2: Extending the Analysis

To Additional Colleges

In addition to the 38 colleges identified in Appendix A that supplied data on the 1982 and 1985 classes for Part 1 of this study, Part 2 includes 7 additional colleges, also identified in Appendix A, that supplied data for the 1985 class only. The findings in Part 2 are based on 1985 data for all 45 colleges and therefore may differ somewhat from those in Part 1.

By Student Group

Part 2 of this study extends the analyses of Part 1 to document differences among student groups in course selection, in the predictive effectiveness of the SAT and HSR for both FGPA and course grade, and in the comparison of actual and predicted grades. To compare groups on predictive effectiveness, correlations are corrected for restriction of range. To compare actual and predicted grades, average over- or underpredictions are provided for each student group using combinations of test scores and/or HSGPA in prediction equations derived for all students. Student groups were defined by academic level, sex, whether English was the student's best language, and ethnic group.

Student groupings by academic level were established for each college. An academic composite index of SAT scores and HSGPA was tabulated for each enrolled freshman. The weights used were averages of the weights for all prediction equations produced in validity studies by the College Board's Validity Study Service using SAT scores and HSGPA on entering classes from 1981 to 1985: .00118 for the verbal score, .00100 for the mathematical score, and .55498 for HSGPA. For each college, students were grouped into equal thirds based on the academic composite index. Any student without at least one A to F course grade was eliminated. Separate analyses were performed for high, middle, and low academic composite groups, first for each college and then averaged across colleges. The averages across colleges are presented.

Student groupings by sex, whether English was the best language, and ethnic background are based on student responses to the SDQ. Chart 1 shows the SDQ responses to the ethnic question and how the ethnic groups are described in this report.

CHART 1

SDQ Responses	*How Described in This Report*
American Indian or Alaskan Native	American Indian
Black or Afro-American or Negro	Black
Mexican-American or Chicano	Hispanic
Oriental or Asian-American or Pacific Islander	Asian-American
Puerto Rican	Hispanic
White or Caucasian	White
Other	Not included

CHART 2

Academic Composite Groups[6]		*Sex*		*English Best Language*		*Ethnic Group*[7]	
High	16,010	Male	22,412	Yes	44,699	American Indian	184
Medium	15,340	Female	23,967	No	1,156	Asian American	3,848
Low	15,029	TOTAL	46,379	No Response	524	Black	2,475
TOTAL	46,379			TOTAL	46,379	Hispanic	1,599
						White	36,743
						Other/No Response	1,530
						TOTAL	46,379

[6]There were fewer students in the lower third of the academic composite groups than in the upper third because any student without at least one A to F course grade was eliminated.

[7]Compared to the most comprehensive recent analyses of SAT predictive validity for each of the ethnic groups, this study includes data for comparable or larger numbers of students:

- American Indian: No known study; this study includes data for 184 American Indian students at 34 of the 45 colleges.
- Asian American: Sue and Abe (1988) include data for 4,113 Asian American students at 8 campuses of the University of California; this study includes data for 3,848 Asian American students at 43 of the 45 colleges.
- Black: Nettles, Thoeny, and Gosman (1986) include data for 664 black students at 30 colleges; this study includes data for 2,475 black students at 45 colleges.
- Hispanic: Pennock-Román (1990) includes data for 1,447 Hispanic students at 6 colleges; this study includes data for 1,599 Hispanic students at 44 of the 45 colleges.

Note that Mexican-American or Chicano and Puerto Rican respondents are grouped as Hispanic. "Other" respondents are not include in the ethnic analysis.

Records for 46,379 students were analyzed. The numbers of students in the various groups are shown in Chart 2 (the numbers of students at each of the 45 colleges whose best language was not English and in each of the ethnic groups are included in Appendix A).

Course Selection and Grading

Course selection for each student group is described by several variables and analyses that were included in Part 1 and by some that were introduced in Part 2. The course-selection variables included in both Parts 1 and 2[8] are:

- the average number of courses taken by a student;
- the percentage of credits in advanced courses;
- the percentage of credits in remedial courses;
- the percentage of students taking a course in each category;
- the number of courses that would account for half of all the credits accumulated, indicating variety of courses taken;
- the standard deviation of course SAT means, indicating variation in student aptitude levels among courses;
- the correlation between course grade mean and course SAT mean, indicating appropriateness of the average course grade with respect to aptitude level; and
- the grade mean residual for each course, indicating the difference between the average course grade and the predicted FGPA of the students in the course based on their SAT scores and HSR.

Several course selection variables and analyses for each student group were introduced in Part 2 of the study. In addition to the three variables associated with grade comparability, grade comparability is documented by the difference in correlations for predicting FGPA and course grade. A small difference indicates less comparable course grades (the natural increase in the correlation due to the increased number of course grades in FGPA is not taking place); a large difference indicates more comparable course grades.

For each course, course grade was predicted from the SAT-Verbal (SAT-V) score, the SAT-Mathematical (SAT-M) score, and HSGPA for all students in the course, and the predictive weights were standardized. After converting negative weights to zero, the proportion of the total of the standardized weights associated with each predictor indicated the proportional contribution of the predictor. Averaging proportional contributions over all courses selected by the students in a group, weighted by the number of students in the group in the course, provided an indication of the relevance of each of the variables in predicting FGPA in courses selected by the group.

For each student, the grade mean residuals for all the courses selected were averaged. This variable describes the average grading difficulty of the courses selected. The grade mean residuals were averaged for each student group and weighted by the number of students in the group in the course to determine how grading difficulty was influenced by the courses selected by the group.[9]

To indicate a student's competitive advantage or disadvantage, his or her SAT scores, TSWE score, and HSGPA were compared to the average SAT scores, TSWE score, and HSGPA for all students in each course selected by the student, and then these differences were averaged over all courses selected by the student. Weighted averaging over all students in a student group provided the average competitive advantage or disadvantage associated with the courses selected by the group.

Predictive Effectiveness

The predictive effectiveness of the SAT and HSR was evaluated for each student group in terms of the prediction of FGPA for each college (averaged over all colleges and weighted by the number of students in the group at the college) and course grade (averaged over all courses and weighted by the number of students in the group in the course). In order to ensure that all students had comparable HSR, HSGPA from the SDQ was used. Correlation coefficients were tabulated.[10] To make them comparable for each student group, college, type of college, and type of course, they were corrected for restriction of range to the full national SAT-taking group using the Pearson-Lawley multivariate correction.[11] To eliminate the artificial reduction of the correlations due to criterion

[8]In Part 1, because the emphasis was on all students at a college, variables were tabulated for each of the 38 colleges, and then averaged, weighting each college equally. In Part 2, the results may differ not only because there are 45 colleges, but also because, due to the emphasis on student groups that in many cases are disproportionate in number across colleges, the averages for each group (including all students) are weighted by the number of students in the group at each college.

[9]To underscore relative group differences, grade mean residuals were centered at .00 for all students.

[10]For FGPA at every college and separately for every course with at least seven freshmen in 1985.

[11]See Gulliksen (1950), pp. 165–66. The estimates of standard deviations for all test takers were 109 for the SAT-Verbal score, 120 for the SAT-Mathematical score, and .61 for HSGPA. The estimates of correlations among the predictors for all test takers were .67 for the SAT-Verbal and SAT-Mathematical scores, .45 for the SAT-Verbal score and HSGPA, and .50 for the SAT-Mathematical score and HSGPA. Estimates were based on SAT scores and on HSGPA from the SDQ for all SAT takers during the 1982–1985 period.

unreliability, they were also corrected for criterion unreliability. To indicate the SAT's utility, the SAT increment to the correlation, over HSGPA, was tabulated.

Correlation coefficients for predicting course grade were averaged over all 7,786 courses with at least 7 freshmen in 1985 and also for each type of course. Types of courses for which the verbal score was most important and types for which the mathematical score was most important were identified.

The TSWE score and the average grade mean residual of the courses selected by the student were used as additional variables for predicting FGPA. They were used as single predictors and combined with SAT scores and HSGPA. TSWE scores were used to predict course grades, overall and for each type of course. Only uncorrected correlations involving these predictors were tabulated.

As a means of adjusting for the lack of criterion comparability caused by course selection, individual predictions of grades in each course taken by a student were averaged and correlated with the student's FGPA. These correlations were not only based on predictions and actual grades in the same courses, but also allowed for the benefits of increased course sample size from one course to eight or nine courses. They were tabulated for course grade predictions based on SAT scores and HSGPA, separately and combined, and both uncorrected correlations and correlations corrected for predictor restriction of range[12] and criterion unreliability were reported.

Lack of criterion reliability reduces the correlation between a predictor and the criterion. In order to eliminate the effect of criterion unreliability on correlations for the criterion of one course grade, 44 course sections were identified from one university for which the first part of the course was in one term and the second part in another term. First- and second-term grades for each course were correlated to estimate the reliability of a single course grade, and these correlations were corrected for restriction of range.[13] In order to eliminate the effect of criterion unreliability on correlations for the criterion of FGPA, at each college separate averages for each of two randomly created sets of half of the course grades of each student were correlated, the correlations were adjusted for the full FGPA, and then the correlations were corrected for restriction of range.

Over- and Underpredictions

For each student group, grades and predictions were compared for both FGPA and course grade (overall and for each type of course). If, on average, grades exceeded predictions, then the predictor(s) was (were) underpredicting for that group; if, on average, predictions exceeded grades, then the predictor(s) was (were) overpredicting for that group.

Over- and underpredictions for both FGPA and course grade were evaluated for the single predictors of HSGPA, SAT-Verbal score, and SAT-Mathematical score, and multiple predictor combinations of SAT scores, HSGPA, and TSWE score. Over- and underpredictions for FGPA were also determined after combining the average grade mean residual of the courses selected with HSGPA and SAT scores.

Differences Among Colleges

Analyses of course selection, predictive effectiveness of the SAT and HSGPA for FGPA and course grade, and over- and underpredictions, by type of course and by student group, were carried out for all 45 colleges and then for groups of 15 colleges based on selectivity and size. Three groups of 15 colleges were defined by their SAT-V+M means: the most selective third with means of 1121 or higher, the middle third with means of 986 to 1120, and the least selective third with means below 986. Three other groups of 15 colleges were defined by the number of freshmen: the largest third with 900 or more entering freshmen who took the SAT, the middle third with 485 to 899, and the smallest third with fewer than 485.

Statistical Precision of Results

The analyses carried out in this study were exploratory in nature. As a consequence, statistical tests of hypotheses were not carried out, and standard errors of statistics were not estimated. Instead, results should be interpreted as descriptive of characteristics observed in a data base of 46,379 students in the 1985 freshman class, taking a total of 7,786 courses offered at 45 colleges and receiving 395,106 course grades.

It may be expected that most of the results reported here are similar to what would be found for another, similar group of students, courses, and colleges. This expectation is stronger with respect to the 23,967 female students in the study than for the 184 American Indian students. Nonetheless, the findings for American Indian students were not ignored. The analysis of over- and underpredictions for course grades of American Indians was based on 1,509 course grades. All other groups had substantially larger numbers of course grades.

[12-13]These correlations were corrected using an extension of the Pearson-Lawley multivariate correction to the problem of implicit selection (see Gulliksen, 1950, p. 165).

Variations by Academic Composite Group

Student Characteristics

Table 1 shows the student composition of the high, middle, and low academic composite groups. The high composite group has almost equal numbers of males and females; the other composite groups have more females. Students whose best language was not English are more strongly represented in the low composite group, as are black, Hispanic, and American Indian students. Asian American, white, and students whose best language was English are more strongly represented in the high composite group.

Course Selection and Grading

Table 2 displays the courses selected by each of the student groups. The average student took 8.5 courses for credit, 6 percent advanced and 2 percent remedial (remedial courses without credit were excluded). As expected, students in the low academic composite group averaged fewer courses (8.1 compared to 8.8 for the high composite group), fewer advanced (4 percent compared to 7 percent), and more remedial (3 percent compared to 1 percent).

TABLE 1

Student Composition of Each Academic Composite Group

All Students			*Academic Composite Group*		
N	*Percent*		*High*	*Middle*	*Low*
		SEX			
22,412	48.3	Males	49.9%	47.2%	47.8%
23,967	51.7	Females	50.1%	52.8%	52.2%
46,379		TOTAL			
		ENGLISH BEST LANGUAGE			
44,699	97.5	Yes	98.3%	97.9%	97.1%
1,156	2.5	No	1.7%	2.1%	2.9%
45,855		TOTAL			
		ETHNIC GROUP			
184	0.4	American Indian	0.2%	0.4%	0.7%
3,848	8.6	Asian American	9.5%	8.6%	7.5%
2,475	5.5	Black	1.5%	3.8%	11.7%
1,599	3.6	Hispanic	2.0%	3.2%	5.6%
36,743	81.9	White	86.7%	84.0%	74.5%
44,849		TOTAL			

Among course categories, the highest percentage of course grades were in the broad category of social sciences/humanities (19 percent). Other courses frequently taken were physical sciences/engineering (14 percent: 8 percent lab or for majors and 6 percent nonlab and nonmajor); mathematics (13 percent: 8 percent calculus, 2 percent precalculus, 1 percent advanced math, 1 percent regular math, and 1 percent remedial math); English (12 percent: 6 percent emphasizing writing, 2 percent emphasizing reading/literature, and 4 percent both); foreign language (7 percent: 4 percent beginning and 3 percent beyond entry); art/music/theater (5 percent: 3 percent nonstudio and 2 percent studio); history (5 percent); biological sciences (4 percent: 2 percent lab or for majors and 2 percent nonlab and nonmajor); and economics (4 percent).

Students in the high academic composite group selected mathematics courses at the calculus level or higher at twice the rate (12 percent: 2 percent advanced math and 10 percent calculus) of students in the low composite group (6 percent: almost all calculus) and selected physical sciences/engineering courses with a lab or for majors at more than three times the rate (13 percent compared to 4 percent for students in the low composite group). Students in the low composite group took more mathematics courses below the calculus level (6 percent compared to 2 percent for students in the high composite group), more regular writing courses (7 percent compared to 5 percent), more regular English courses (5 percent compared to 4 percent), and more social sciences/humanities courses (21 percent compared to 17 percent).

Table 3 shows course selection and grading characteristics for each of the student groups. On average, in predicting course grades for all students, 54 percent of the predictive weight was on SAT scores, with more on the mathematical score (28 percent[14]) than the verbal (25 percent[15]), and 46 percent[16] on HSGPA. For courses selected by students in each of the academic composite groups, the SAT and HSGPA contributions were about the same as for all courses, but students in the high composite group selected courses averaging more predictive weight on the mathematical score (29 percent compared to 27 percent for courses selected by students in the low composite group) and less on the verbal score (23 percent compared to 26 percent).

Course grades were much less comparable for the low than for the high academic composite group. The average difference in the corrected correlations predicting course grade and FGPA (using SAT scores and HSGPA) was only .02 for students in the low composite group compared to .12 and .11 for students in the high

[14–16] Percentages may not add to 100 percent because of rounding.

TABLE 2

Courses Selected by Student Group

All Students	Academic Composite Group			Sex			English Best Language		Ethnic Group				
	High	Medium	Low	Male	Female		Yes	No	American Indian	Asian American	Black	Hispanic	White
8.5	8.8	8.5	8.1	8.4	8.6	Average number of courses	8.5	8.1	8.2	7.9	8.0	7.8	8.6
6%	7%	5%	4%	5%	6%	% of courses advanced	6%	7%	3%	6%	4%	5%	6%
2%	1%	1%	3%	2%	2%	% of courses remedial	1%	3%	2%	1%	3%	1%	1%
						COURSE CATEGORIES							
1%	2%	1%	–	2%	1%	Advanced mathematics	1%	3%	–	3%	–	1%	1%
8%	10%	9%	6%	10%	7%	Calculus	8%	12%	6%	13%	6%	9%	8%
2%	1%	2%	3%	2%	2%	Precalculus	2%	2%	3%	2%	3%	3%	2%
1%	1%	2%	2%	2%	2%	Regular mathematics	2%	2%	3%	1%	2%	2%	2%
1%	–	1%	1%	1%	1%	Remedial mathematics	1%	1%	1%	–	2%	–	1%
–	–	–	–	–	–	Advanced English	–	–	–	–	–	–	–
4%	4%	5%	5%	4%	5%	Regular English	4%	4%	6%	3%	4%	5%	5%
–	–	–	–	–	–	Remedial English	–	1%	–	–	–	–	–
–	–	–	–	–	–	Advanced reading/lit.	–	–	–	–	–	–	–
2%	2%	2%	2%	2%	3%	Regular reading/lit.	2%	1%	1%	2%	2%	2%	2%
–	–	–	–	–	–	Remedial reading/lit.	–	–	–	–	–	–	–
–	–	–	–	–	–	Advanced writing	–	–	–	–	–	–	–
6%	5%	6%	7%	6%	6%	Regular writing	6%	5%	6%	5%	8%	5%	6%
–	–	–	1%	–	–	Remedial writing	–	1%	1%	1%	1%	1%	–
–	–	–	–	–	–	Advanced biological sciences	–	–	–	–	–	–	–
2%	3%	3%	2%	2%	3%	Lab or major biological sciences	2%	3%	3%	3%	3%	4%	2%
2%	2%	2%	3%	2%	3%	Nonlab and nonmajor biological sciences	2%	2%	2%	2%	3%	3%	2%
–	1%	–	–	1%	–	Adv. physical sciences/engineering	–	–	–	1%	–	–	–
8%	12%	7%	4%	10%	5%	Lab or major physical sciences/engineering	8%	16%	5%	15%	6%	9%	7%
6%	6%	6%	5%	7%	5%	Nonlab and nonmajor phys. sci./engr.	6%	6%	7%	7%	5%	7%	6%
3%	3%	3%	3%	2%	4%	For. lang.–beyond entry	3%	3%	2%	2%	3%	3%	3%
4%	4%	4%	5%	3%	5%	Beg. foreign language	4%	4%	6%	4%	5%	4%	4%
5%	5%	5%	5%	5%	5%	History	5%	3%	5%	3%	4%	4%	5%
4%	4%	5%	4%	5%	4%	Economics	4%	4%	4%	4%	3%	4%	4%
19%	17%	19%	21%	17%	21%	Social sci./humanities	19%	13%	19%	15%	22%	19%	19%
3%	2%	3%	3%	3%	3%	Bus./communications	3%	1%	2%	1%	2%	2%	3%
2%	2%	2%	3%	2%	3%	Art/music/theater – studio	2%	2%	2%	2%	3%	2%	2%
3%	3%	3%	4%	3%	4%	Art/music/theater – nonstudio	3%	2%	5%	3%	3%	3%	3%
2%	2%	2%	2%	3%	2%	Computer science	2%	4%	2%	3%	2%	2%	2%
1%	–	1%	1%	–	1%	Health/nursing	1%	–	1%	–	1%	–	1%
1%	–	1%	1%	–	1%	Education	1%	1%	–	–	1%	1%	1%
3%	2%	3%	3%	2%	3%	Physical education	3%	2%	4%	2%	4%	2%	3%
1%	1%	1%	1%	1%	1%	Other	1%	1%	1%	1%	1%	1%	1%
1%	1%	1%	–	1%	–	Military science	1%	–	1%	–	1%	–	1%
–	–	–	–	–	1%	Home economics	–	–	–	–	–	1%	–
–	–	–	–	–	–	Technical/vocational	–	–	–	–	–	–	–
–	–	–	–	1%	–	Architecture/env. des.	–	–	–	1%	–	1%	–

TABLE 3

Course Selection and Grading Characteristics by Student Group

All Students	Academic Composite Group			Sex			English Best Language		Ethnic Group				
	High	Medium	Low	Male	Female		Yes	No	American Indian	Asian American	Black	Hispanic	White
25	24	23	23	21	27	Number of courses accounting for half of all credits	25	18	15	24	21	23	24
67	84	43	55	69	67	*SD* of course SAT means	67	96	101	84	128	121	97
.12	.19	−.07	−.05	.17	.14	Correlation between course grade mean and SAT mean	.12	.19	.10	.15	.22	.17	.30
2.63	3.01	2.60	2.26	2.58	2.67	Mean FGPA	2.63	2.67	2.21	2.80	2.14	2.37	2.66
.00	−.03	.00	+.04	−.03	+.03	Average grade mean residual	.00	−.03	+.02	−.07	+.06	+.01	+.01
						PROPORTIONAL CONTRIBUTIONS TO PREDICT GRADES IN STUDENT-SELECTED COURSES:							
25%	23%	25%	26%	23%	26%	SAT-V	25%	20%	25%	20%	26%	22%	25%
28%	29%	28%	27%	30%	27%	SAT-M	28%	32%	28%	32%	28%	30%	28%
46%	47%	46%	46%	47%	46%	HSGPA	46%	48%	47%	47%	45%	47%	46%
						CORRELATIONS USING SAT AND HSGPA:*							
.63	.76	.65	.49	.61	.66	For FGPA	.63	.57	.59	.64	.52	.54	.63
.57	.64	.54	.47	.56	.60	For course grade	.58	.56	.57	.62	.52	.55	.56
.06	.12	.11	.02	.05	.06	Difference	.05	.01	.02	.02	.00	-.01	.07
						CORRELATIONS USING SAT:*							
.53	.59	.53	.43	.52	.58	For FGPA	.53	.49	.46	.54	.46	.40	.52
.49	.51	.46	.42	.48	.52	For course grade	.49	.46	.39	.51	.46	.45	.48
.04	.08	.07	.01	.04	.06	Difference	.04	.03	.07	.03	.00	−.05	.04
						SAT-VERBAL MEAN:							
505	561	497	449	512	499	Group	508	391	462	484	436	462	513
505	516	503	494	508	502	Courses selected	505	492	492	513	488	500	505
0	+45	−6	-45	+4	-3	Difference	+3	−101	−30	−29	−52	−38	+8
						SAT-MATHEMATICAL MEAN:							
559	617	553	499	586	535	Group	559	568	511	595	466	516	564
559	575	556	544	570	549	Courses selected	559	575	543	590	538	559	558
0	+42	−3	−45	+16	−14	Difference	0	−7	−32	+5	−72	−43	+6
						TSWE MEAN:							
49	53	49	45	48	49	Group	49	39	46	46	44	46	50
49	50	49	48	49	49	Courses selected	49	48	48	49	48	49	49
0	+3	0	−3	−1	0	Difference	0	−9	−2	−3	−4	−3	−1
						HSGPA MEAN:							
3.41	3.79	3.42	2.95	3.37	3.44	Group	3.41	3.45	3.26	3.58	3.18	3.43	3.40
3.41	3.47	3.40	3.35	3.42	3.40	Courses selected	3.41	3.45	3.39	3.54	3.36	3.48	3.39
0	+.32	+.02	−.40	−.05	+.04	Difference	.00	.00	−.13	+.04	−.18	−.05	+.01
						PREDICTED GPA MEAN:							
2.63	2.98	2.62	2.24	2.64	2.62	Group	2.64	2.54	2.45	2.76	2.28	2.49	2.65
2.63	2.70	2.62	2.57	2.65	2.62	Courses selected	2.63	2.67	2.60	2.76	2.54	2.62	2.62
0	+.28	.00	−.33	−.01	.00	Difference	+.01	−.13	−.15	.00	−.26	−.13	+.03
						COURSE GRADE MEAN:							
2.70	3.05	2.66	2.34	2.66	2.74	Group	2.70	2.75	2.29	2.85	2.24	2.44	2.73
2.70	2.73	2.69	2.68	2.69	2.71	Courses selected	2.70	2.71	2.69	2.76	2.67	2.69	2.70
.00	+.32	−.03	−.34	−.03	+.03	Difference	.00	+.04	−.40	+.09	−.43	−.25	+.03

*Correlations corrected for restriction of range

and medium composite groups. The negative correlation (–.05) between course grade mean and course SAT mean for students in the low composite group is consistent with lack of grade comparability. While not high, the correlation between course grade and course SAT mean, a moderately positive +.19 for students in the high composite group, is consistent with greater grade comparability.

The courses selected by students in the high academic composite group were more strictly graded (average grade mean residual of –.03) than the courses selected by students in the low composite group (+.04). In other words, compared with the high composite group, course selection gave students in the low composite group a "bonus" on FGPA of .07. Despite this bonus, they still averaged .75 lower on FGPA (2.26 compared to 3.01).

As expected, compared to students in the low composite group, students in the high composite group had higher average test scores (by about 1.3 standard deviations: SAT-Verbal by 112 points, SAT-Mathematical by 118 points, and TSWE by 8 points), HSGPA (by 2 standard deviations: .84), and predicted FGPA (by .74). They also selected courses in which students had higher average test scores (SAT-Verbal by 22 points, SAT-Mathematical by 31 points, and TSWE by 2 points), HSGPA (by .12), and predicted FGPA (by .13) than those in the courses selected by students in the low composite group, by about .3 of a standard deviation. Despite this pattern of course selection, students in the high composite group were at a substantial competitive advantage compared to students in the courses they selected (SAT-Verbal mean by 44 points, .5 standard deviation; SAT-Mathematical mean by 42 points, .5 standard deviation; TSWE mean by 3 points; HSGPA mean by .32, .8 standard deviation; and predicted FGPA by .28). Students in the low composite group were at a substantial competitive disadvantage compared with students in the courses they selected (SAT-Verbal mean by 44 points, .5 standard deviation; SAT-Mathematical mean by 45 points, .5 standard deviation; TSWE mean by 3 points; HSGPA mean by .40, 1.0 standard deviation; and predicted FGPA by .33).

Predictive Effectiveness

Table 4 presents correlations with FGPA, with course grade, and between average predicted course grade and FGPA for each of the student groups. To permit comparisons among the groups and to eliminate the effect of criterion unreliability, the FGPA correlations for SAT scores and HSGPA corrected for predictor restriction of range and criterion unreliability are shown in addition to the uncorrected correlations. Because of the great restriction of range in individual course grade correlations, they are presented only after correction. In addition to SAT scores and HSGPA, an uncorrected FGPA correlation for the TSWE score and also for the average grade mean residual (grading difficulty) of the courses selected by the student (abbreviated "Z") are shown. Increments in correlations are shown for SAT scores over HSGPA and for the average grade mean residual (Z increment) over SAT scores and HSGPA.

A correlation between a predictor and a criterion is proportional to the square root of the reliability of the criterion. To determine the reliability of the criterion of one course grade, the parallel forms method was used. A sample of 44 course sections from one university was identified for which the first part of the course was in one term and the second part was in another term. An average of 25 students per course section took both parts of the course. The correlation of the first- and second-term grades was $r = .59$, which is an estimate of the reliability of a course grade. An estimate of the reliability of a course grade corrected for restriction of range is .66.[17]

To determine the reliability of the criterion of FGPA, the split-halves method was used. At each college, separate averages for each of two randomly created sets of half of the course grades of each student were created. The average correlation of the two sets was .69. Adjusted for the full FGPA, an estimate of the average reliability of FGPA is .82. Corrected for restriction of range, the estimate becomes .87.

Comparing corrected correlations for students at different academic aptitude levels, the high composite group had higher correlations than the low composite group for both FGPA and course grade. Using SAT scores as predictors, high composite group correlations were .63 for FGPA and .63 for course grade, compared to .46 and .52 in the low composite group. But the differences between

[17]Corrected reliability

$$= (\sigma_{\hat{G}}^2+\sigma_r^2-\sigma_e^2)/(\sigma_{\hat{G}}+\sigma_r^2)$$
$$= R^2_{G.VMH}+[(\sigma_r^2-\sigma_e^2)/(\sigma_{\hat{G}}+\sigma_r^2)]$$
$$= R^2_{G.VMH}+[1-R^2_{G.VMH}]\,[(\rho_g-R^2_{g.vmh})/(1-R^2_{g.vmh})]$$
$$= .57^2+[1-.57^2]\,[(.59-.42^2)/(1-.42^2)]$$
$$= .66$$

Where $\sigma_{\hat{G}}^2$ = the variance of a predicted course grade, corrected for restriction of range

σ_r^2 = the residual variance

σ_e^2 = the error variance

ρ_g = reliability of a course grade, uncorrected for restriction of range = .59

$R_{g.vmh}$ = multiple correlation of predicting course grade from SAT scores and HSR, uncorrected for restriction of range = .42

$R_{G.VMH}$ = multiple correlation of predicting course grade from SAT scores and HSR, corrected for restriction of range = .57

TABLE 4

Predictive Effectiveness by Student Group

All Students	Academic Composite Group			Sex			English Best Language		Ethnic Group				
	High	Medium	Low	Male	Female		Yes	No	American Indian	Asian American	Black	Hispanic	White
						FGPA CORRELATIONS							
						UNCORRECTED:							
.30	.12	.04	.12	.26	.34	SAT-Verbal	.31	.17	.22	.24	.26	.24	.27
.31	.14	.05	.14	.31	.36	SAT-Mathematical	.31	.37	.25	.37	.24	.21	.26
.28	.15	.08	.15	.24	.32	TSWE	.30	.19	.23	.23	.28	.24	.26
.22	.31	.34	.33	.19	.24	Average grade mean residual (Z)	NA	NA	NA	NA	NA	NA	NA
.36	.18	.08	.17	.35	.41	SAT (V,M)	.36	.39	.34	.39	.30	.27	.32
.39	.19	.06	.11	.38	.40	HSGPA	.40	.34	.42	.37	.28	.35	.38
+.09	+.12	+.12	+.14	+.08	+.10	SAT Increment	+.08	+.14	+.13	+.11	+.11	+.08	+.07
.48	.31	.18	.25	.46	.50	VMH	.48	.48	.55	.48	.39	.43	.45
+.10	+.15	+.22	+.20	+.09	+.11	Z Increment	+.10	+.09	+.12	+.10	+.17	+.12	+.11
.58	.46	.40	.45	.55	.60	VMHZ	.58	.57	.67	.58	.56	.55	.56
						CORRECTED:*							
.50	.56	.50	.40	.48	.55	SAT-Verbal	.52	.40	.42	.47	.44	.39	.50
.53	.59	.53	.42	.53	.58	SAT-Mathematical	.53	.52	.36	.56	.44	.38	.52
.57	.63	.57	.46	.56	.62	SAT (V,M)	.57	.53	.49	.58	.49	.43	.56
.61	.76	.63	.44	.58	.61	HSGPA	.61	.50	.49	.60	.46	.53	.61
+.07	+.06	+.07	+.09	+.07	+.10	SAT Increment	+.07	+.11	+.14	+.09	+.10	+.05	+.07
.68	.82	.70	.53	.65	.71	VMH	.68	.61	.63	.69	.56	.58	.68
						COURSE GRADE CORRELATIONS							
						CORRECTED:*							
.50	.52	.48	.42	.48	.53	SAT-Verbal	.50	.43	.39	.49	.47	.44	.49
.54	.58	.52	.46	.53	.57	SAT-Mathematical	.54	.53	.32	.59	.48	.48	.53
.60	.63	.57	.52	.59	.64	SAT (V,M)	.60	.57	.48	.63	.57	.55	.59
.58	.68	.57	.44	.57	.59	HSGPA	.58	.54	.59	.63	.46	.55	.57
+.12	+.11	+.09	+.14	+.12	+.15	SAT Increment	+.13	+.15	+.11	+.13	+.18	+.13	+.12
.70	.79	.66	.58	.69	.74	VMH	.71	.69	.70	.76	.64	.68	.69
						CORRELATION OF PREDICTED AND ACTUAL FGPA							
						UNCORRECTED:							
						Predicted FGPA based on:							
.44	.33	.26	.35	.40	.48	SAT-Verbal	.45	.41	.38	.40	.44	.41	.42
.46	.36	.27	.36	.43	.50	SAT-Mathematical	.46	.47	.38	.46	.44	.37	.43
.50	.39	.29	.38	.46	.54	SAT (V,M)	.50	.50	.39	.49	.48	.41	.47
.51	.41	.27	.34	.49	.53	HSGPA	.51	.45	.51	.47	.43	.48	.51
+.09	+.08	+.14	+.15	+.08	+.10	SAT Increment	+.09	+.11	+.05	+.11	+.13	+.06	+.07
.60	.49	.41	.49	.57	.63	VMH	.60	.56	.56	.58	.56	.54	.58
						CORRECTED:*							
						Predicted FGPA based on:							
.60	.62	.60	.55	.57	.63	SAT-Verbal	.60	.56	.53	.56	.57	.52	.59
.62	.67	.61	.56	.60	.65	SAT-Mathematical	.62	.57	.50	.59	.58	.48	.61
.65	.70	.64	.59	.63	.70	SAT (V,M)	.65	.61	.50	.63	.62	.53	.64
.69	.80	.69	.57	.65	.70	HSGPA	.69	.58	.59	.65	.57	.61	.69
+.07	+.05	+.07	+.08	+.08	+.09	SAT Increment	+.07	+.11	+.04	+.10	+.11	+.04	+.06
.76	.85	.76	.65	.73	.79	VMH	.76	.69	.63	.75	.68	.65	.75

*Correlations corrected for restriction of range and criterion unreliability

the high and low composite groups were greater using HSGPA as a predictor: .76 for FGPA and .68 for course grade in the high composite group, compared to .44 and .44 in the low composite group. Although the SAT predicted better in the high composite group, the SAT increment was greatest in the low composite group: .09 for FGPA and .14 for course grade, compared to .06 and .11 in the high composite group.

A very surprising result was the large increment (+.10) for the average grade mean residual over HSGPA and SAT scores, raising the uncorrected multiple correlation for FGPA from .48 to .58. This result indicates the importance of the degree of grading difficulty, as a result of course selection, in predicting FGPA. Instead of two major types of predictors, SAT scores and HSR, there are really three. The three factors contributing to a high FGPA are high aptitude for college work, good performance in high school, and selection of courses that are graded leniently. To eliminate the effect of course selection, SAT and HSR correlations for FGPA that indicate their predictive effectiveness for a given degree of grading difficulty could be obtained from partial correlations, controlling for average grade mean residual. The uncorrected correlation between the SAT and FGPA was increased by .06 as a result of controlling for average grade mean residual.[18]

Instead of establishing a predictor to describe course selection and controlling for it, to eliminate the understatement of predictive validity of SAT scores and HSR for course by course selection, Young (1990) used Item Response Theory methods to correct the FGPA criterion. Perhaps the simplest way to eliminate the effects of course selection and varying degrees of grading difficulty would be to use course grade instead of FGPA as the criterion.

Another approach to this problem is to extend use of the course grade criterion for all courses taken by a student: obtaining an average predicted course grade (predicted FGPA) that is appropriate for the courses taken by each student; correlating this predicted FGPA with the student's actual FGPA (eliminating courses with fewer than seven freshmen from both predicted and actual FGPA) for every college with at least five students in a student group; and averaging these correlations over all colleges, weighted by the number of students in a student group at the college. These average correlations are also shown in Table 4, with predicted FGPA based on SAT scores, HSGPA, and both, uncorrected and corrected for restriction of range. The average uncorrected correlation of predicted and actual FGPA using both SAT scores and HSGPA was .60, slightly higher than the average FGPA correlation of .58 using the average grade mean residual with both SAT scores and HSGPA. All other correlations of predicted and actual FGPA, for SAT scores and for HSGPA, uncorrected or corrected, were significantly higher than their FGPA or course grade counterparts. For example, using both SAT scores in the predicted FGPA, the uncorrected correlation of predicted and actual FGPA was .50, compared to .36 for the uncorrected multiple correlation using both SAT scores to predict FGPA directly; and the corrected correlation of predicted and actual FGPA was .65, compared to .57 for the corrected multiple correlation using both SAT scores to predict FGPA directly and .60 for the corrected multiple correlation using both SAT scores to predict a single course grade.

Table 5 summarizes for all students selected data from Tables 3 and 4. It shows average correlations using SAT scores and HSGPA to predict FGPA, one specified course grade, and all the grades in the courses taken by a student. The correlations are shown without any correction (as in Table 4), corrected for restriction of range alone (as in Table 3), and corrected for criterion unreliability as well as restriction of range (as in Table 4).

In every case, the average correlation using the SAT was within .04 of the corresponding average correlation using HSGPA. The SAT was slightly better (by .02) in predicting one specified course grade. HSGPA was slightly better (by .01 to .04) in predicting FGPA, either directly or by predicting each of a student's course grades separately. In every case, the average correlation using the SAT-Mathematical score was slightly better (by .01 to .04) than the average correlation using the SAT-Verbal score.

For predicting one specified course grade, the average SAT increment to the correlation over use of HSR alone was .12 when uncorrected or fully corrected, and .10 when corrected only for restriction of range. Table 6 displays the course categories with the highest and lowest SAT increments to the correlation corrected for restriction of range (based on at least 25 courses). The SAT increment for every type of course equaled or exceeded the .07 SAT increment for FGPA (with the exception, by a small amount, of physical education). The highest was .16 for advanced physical sciences/engineering. The lowest were .06 for physical education and .07 for foreign language, either beginning or beyond entry.

[18] r_{sf} = the correlation between the SAT total and FGPA = +.36

r_{sz} = the correlation between the SAT total and the average grade mean residual (z) = −.18

r_{fz} = the correlation between FGPA and z = +.22

$r_{sf.z}$ = the correlation between the SAT total and FGPA, controlled for z

$$= \frac{r_{sf} - r_{sz}\, r_{fz}}{\sqrt{(1-r_{sz})\,(1-r_{fz})}} = \frac{(.36) - (-.18)\,(+.22)}{\sqrt{(1.18)\,(.78)}} = .42$$

$r_{sf.z} - r_{sf} = .42 - .36 = .06$

TABLE 5

Average Correlations for All Students

	PREDICTORS					
Type of Correlation and Criterion	*SAT-V*	*SAT-M*	*SAT*	*HSGPA*	*SAT Increment*	*SAT HSGPA Multiple*
UNCORRECTED CORRELATIONS						
FGPA	.30	.31	.36	.39	+.09	.48
One specified course grade	.23	.26	.32	.30	+.12	.42
FGPA correlation with average of predictions of grades in courses taken	.44	.46	.50	.51	+.09	.60
CORRELATIONS CORRECTED FOR RESTRICTION OF RANGE						
FGPA	.47	.49	.53	.57	+.06	.63
One specified course grade	.41	.44	.49	.47	+.10	.57
FGPA correlation with average of predictions of grades in courses taken	.56	.58	.61	.64	+.07	.71
CORRELATIONS CORRECTED FOR RESTRICTION OF RANGE AND CRITERION UNRELIABILITY						
FGPA	.50	.53	.57	.61	+.07	.68
One specified course grade	.50	.54	.60	.58	+.12	.70
FGPA correlation with average of predictions of grades in courses taken	.60	.62	.65	.69	+.07	.76

Table 7 shows the types of courses for which the two SAT scores predict differently. The difference favoring the verbal score was greatest for all three types of regular courses in English and for history. The difference favoring the mathematical score was greatest for all five mathematics course categories and all three physical sciences/engineering course categories.

Because of greater restriction of range of SAT scores and HSGPA for students taking a typical course, and because of lower reliability for one course grade (.66), Table 5 shows that average uncorrected correlations for one specified course grade are the lowest among the nine types of correlations presented (.32 for the SAT, .30 for HSGPA, and .42 for the SAT-HSGPA multiple). But this type of correlation shows the biggest boost in size (.27 or .28) as a result of correcting for restriction of range and criterion unreliability, to .60 for the SAT, .58 for HSGPA, and .70 for the SAT and HSGPA.

Because of somewhat lower restriction of range and higher reliability (.87), average uncorrected correlations for directly predicting FGPA are somewhat higher (.36 for the SAT, .39 for HSGPA, and .48 for the SAT and HSGPA) than for predicting one specified course grade (.32 for the SAT, .30 for HSGPA, and .42 for the SAT and HSGPA). But because of a smaller boost in size (.18 to .21) as a result of correcting for restriction of range and criterion unreliability, fully corrected average correlations using the SAT are somewhat lower for predicting FGPA directly (.57 for the SAT, .61 for HSGPA, and .68 for the SAT and HSGPA) than for predicting one specified course grade (.60 for the SAT, .58 for HSGPA, and .70 for the SAT and HSGPA).

In all cases, because they allow for criterion comparability by utilizing specific course grades and also allow for multiple courses, thereby reducing criterion unreliability, average correlations between FGPA and the average of the predictions of grades in courses taken are the highest type of correlation. They range from uncorrected correlations of .50 for the SAT, .51 for HSGPA, and .60 for the SAT and HSGPA, to fully corrected correlations of .65 for the SAT, .69 for HSGPA, and .76 for the SAT and HSGPA. The latter three correlations may be the closest estimates ever made of the effectiveness of the SAT and HSGPA for predicting FGPA because they are based on a large cross section of colleges and on comparable course grades for all courses taken by a student, and because the effects of both predictor restriction of range and criterion unreliability are removed.

TABLE 6

SAT Increment by Type of Course

Number of Courses	*Number of Grades*	*High*	*SAT Increment*
56	1,717	Advanced physical sciences/engineering	.16
43	692	Remedial reading/literature	.14
212	7,809	Computer science	.14
531	17,439	History	.13
37	1,106	Home economics	.13
Number of Courses	*Number of Grades*	*Low*	*SAT Increment*
261	9,177	Physical education	.06
558	14,760	Beginning foreign language	.07
404	9,540	Foreign language—beyond entry	.07
47	1,326	Architecture/environmental design	.08
49	3,149	Remedial mathematics	.08

TABLE 7

Comparison of the Effectiveness of Verbal and Mathematical Scores in Predicting Course Grade, by Type of Course

Number of Courses	*Number of Grades*	*Better Prediction Using the Verbal Score*	*Correlations with Course Grade*			*Proportional Contributions*		
			SAT-V	*SAT-M*	*Difference*	*SAT-V*	*SAT-M*	*Difference*
249	7,999	Regular reading/literature	.44	.37	.07	43%	11%	32%
531	17,439	History	.48	.42	.06	44%	16%	28%
226	23,161	Regular writing	.43	.38	.05	41%	12%	29%
213	17,112	Regular English	.43	.38	.05	41%	12%	29%
Number of Courses	*Number of Grades*	*Better Prediction Using the Mathematical Score*	*Correlations with Course Grade*			*Proportional Contributions*		
			SAT-V	*SAT-M*	*Difference*	*SAT-V*	*SAT-M*	*Difference*
93	7,868	Precalculus	.32	.50	.18	2%	44%	42%
49	3,149	Remedial math	.26	.43	.17	4%	41%	37%
128	7,326	Regular math	.35	.50	.15	4%	45%	41%
332	33,346	Calculus	.35	.50	.15	4%	44%	40%
101	4,080	Advanced math	.41	.54	.13	9%	40%	31%
56	1,717	Advanced physical sciences/engineering	.42	.55	.13	17%	45%	28%
400	30,288	Lab/major physical sciences/engineering	.39	.51	.12	8%	42%	34%
209	22,098	Nonlab physical sciences/engineering	.43	.53	.10	15%	43%	28%

Over- and Underpredictions

Table 8 presents average over- and underpredictions, by student group, using the prediction equations for all students. The averages are the differences between the actual level of the criterion and the predicted level of the criterion; positive numbers indicate underpredictions and negative numbers indicate overpredictions. Two different criteria were used: FGPA and course grade. The FGPA prediction equations were developed separately for each college using one FGPA per student. The course grade prediction equations were developed for all courses. For both criteria, over- and underpredictions are displayed in Table 8 using four single predictors—HSGPA, SAT-V, SAT-M, and TSWE—and five sets of multiple predictors: SAT-V and SAT-M; HSGPA, SAT-V, and SAT-M; SAT-V, SAT-M, and TSWE; HSGPA, SAT-V, SAT-M, and TSWE; and HSGPA, SAT-V, SAT-M, and average grade mean residual (for FGPA only).

Over- and underpredictions by academic composite group give apparently paradoxical results, which—it turns out—can be explained as selection artifacts. There is minimal over- or underprediction when HSGPA and test scores are used together, but substantial overprediction for the low composite group and underprediction for the high composite group when HSGPA or test scores are used separately.

The groupings within each college were based on an academic composite index, which is a linear combination of SAT scores and HSGPA. Thus, the low composite group in each college had relatively low mean test scores and a relatively low mean HSGPA. Similarly, the high composite group had relatively high mean test scores and HSGPA. The mean on the criterion (course grade or FGPA) in each group should correspond to the mean prediction based on HSGPA and test scores together, resulting in (ideally) no under- or overprediction in any of the three groups. For the low composite group, however, the mean on the criterion will be lower than would be predicted based on HSGPA or test scores alone (since these predictions do not take into account the low scores on the predictor not used). Similarly, the mean on the criterion in the high composite group will be higher than would be predicted from HSGPA or test scores used separately (since these predictions ignore the high scores on the other predictor). Consequently, entirely because of selection, if predictions are based on HSGPA or test scores separately, overprediction would be expected for the low composite group and underprediction would be expected for the high composite group.

Table 9 shows the largest average over- and underpredictions for the high and low academic composite groups for all of the 37 course categories with at least 100 grades. The over- and underpredictions were determined by using HSGPA and SAT scores to predict course grades for all students, and by comparing each predicted grade with each actual grade. They indicate nonlinearities in the prediction as a result of special factors in course selection. The underprediction in the low composite group, averaging +.12 for technical/vocational courses indicates a

TABLE 8

Average Over- (–) and Underpredictions (+) (Actual – Predicted) by Student Group, Using Prediction Equations for All Students

All Students	*Academic Composite Group*			*Sex*			*English Best Language*		*Ethnic Group*				
	High	*Medium*	*Low*	*Male*	*Female*		*Yes*	*No*	*American Indian*	*Asian American*	*Black*	*Hispanic*	*White*
						FGPA CRITERION*							
.00	+.13	–.05	–.09	–.02	+.02	HSGPA	.00	+.01	–.32	+.02	–.35	–.24	+.03
.00	+.25	–.01	–.25	–.06	+.06	SAT–V	–.01	+.29	–.33	+.14	–.34	–.18	+.02
.00	+.23	–.02	–.23	–.11	+.10	SAT–M	.00	+.01	–.32	+.03	–.28	–.16	+.02
.00	+.19	–.01	–.19	–.10	+.09	SAT	.00	+.18	–.29	+.08	–.23	–.13	+.01
.00	+.02	–.03	+.01	–.06	+.06	HSGPA, SAT	.00	+.15	–.24	+.04	–.16	–.13	+.01
.00	+.28	–.03	–.27	–.03	+.03	TSWE	–.01	+.28	–.35	+.17	–.36	–.20	+.02
.00	+.19	–.02	–.18	–.09	+.08	SAT, TSWE	–.01	+.22	–.29	+.11	–.22	–.12	+.01
.00	+.03	–.04	+.01	–.05	+.05	HSGPA, SAT, TSWE	.00	+.18	–.24	+.06	–.15	–.12	+.01
.00	+.02	–.03	+.01	–.03	+.03	HSGPA, SAT, Z**	.00	+.15	–.23	+.09	–.16	–.10	+.01
						COURSE GRADE CRITERION***							
.00	+.11	–.04	–.08	+.01	–.01	HSGPA	.00	+.04	–.31	+.06	–.31	–.22	+.03
.00	+.22	–.02	–.23	–.03	+.03	SAT–V	–.01	+.25	–.32	+.15	–.30	–.16	+.01
.00	+.20	–.02	–.28	–.07	+.07	SAT–M	.00	+.07	–.31	+.08	–.22	–.13	+.01
.00	+.16	–.01	–.17	–.06	+.06	SAT	.00	+.18	–.27	+.12	–.18	–.10	+.01
.00	.00	–.02	+.03	–.03	+.03	HSGPA, SAT	.00	+.16	–.22	+.08	–.12	–.09	.00
.00	+.25	–.03	–.25	–.01	+.01	TSWE	–.01	+.23	–.34	+.18	–.32	–.18	+.01
.00	+.15	–.01	–.16	–.05	+.05	SAT, TSWE	–.01	+.21	–.26	+.13	–.17	–.09	.00
.00	.00	–.02	+.03	–.02	+.02	HSGPA, SAT, TSWE	.00	+.18	–.21	+.10	–.11	–.08	.00

*Mean standard deviation of FGPA = .71
**Z = Average grade mean residual
***Mean standard deviation of course grade = .81

TABLE 9

Largest Average Over- (–) and Underpredictions (+) for High and Low Academic Composite Groups, by Type of Course,* Using HSGPA and SAT Scores to Predict Course Grade**

HIGH ACADEMIC COMPOSITE GROUP (AVERAGE .00 UNDER- OR OVERPREDICTION)					
Number of Grades	*Underprediction (+)*		*Number of Grades*	*Overprediction (–)*	
5,700	Economics	+.03	102	Remedial writing	–.07
2,501	Nonlab biological sciences	+.03	204	Remedial reading/literature	–.06
762	Advanced physical sciences/engineering	+.03	2,853	Physical education	–.05
LOW ACADEMIC COMPOSITE GROUP (AVERAGE +.03 UNDERPREDICTION)					
Number of Grades	*Underprediction (+)*		*Number of Grades*	*Overprediction (–)*	
131	Technical/vocational	+.12	206	Advanced biological sciences	–.06
389	Advanced mathematics	+.10			
372	Advanced physical sciences/engineering	+.09			

* With 100 or more grades in the group
** Mean standard deviation of FGPA = .71

TABLE 10

Student Groups at More Selective and Less Selective Colleges and at Large and Small Colleges, by Academic Composite Group

Selectivity							*Size*					
High (1121+)			*Low (985–)*				*Large (900+)*			*Small (484–)*		
All	*High Academic Composite Group*	*Low Academic Composite Group*	*All*	*High Academic Composite Group*	*Low Academic Composite Group*		*All*	*High Academic Composite Group*	*Low Academic Composite Group*	*All*	*High Academic Composite Group*	*Low Academic Composite Group*
						SEX						
50.4%	53.0%	50.0%	45.1%	41.0%	48.4%	Males	49.3%	53.1%	46.2%	40.7%	37.6%	44.3%
49.6%	47.0%	50.0%	54.9%	59.0%	51.6%	Females	50.7%	46.9%	53.8%	59.3%	62.4%	55.7%
						ENGLISH BEST LANGUAGE						
98.5%	98.9%	97.8%	98.0%	98.4%	97.4%	Yes	97.1%	95.7%	96.1%	98.4%	99.0%	97.7%
1.5%	1.1%	2.2%	2.0%	1.6%	2.6%	No	2.9%	4.3%	3.9%	1.6%	1.0%	2.3%
						ETHNIC GROUP						
0.4%	0.1%	0.9%	0.5%	0.4%	0.6%	American Indian	0.4%	0.2%	0.8%	0.3%	0.1%	0.5%
11.2%	12.6%	9.3%	3.2%	3.4%	3.6%	Asian American	11.0%	12.3%	9.4%	3.2%	3.5%	3.1%
5.4%	0.6%	14.6%	4.8%	1.8%	8.4%	Black	6.5%	1.8%	14.0%	3.9%	1.1%	8.5%
3.0%	0.8%	5.8%	1.6%	1.3%	1.8%	Hispanic	4.7%	2.6%	7.5%	0.8%	0.2%	1.5%
80.0%	85.9%	69.4%	89.8%	93.2%	85.6%	White	77.3%	83.1%	68.4%	91.8%	95.1%	86.4%

high level of technical/vocational skill that compensates for lower academic credentials. In the same way, underprediction in the low composite group, averaging +.10 for advanced mathematics and +.09 for advanced physical sciences/engineering, indicates some skills in those areas that account for selection of an advanced course despite lower academic credentials. Analogously, overprediction in the high composite group, averaging –.07 for remedial writing and –.06 for remedial reading/literature indicates some weakness in these areas that accounts for taking a remedial course despite higher academic credentials.

Differences Among Colleges

Table 10 shows the student composition of the most selective third of the colleges (SAT-V+M mean of at least 1121), the least selective third (SAT-V+M mean below 986), the largest third (900 or more SAT takers), and the smallest third (fewer than 485 SAT takers). The table also shows the student composition of the highest third and lowest third of the academic composite groups for each of these colleges.

Females were much more heavily represented in the small (59 percent) and less selective (55 percent) colleges, where they tended to be in the high academic composite group (62 percent at small colleges and 59 percent at less selective colleges). At the larger and more selective colleges, the male-female ratios were almost even, with more males than females in the high academic composite group. These distributions are consistent with the fact that, for given academic credentials, females are more likely than males to select colleges closer to home, rather than to attend larger or more selective colleges further away.[19]

Students whose best language was not English were more strongly represented in large colleges (3 percent), where they tended to be in either the high or the low (not the middle) academic composite group. In small or medium-size colleges, they were more likely to be in the low academic group.

Asian American students were much more heavily represented in the more selective and larger colleges, where they tended to be in the high academic composite group. Black and Hispanic students were also more strongly represented in the larger and more selective colleges, but they tended to be in the low academic composite group at all types of colleges. White students made up 90 percent or more of the small (92 percent) and less selective (90 percent) college samples, but not more than 80 percent of the large (77 percent) and more selective (80 percent) college samples. At all types of colleges, they tended to be in the high academic composite group.

Table 11 shows characteristics, selected from those presented in Tables 2 to 4, for more selective and less selective colleges and for large and small colleges, by academic composite group (only the highest and lowest thirds are shown). Students carried a heavier average course load at less selective (8.4) and large (8.4) colleges than at more

[19]See Ramist (1978).

TABLE 11

Characteristics of More Selective and Less Selective Colleges and of Large and Small Colleges, by Academic Composite Group

Selectivity							*Size*					
High (1121+)			*Low (985–)*				*Large (900+)*			*Small (484–)*		
All	*High Academic Composite Group*	*Low Academic Composite Group*	*All*	*High Academic Composite Group*	*Low Academic Composite Group*		*All*	*High Academic Composite Group*	*Low Academic Composite Group*	*All*	*High Academic Composite Group*	*Low Academic Composite Group*
8.0	8.1	7.9	8.4	8.9	8.0	Average number of courses	8.4	8.8	8.0	7.9	8.2	7.6
9%	10%	7%	2%	3%	1%	% of courses advanced	6%	8%	4%	7%	8%	5%
1%	0%	2%	4%	3%	6%	% of courses remedial	1%	0%	3%	1%	1%	2%
32	29	28	20	18	18	Number of courses accounting for half of all credits	29	27	26	23	22	18
58	89	66	64	92	47	*SD* of course SAT means	74	84	58	54	95	56
.12	.17	.05	.11	.22	–.08	Correlation between course grade mean and SAT mean	.13	.20	-.05	.09	.21	–.04
2.87	3.15	2.55	2.32	2.76	1.96	Mean FGPA	2.64	3.02	2.25	2.64	2.98	2.32
.00	–.03	+.03	.00	–.03	+.04	Average grade mean residual (Z)	.00	–.04	+.04	.00	–.02	+.03
.11	.20	.16	.13	.17	.28	Z Increment*	.09	.14	.20	.14	.18	.24
						PROPORTIONAL CONTRIBUTIONS TO PREDICT GRADES IN STUDENT-SELECTED COURSES						
27	26	29	26	25	26	SAT-V	23	25	21	30	29	30
29	30	28	25	25	24	SAT-M	30	29	31	24	25	24
43	43	43	49	49	48	HSGPA	46	45	47	46	45	46
						CORRELATIONS USING SAT AND HSGPA:**						
.69	.83	.60	.58	.71	.36	For FGPA	.62	.74	.50	.65	.79	.54
.62	.65	.55	.53	.61	.40	For course grade	.58	.65	.48	.56	.61	.46
.07	.18	.05	.05	.10	–.04	Difference	.04	.09	.02	.09	.18	.08
						CORRELATIONS USING SAT:**						
.59	.64	.53	.48	.55	.32	For FGPA	.52	.58	.43	.55	.62	.48
.53	.50	.48	.45	.48	.35	For course grade	.50	.52	.43	.48	.49	.40
.06	.14	.05	.03	.07	–.03	Difference	.02	.06	.00	.07	.13	.08
						COURSE GRADE CORRELATIONS:***						
.54	.50	.48	.47	.50	.36	SAT-Verbal	.50	.53	.43	.49	.50	.42
.58	.57	.52	.48	.52	.37	SAT-Math	.55	.59	.47	.50	.52	.43
.65	.62	.59	.55	.59	.43	SAT (V,M)	.62	.64	.53	.59	.60	.49
.62	.68	.52	.53	.64	.34	HSGPA	.59	.69	.46	.54	.63	.41
.14	.12	.16	.12	.11	.15	SAT Increment	.12	.11	.13	.15	.12	.16
.76	.80	.68	.65	.75	.49	VMH	.71	.80	.59	.69	.75	.57

* Z increment = The difference between "the uncorrected correlation of SAT, HSGPA, and average grade mean residual prediction of FGPA" and "the uncorrected correlation of SAT, HSGPA prediction of FGPA."

**Correlations corrected for restriction of range

***Correlations corrected for restriction of range and for criterion unreliability

TABLE 12

Student Characteristics by Sex

	Male		*Female*	
	N	*Percent*	N	*Percent*
ACADEMIC COMPOSITE				
High	7,988	35.6	8,022	33.5
Middle	7,245	32.3	8,095	33.8
Low	7,179	32.0	7,850	32.8
TOTAL	22,412		23,967	
ENGLISH BEST LANGUAGE				
Yes	21,526	97.2	23,173	97.8
No	630	2.8	526	2.2
TOTAL	22,156		23,699	
ETHNIC GROUP				
American Indian	89	0.4	95	0.4
Asian American	1,902	8.8	1,946	8.4
Black	918	4.2	1,557	6.7
Hispanic	812	3.7	787	3.4
White	17,941	82.8	18,802	81.1
TOTAL	21,662		23,187	

selective (8.0) and small (7.9) colleges. As expected, however, the courses taken were more frequently advanced and less frequently remedial for the high composite group at more selective colleges (10 percent advanced, 0 percent remedial) than for the low composite group at less selective colleges (1 percent advanced, 6 percent remedial).

Most of the predictive weight was on the SAT for courses at more selective colleges (57 percent) compared to slightly more than half at less selective colleges (51 percent). At small colleges, there was more weight on the SAT-Verbal score (30 percent) than on the SAT-Mathematical score (24 percent); at large colleges, the emphasis was reversed (30 percent on the mathematical score and 23 percent on the verbal score).

Grades were the least comparable for the low academic composite group at less selective colleges. They were so lacking in comparability that the correlation using SAT scores and HSGPA to predict FGPA (.36) was lower than the correlation to predict a single course grade (.40). Course selection and degree of grading difficulty were so important for this group that the increment in correlation using the average grade mean residual was +.28: uncorrected correlations for FGPA were only .06 for HSR, .12 for the SAT, and .17 for the HSR-SAT multiple, jumping to .45 for the HSR-SAT average grade mean residual multiple![20] The correlation between course grade and course SAT mean for this group was negative (−.08), which is associated with lack of comparability of grades. Grades were also less comparable at large colleges than at small colleges.

Mean FGPA was higher by about half a grade at more selective (2.87) than at less selective (2.32) colleges. The difference in mean FGPA between the high and low academic composite groups was greater at less selective (.80) than at more selective (.60) colleges (a typical standard deviation of FGPA was about .70).

For both SAT scores and HSGPA, prediction of course grade was better at more selective than at less selective colleges and better for the high than for the low academic composite groups. For example, the multiple correlation corrected for restriction of range was .65 for the high composite group at more selective colleges, but only .40 for the low composite group at less selective colleges.

For all types of colleges, the correlations using SAT scores to predict course grade exceeded the correlations using HSGPA. The SAT increment was highest at small (+.15) and more selective (+.14) colleges and lowest at large (+.12) and less selective (+.12) colleges.

Differences by Sex

Student Characteristics

Table 12 shows student characteristics by sex. There were more males (2.8 percent) than females (2.2 percent) whose best language was not English. There were more black females (6.7 percent of all females) than black males (4.2 percent of all males). Males (35.6 percent) were more likely than females (33.5 percent) to be in the high academic composite group.

Course Selection and Grading

Table 2 showed the courses selected by males and females. Males took more courses in the physical sciences/engineering (18 percent of all courses compared to 10 percent for females), mathematics at the calculus level or higher (12 percent compared to 8 percent), computer science (3 percent compared to 2 percent), economics (5 percent compared to 4 percent), military science (1 percent compared to less than 0.5 percent), and architecture/environmental design (1 percent compared to less than 0.5 percent). Females took more courses in the social sciences/humanities (21 percent compared to 17 percent for males), foreign language (9 percent compared to 5 percent), biological sciences (6 percent compared to 4 percent), art/music/theater (7 percent compared to 5 percent), English (14 percent compared to 12 percent), physical education (3 percent compared to 2 percent), health/nursing (1 percent compared to less than

[20]These data are not displayed in any of the tables; only the average grade residual of +.28 (.45−.17) is shown in Table 11.

0.5 percent), education (1 percent compared to less than 0.5 percent), and home economics (1 percent compared to less than 0.5 percent).

Table 3 displayed course selection and grading characteristics for males and females. On average, approximately the same optimal predictive weight was on HSGPA in the courses selected by males and females. Because of the more quantitative nature of the courses selected more frequently by males, more of the SAT weight was on the mathematical score (30 percent) than on the verbal score (23 percent) for males. For females, the weights were almost the same: mathematical score 27 percent, verbal score 26 percent.

Grades for females were slightly more comparable than grades for males. The average increase in the corrected correlation predicting FGPA instead of course grade using SAT scores was .06 for females compared to .04 for males.

The courses selected by males were more strictly graded (average grade mean residual of −.03) than the courses selected by females (+.03). Course selection gave females a "bonus" of .03 on FGPA and gave males a "liability" of .03, a total benefit to females over males of .06. This benefit accounted for .06 of the .09 higher FGPA for females (2.67 compared to 2.58, which amounts to about one-eighth of a standard deviation).[21]

Males had higher average SAT scores than females (verbal by 13 points, one-seventh of a standard deviation, and mathematical by 51 points, more than half of a standard deviation) and also higher predicted FGPA (by .02). Females had higher HSGPA (by .07, one-seventh of a standard deviation) and TSWE scores (by 1 point, one-tenth of a standard deviation). Males selected courses in which students had higher average test scores (SAT-Verbal by 6 points and SAT-Mathematical by 21 points), HSGPA (by .02), and predicted FGPA (by .03) than students in the courses selected by females. As a result of their academic credentials and course selection, in terms of average predicted FGPA, neither males nor females were at any competitive advantage or disadvantage in the courses they selected. Males had higher average SAT scores (verbal by 4 points and mathematical by 16 points) and females had a higher HSGPA (by .05) than the averages of all students in their selected courses.

Comparisons of course selection and grading by sex were relatively consistent for both high and low academic composite groups, as shown in Tables 13 and 14. One difference by group stands out however. Table 2 showed that males and females were equally likely to take a mathematics course below the calculus level (5 percent of all course grades for both males and females). For males, the likelihood of taking a mathematics course below the calculus level differed by academic level: only 2 percent for males in the high composite group (compared to 15 percent taking calculus or higher) and 6 percent for males in the low composite group (compared to 8 percent taking calculus or higher). For females, there was no difference by academic level: 5 percent for females in both the high composite group (with 10 percent taking calculus or higher) and low composite group (with 5 percent taking calculus or higher).

To adjust for the relatively large number of black students who were female and for students whose best language was not English who were male, data on white students whose best language was English were included in Tables 13 and 14. These data show comparisons similar to those Tables 2 and 3 presented for all males and all females.

Predictive Effectiveness

Table 4 displayed correlations with FGPA, with course grade, and between predicted and actual FGPA for males and females. The correlations for females were higher than those for males, more so for the SAT (by .05 to .08) than for HSGPA (by .02 to .05), and more so for the verbal score (by .05 to .08) than for the mathematical score (by .04 to .07).[22]

For both uncorrected and corrected correlations with FGPA for males, the correlation for HSGPA exceeded the correlation for the SAT. But for females, in each case the correlation for the SAT exceeded the correlation for HSGPA. The SAT increment was also greater for females than for males in absolute terms (by .01 to .03) and in percentage of improvement (by about 20 percent of the HSGPA correlation). Using both HSGPA and SAT scores, the corrected correlation for predicting FGPA was higher for females (.71) than for males (.65), as was the corrected correlation for predicting course grade (.74 for females compared to .69 for males) and the corrected correlation between predicted and actual FGPA (.79 for

[21]McCornack and McLeod (1988) hypothesized that "A possible explanation for the higher GPA is that women select less stringently graded courses in greater numbers" (p. 329). They found a correlation of +.40 between course grade and percentage of women in a course.

[22]There are indications that these differences had decreased over time. Ramist (1984) reported results from over 500 colleges during the period 1964 to 1981, in which correlations for females exceeded correlations for males by an average of .08 for the SAT and .05 for HSR. Clark and Grandy (1984) reported results from 30 colleges for 1980, in which the SAT difference was .06 to .07. But Ramist and Weiss (1990) reported data that indicate increases in these differences in the late 1980s after a decade-long decline.

TABLE 13

Comparisons of Courses Selected for High and Low Academic Composite Groups by Sex and for White Students Whose Best Language Is English

Academic Composite Group					*White, English Best Language*	
High		*Low*				
Male	*Female*	*Male*	*Female*		*Male*	*Female*
8.8	8.9	8.1	8.1	Average number of courses	8.6	8.7
7%	8%	3%	5%	% of courses advanced	5%	7%
1%	1%	3%	2%	% of courses remedial	1%	1%
				COURSE CATEGORIES:		
3%	1%	–	–	Advanced mathematics	1%	1%
12%	9%	8%	5%	Calculus	10%	7%
1%	2%	3%	2%	Precalculus	2%	2%
1%	2%	2%	2%	Regular mathematics	2%	2%
–	1%	1%	1%	Remedial mathematics	1%	1%
–	–	–	–	Advanced English	–	–
3%	4%	5%	5%	Regular English	4%	5%
–	–	–	–	Remedial English	–	–
–	–	–	–	Advanced reading/literature	–	–
2%	3%	2%	2%	Regular reading/literature	2%	3%
–	–	–	–	Remedial reading/literature	–	–
–	–	–	–	Advanced writing	–	–
5%	5%	6%	7%	Regular writing	6%	6%
–	–	1%	1%	Remedial writing	–	–
–	–	–	–	Advanced biological sciences	–	–
2%	3%	2%	3%	Lab/major biological sciences	2%	3%
1%	2%	2%	3%	Nonlab and nonmajor biological sciences	2%	3%
1%	–	–	–	Advanced physical sciences/engineering	1%	–
15%	8%	6%	3%	Lab or major physical sciences/engineering	10%	5%
7%	5%	6%	4%	Nonlab and nonmajor physical sciences/engineering	7%	5%
2%	4%	2%	4%	Foreign language–beyond entry	2%	4%
3%	4%	4%	5%	Beginning foreign language	4%	5%
5%	5%	6%	5%	History	6%	5%
4%	4%	5%	3%	Economics	5%	4%
15%	19%	19%	23%	Social sciences/humanities	17%	21%
2%	3%	3%	3%	Business/communications	3%	3%
2%	3%	2%	3%	Art/music/theater–studio	2%	3%
2%	4%	3%	4%	Art/music/theater–nonstudio	3%	4%
3%	2%	2%	1%	Computer science	3%	1%
–	1%	–	1%	Health/nursing	–	1%
–	1%	1%	1%	Education	–	1%
2%	3%	3%	3%	Physical education	2%	3%
1%	1%	1%	1%	Other	1%	1%
1%	–	1%	–	Military science	1%	–
–	–	–	1%	Home economics	–	1%
–	–	–	–	Technical/vocational	–	–
–	–	1%	–	Architecture/environmental design	1%	–

TABLE 14

Comparisons of Course Selection and Grading Characteristics for High and Low Academic Composite Groups by Sex and for White Students Whose Best Language Is English

Academic Composite Group					*White, English Best Language*	
High		*Low*				
Male	*Female*	*Male*	*Female*		*Male*	*Female*
21	24	19	24	Number of courses accounting for half of all credits	21	25
85	87	62	56	*SD* of course SAT means	99	97
.27	.15	−.02	.00	Correlation between course grade mean and SAT mean	.36	.25
2.96	3.05	2.21	2.31	Mean FGPA	2.60	2.71
−.06	.00	+.02	+.06	Average grade mean residual	−.02	+.03
				PROPORTIONAL CONTRIBUTIONS TO PREDICT GRADES IN STUDENT-SELECTED COURSES:		
24%	25%	24%	25%	SAT-V	24%	25%
28%	24%	24%	24%	SAT-M	24%	24%
51%	51%	52%	51%	HSGPA	52%	51%
				CORRELATIONS USING SAT AND HSGPA:*		
.74	.76	.45	.54	For FGPA	.61	.65
.63	.64	.47	.50	For course grade	.55	.59
.11	.12	−.02	.04	Difference	.06	.06
				CORRELATIONS USING SAT:*		
.60	.60	.39	.49	For FGPA	.52	.56
.51	.51	.41	.45	For course grade	.47	.50
.09	.09	−.02	.04	Difference	.05	.06
				SAT-VERBAL MEAN:		
566	556	456	442	Group	521	508
522	511	496	492	Courses selected	509	502
+44	+45	−40	−50	Difference	+12	+6
				SAT-MATH MEAN:		
643	592	525	475	Group	590	540
593	561	552	537	Courses selected	568	548
+50	+31	−27	−62	Difference	+22	−8
				TSWE MEAN:		
52	53	44	45	Group	49	50
50	50	48	48	Courses selected	49	49
+2	+3	−4	−3	Difference	−	+1
				HSGPA MEAN:		
3.78	3.81	2.88	3.01	Group	3.36	3.45
3.50	3.44	3.34	3.35	Courses selected	3.40	3.38
+.28	+.37	−.46	−.34	Difference	−.04	+.07
				PREDICTED GPA MEAN:		
3.00	2.96	2.24	2.24	Group	2.66	2.64
2.73	2.67	2.57	2.57	Courses selected	2.64	2.61
+.27	+.29	−.33	−.33	Difference	+.02	+.03
				COURSE GRADE MEAN:		
3.02	3.09	2.30	2.38	Group	2.69	2.77
2.73	2.73	2.65	2.70	Courses selected	2.68	2.71
+.29	+.36	−.35	−.32	Difference	+.01	+.06

*Correlations corrected for restriction of range

TABLE 15

Comparisons of Predictive Effectiveness for High and Low Academic Composite Groups by Sex and for White Students Whose Best Language Is English

Academic Composite Group					*White, English Best Language*	
High		*Low*				
Male	*Female*	*Male*	*Female*		*Male*	*Female*
				FGPA CORRELATIONS		
				UNCORRECTED:		
.11	.15	.10	.16	SAT-Verbal	.25	.31
.18	.17	.14	.19	SAT-Math	.28	.31
.13	.16	.12	.18	TSWE	.23	.28
.28	.32	.27	.37	Average grade mean residual (Z)	.NA	.NA
.21	.22	.17	.23	SAT (V,M)	.32	.36
.19	.18	.11	.10	HSGPA	.37	.38
+.13	+.14	+.14	+.19	SAT Increment	+.07	+.09
.32	.32	.25	.29	VMH	.44	.47
+.13	+.15	+.17	+.21	Z Increment	+.10	+.11
.45	.47	.42	.50	VMHZ	.54	.58
				CORRECTED:*		
.56	.57	.35	.45	SAT-Verbal	.48	.54
.58	.60	.39	.52	SAT-Math	.52	.56
.64	.64	.42	.53	SAT (V,M)	.56	.60
.73	.75	.40	.47	HSGPA	.59	.62
+.06	+.07	+.08	+.11	SAT Increment	+.06	+.08
.79	.82	.48	.58	VMH	.65	.70
				COURSE GRADE CORRELATIONS		
				CORRECTED:*		
.50	.53	.41	.46	SAT-Verbal	.48	.52
.58	.57	.44	.49	SAT-Math	.52	.55
.63	.63	.50	.55	SAT (V,M)	.58	.62
.66	.65	.44	.46	HSGPA	.55	.58
+.11	+.14	+.14	+.16	SAT Increment	+.13	+.15
.77	.79	.58	.62	VMH	.68	.73
				CORRELATION OF PREDICTED AND ACTUAL FGPA		
				UNCORRECTED:		
				Predicted FGPA based on:		
.31	.36	.31	.38	SAT-Verbal	.39	.45
.35	.39	.32	.43	SAT-Math	.41	.47
.37	.42	.34	.44	SAT (V,M)	.44	.51
.38	.43	.32	.35	HSGPA	.48	.52
+.09	+.08	+.13	+.17	SAT Increment	+.08	+.08
.47	.51	.45	.52	VMH	.56	.60
				CORRECTED:*		
				Predicted FGPA based on:		
.60	.62	.49	.59	SAT-Verbal	.57	.62
.65	.67	.49	.62	SAT-Math	.59	.64
.69	.70	.53	.64	SAT (V,M)	.62	.69
.78	.80	.52	.59	HSGPA	.65	.70
+.06	+.05	+.08	+.11	SAT Increment	+.08	+.08
.84	.85	.60	.70	VMH	.73	.78

*Correlations corrected for restriction of range and criterion unreliability

TABLE 16

Comparisons of Over- (–) and Underpredictions (+) by Sex (Actual–Predicted) for High and Low Academic Composite Groups and for White Students Whose Best Language Is English

Academic Composite Group					*White, English Best Language*	
High		*Low*				
Male	*Female*	*Male*	*Female*		*Male*	*Female*
				FGPA CRITERION		
+.09	+.17	–.09	–.08	HSGPA	+.01	+.06
+.18	+.31	–.31	–.18	SAT–V	–.05	+.09
+.12	+.35	–.35	–.13	SAT–M	–.09	+.14
+.09	+.29	–.29	–.09	SAT	–.09	+.11
–.05	+.09	–.04	+.06	HSGPA, SAT	–.05	+.07
+.24	+.32	–.31	–.24	TSWE	–.03	+.06
+.10	+.27	–.27	–.10	SAT, TSWE	–.08	+.10
–.04	+.09	–.03	+.05	HSGPA, SAT, TSWE	–.05	+.06
–.02	+.07	–.02	+.04	HSGPA, SAT, Z*	–.03	–.04
				COURSE GRADE CRITERION		
+.09	+.12	–.05	–.11	HSGPA	+.03	+.02
+.18	+.25	–.26	–.19	SAT–V	–.03	+.05
+.12	+.28	–.28	–.15	SAT–M	–.06	+.09
+.09	+.22	–.23	–.10	SAT	–.06	+.07
–.04	+.04	+.01	+.04	HSGPA, SAT	–.03	+.03
+.23	+.26	–.26	–.25	TSWE	.00	+.03
+.09	+.20	–.21	–.10	SAT, TSWE	–.06	+.06
–.04	+.03	+.02	+.04	HSGPA, SAT, TSWE	–.02	+.02

*Z = Average grade mean residual

females compared to .73 for males).

As shown in Table 15, sex differences in predictive effectiveness were essentially the same when the groups were limited to white students whose best language was English. On the other hand, when the high academic composite group was isolated, sex differences in predictive effectiveness were almost eliminated.

Predictive effectiveness was greater for all females than for all males in almost every course category for which there were at least 25 courses analyzed. Using SAT scores to predict course grade and corrected correlations, the largest differences were in architecture/environmental design (.56 for females and .34 for males), home economics (.46 for females and .28 for males), and remedial reading/literature (.50 for females and .36 for males). Prediction was slightly better for males in only three course categories: studio art/music/theater (.36 for males and .31 for females), advanced mathematics (.53 for males and .52 for females), and education (.39 for males and .38 for females).

Over- and Underpredictions

Table 8 displayed average over- and underpredictions for males and females using prediction equations for all students. Over- and underpredictions are shown for FGPA and course grade in terms of eight different sets of predictors (four single predictors and four sets of multiple predictors), plus one set using average grade mean residual with HSGPA and SAT scores for FGPA only.

FGPA for females was underpredicted for all nine sets of predictors, ranging from an average of .02 for HSGPA to .10 for the SAT-Mathematical score (with FGPA for males overpredicted by almost the same amounts). Prediction of course grade as opposed to FGPA eliminated the average grade mean residual "bonus" for females of .03 and the "liability" for males of .03. With the exception of a few cases of a .02 or .04 change due to rounding, all average underpredictions for females and overpredictions for males were reduced by .03 in predicting course grade as opposed to FGPA.

Using HSGPA, the average underprediction for females of .02 for FGPA was changed to an average overprediction of .01 for course grade; the average overprediction for males of .02 for FGPA was changed to an average underprediction of .01 for course grade.

Using the two SAT scores, the average underprediction for females of .09 for FGPA was reduced to .06 for course grade; the average overprediction for males

TABLE 17

Average Over- (–) and Underpredictions (+) of Course Grades for Females by Type of Course

Single Predictors					*Multiple Predictors*		
HSGPA	*SAT-V*	*SAT-M*	*TSWE*		*SAT*	*HSGPA SAT*	*HSGPA/ SAT/TSWE*
–.01	+.03	+.07	+.01	Overall average	+.06	+.03	+.02
				TYPES OF COURSES WITH OVERPREDICTIONS (–)			
–.21	–.17	–.13	–.20	Technical/vocational	–.14	–.17	–.18
–.10	–.08	–.05	–.09	Advanced physical sciences/engineering	–.05	–.07	–.08
–.12	–.07	.00	–.09	Nonlab physical sciences/engineering	–.01	–.05	–.05
–.10	–.05	+.02	–.08	Economics	.00	–.04	–.04
–.09	–.03	+.02	–.06	Computer science	+.01	–.03	–.04
–.03	–.02	.00	–.02	Physical education	–.01	–.02	–.02
–.06	–.02	+.03	–.05	Lab/major biological sciences	+.02	–.01	–.02
–.07	–.02	+.04	–.04	Lab or major physical sciences/engineering	+.04	–.01	–.02
–.03	.00	+.04	–.02	History	+.02	–.01	–.01
–.03	+.04	+.04	+.01	Architecture	+.05	.00	–.01
				TYPES OF COURSES WITH UNDERPREDICTIONS (+)			
+.11	+.15	+.18	+.13	Remedial reading/literature	+.15	+.11	+.10
+.06	+.09	+.12	+.07	Beginning language	+.11	+.08	+.07
+.04	+.07	+.10	+.05	Remedial writing	+.10	+.08	+.07
+.05	+.09	+.10	+.06	Regular English	+.10	+.08	+.06
+.05	+.08	+.10	+.05	Regular writing	+.09	+.07	+.05
+.05	+.07	+.09	+.04	Advanced English	+.09	+.07	+.05
+.03	+.07	+.09	+.05	Regular reading/literature	+.08	+.06	+.05
+.04	+.05	+.07	+.04	Advanced language	+.07	+.05	+.05
.00	+.07	+.12	+.05	Precalculus	+.12	+.05	+.05
.00	+.07	+.12	+.05	Regular math	+.11	+.05	+.05
–.01	+.07	+.11	+.05	Remedial math	+.11	+.05	+.05
–.01	+.06	+.11	+.04	Calculus	+.11	+.05	+.05
+.04	+.07	+.10	+.05	Health/nursing	+.09	+.05	+.04
+.03	+.06	+.07	+.03	Other	+.07	+.05	+.04
+.04	+.04	+.05	+.01	Advanced biological sciences	+.05	+.05	+.04
+.01	+.06	+.09	+.03	Business/communications	+.09	+.04	+.04
+.03	+.04	+.07	+.03	Remedial English	+.06	+.04	+.03
+.02	+.05	+.07	+.03	Nonstudio art/music/theater	+.06	+.04	+.03
+.02	+.05	+.06	+.05	Education	+.06	+.03	+.03
+.02	+.03	+.05	+.02	Studio art/music/theater	+.04	+.03	+.02
–.02	+.04	+.08	+.01	Military science	+.07	+.03	+.02
–.01	+.03	+.06	.00	Social sciences/humanities	+.06	+.03	+.02
+.02	+.05	+.08	+.04	Advanced writing	+.06	+.02	+.02
+.01	+.02	+.02	+.01	Home economics	+.02	+.01	+.01
–.04	+.01	+.07	.00	Advanced math	+.06	+.01	+.01
+.01	+.01	+.04	–.01	Advanced reading/literature	+.02	+.01	.00
–.04	+.01	+.05	–.02	Nonlab and nonmajor biological sciences	+.04	+.01	.00

TABLE 18

Characteristics of More Selective and Less Selective Colleges and of Large and Small Colleges by Sex

Selectivity					*Size*			
High (1121+)		*Low (985–)*			*Large (900+)*		*Small (484–)*	
Males	*Females*	*Males*	*Females*		*Males*	*Females*	*Males*	*Females*
8.1	8.0	8.3	8.6	Average number of courses	8.3	8.5	7.9	8.0
8%	9%	2%	3%	% of courses advanced	5%	6%	4%	8%
1%	1%	4%	4%	% of courses remedial	1%	1%	2%	1%
25	35	18	20	Number of courses accounting for half of all credits	24	31	17	26
61	62	69	67	*SD* of course SAT mean	74	74	60	59
.17	.12	.13	.14	Correlation between course grade mean and SAT mean	.18	.15	.10	.13
2.85	2.90	2.18	2.42	Mean FGPA	2.61	2.67	2.48	2.75
–.02	+.01	–.03	+.03	Average grade mean residual (Z)	–.04	+.04	–.01	+.01
.09	.12	.14	.13	Z Increment*	.09	.10	.13	.15
				PROPORTIONAL CONTRIBUTIONS TO PREDICT GRADES IN STUDENT-SELECTED COURSES:				
26%	28%	25%	27%	SAT-V	22%	25%	29%	30%
31%	28%	26%	24%	SAT-M	31%	28%	25%	24%
42%	43%	49%	48%	HSGPA	47%	46%	46%	45%
				CORRELATIONS USING SAT AND HSGPA:**				
.70	.69	.54	.62	For FGPA	.61	.66	.63	.66
.62	.62	.51	.57	For course grade	.57	.60	.55	.57
.08	.07	.03	.05	Difference	.04	.06	.08	.09
				CORRELATIONS USING SAT:**				
.60	.60	.45	.54	For FGPA	.51	.57	.53	.58
.53	.53	.43	.49	For course grade	.48	.52	.47	.49
.07	.07	.02	.05	Difference	.03	.05	.06	.09
				SAT-V + M MEAN:				
1241	1190	938	897	Group	1110	1036	1058	1041
1227	1205	921	911	Courses selected	1087	1058	1043	1051
+14	–15	+17	–14	Difference	+23	–22	+15	–10
				HSGPA MEAN:				
3.61	3.66	2.98	3.15	Group	3.43	3.49	3.16	3.35
3.64	3.62	3.06	3.08	Courses selected	3.47	3.45	3.23	3.30
–.03	+.04	–.08	+.07	Difference	–.04	+.04	–.07	+.05
				COURSE GRADE CORRELATIONS:***				
.54	.53	.43	.50	SAT-Verbal	.48	.54	.49	.50
.59	.58	.47	.52	SAT-Mathematical	.54	.59	.50	.52
.65	.65	.53	.60	SAT (V,M)	.59	.64	.58	.60
.60	.60	.49	.54	HSGPA	.58	.60	.53	.54
+.16	+.16	+.14	+.16	SAT increment	+.12	+.14	+.15	+.16
.76	.76	.63	.70	VMH	.70	.74	.68	.70
				OVER(–)/UNDERPREDICTIONS (+) FOR COURSE GRADE:				
+.02	–.02	–.03	+.02	HSGPA	+.02	–.02	–.01	+.01
–.03	+.04	–.12	+.10	SAT	–.06	+.06	–.08	+.06
–.01	+.01	–.06	+.05	HSGPA, SAT	–.02	+.02	–.04	+.03
.00	.00	–.05	+.04	HSGPA, SAT, TSWE	–.02	+.02	–.03	+.02

* Z Increment = The difference between "the uncorrected correlation of SAT, HSGPA, and average grade mean residual prediction of FGPA" and "the uncorrected correlation of SAT, HSGPA prediction of FGPA."

**Correlations corrected for restriction of range

***Correlations corrected for restriction of range and criterion unreliability

of .10 for FGPA[23] was also reduced to .06 for course grade. Using HSGPA and the two SAT scores, the average underprediction for females of .06 for FGPA[24] was reduced to .03 for course grade; the average overprediction for males of .06 for FGPA[25] was also reduced to .03 for course grade.

The latter reduction of average female underprediction and male overprediction from .06 to .03 using HSGPA and the two SAT scores was accomplished not only by predicting course grade instead of FGPA, but also by using the average grade mean residual (Z) as an additional predictor of FGPA. The predictors HSGPA, SAT-V, SAT-M, and Z for FGPA resulted in average underprediction for females and overprediction for males of .03.[26]

Use of the TSWE score with the two SAT scores, with or without HSGPA, reduced the average underprediction for females and overprediction for males by an additional .01. Using TSWE with the SAT scores, the reduction was from .06 to .05. Using TSWE with SAT scores and HSGPA, the reduction was from .03 to .02 (.02 represents only about 3 percent of a standard deviation). This reduction is due to the fact that women tend to write better, writing ability is important in academic performance, and SAT scores and HSR do not fully measure writing ability (Breland and Griswold 1982; Ramist 1984).

As shown in Table 16, over- and underpredictions by sex were not due to differences related to ethnic background or whether English was the best language. Similar differences were found for white students whose best language was English and for both high and low academic composite groups.

Although the average underprediction of course grade for females using HSGPA and SAT scores was .03, there were large differences among the course categories. Table 17 displays under- and overpredictions for the course categories. The largest overpredictions were in technical courses other than math: technical/vocational, physical sciences/engineering, economics, and computer science. The largest underpredictions were in the various English courses and in beginning foreign language.

TABLE 19

Summary of Over- (–) and Underpredictions (+) of Course Grade for Females, and of FGPA, Average Grade Mean Residual, and SAT Mean, by Sex and by Selectivity of College

All Colleges		*Selectivity*		
		More	*Average*	*Less*
	PREDICTOR(S)			
–.01	HSGPA	–.02	–.02	+.02
+.03	SAT-Verbal	+.01	+.03	+.07
+.07	SAT-Math	+.04	+.06	+.10
+.01	TSWE	–.01	.00	+.05
+.06	SAT	+.04	+.06	+.10
+.03	HSGPA,SAT	+.01	+.02	+.05
+.02	HSGPA,SAT,TSWE	.00	+.02	+.04
	FGPA			
2.67	Females	2.90	2.67	2.42
2.58	Males	2.85	2.59	2.18
+.09	Difference	+.05	+.08	+.24
	Z*			
+.03	Females	+.01	+.03	+.03
–.03	Males	–.02	–.03	–.03
+.06	Difference	+.03	+.06	+.06
	SAT MEAN			
1031	Females	1189	1015	893
1095	Males	1242	1078	937
–64	Difference	–53	–63	–44

*Z = Average grade mean residual

Differences Among Colleges

Table 18 shows the characteristics of more selective and less selective colleges and of large and small colleges, by sex. The finding of slightly more comparable grades for females and slightly less comparable grades for males did not hold at more selective colleges. Although the predictive effectiveness of HSGPA and SAT scores was usually better for females, predictive effectiveness for males and for females was almost identical at more selective colleges, actually slightly better for males.

Table 19 summarizes average over- and under-

[23]Stricker, Rock, and Burton (1991) found the same overprediction for males of .10 for FGPA at a large state university.

[24-25]These correspond with the underpredictions for females and overpredictions for males found by Clark and Grandy (1984) for 41 colleges and by Stricker, Rock, and Burton (1991) for a large state university. See also Linn (1973), Wild (1977), and Ramist (1984, p. 154). Using HSGPA and ACT scores, ACT (1973), Gamache and Novick (1985), and Sawyer (1986) reported underpredictions for females and overpredictions for males ranging from .05 to .12. Reductions in under- and overpredictions were reported as a result of separate analyses by area of study for the ACT in Gamache and Novick (1985) and for the SAT in Clark and Grandy (1984).

[26]By predicting scores in two standard courses and also by adjusting FGPA with a pairwise matching technique, Elliott and Strenta (1988) reduced, but did not eliminate, the underprediction for women. McCornack and McLeod (1988) reported complete elimination of underprediction and slight overprediction for women by predicting grades in 88 introductory courses at a large state university. But FGPA adjustment at another large state university, reported by Stricker, Rock, and Burton (1991), had a negligible effect on underprediction for women.

predictions of course grade (with FGPA, average grade mean residual, and SAT mean) by sex for all, more selective, and less selective colleges. At more selective colleges, the typical underprediction for females and overprediction for males was substantially reduced; there was no under- or overprediction using HSGPA, SAT scores, and TSWE to predict course grade. At a very selective college, Young (1991) found no significant under- or overprediction for females or males based on HSGPA and SAT scores when FGPA was adjusted by means of Item Response Theory.

At less selective colleges, males earned a very low mean FGPA and there was greater than average overprediction for males. The mean FGPA of only 2.18 for males was about one-quarter of a grade, or one-third of a standard deviation, below the mean of 2.42 for females. This difference was five times the difference at more selective colleges (.05) and three times the difference at colleges with average selectivity (.08). Especially since males at less selective colleges had higher mean SAT scores than their female classmates by 44 total SAT points, it is clear that for some reason males at less selective colleges did not perform up to their potential.[27]

Why do women obtain higher grades than predicted, especially at less selective colleges? Caldwell and Hartnett (1967) speculate about attendance, attitude, punctuality, and neatness. Stricker, Rock, and Burton (1991) suggest that it is a result of differences in studiousness and in how males and females handle lack of confidence. With respect to students who have lower levels of academic ability and who are likely to attend less selective colleges where males receive especially low grades given their academic ability, the authors state: "Women lacking confidence in their academic ability may compensate by working harder while men in that situation simply give up."

Differences by English as Best or Not Best Language

Student Characteristics

Table 20 shows the characteristics of students whose best language was not English. These students tended to be male (55 percent) and nonwhite (70 percent); 57 percent were Asian American. They were twice as likely to be in the low academic composite group (44 percent) than in the high composite group (23 percent).

[27]Females also outperformed males at small colleges by an average of one-quarter of a grade, but the difference in SAT score means was only 17 total points.

TABLE 20

Student Characteristics by English as Best or Not Best Language

	English Best Language		*English Not Best Language*	
	N	*Percent*	N	*Percent*
ACADEMIC COMPOSITE				
High	15,614	34.9	262	22.7
Middle	14,791	33.1	389	33.7
Low	14,294	32.0	505	43.7
TOTAL	44,699		1,156	
SEX				
Male	21,526	48.2	630	54.5
Female	23,173	51.8	526	45.5
TOTAL	44,699		1,156	
ETHNIC GROUP				
American Indian	184	0.4	0	0.0
Asian American	3,218	7.4	614	56.9
Black	2,424	5.6	42	3.9
Hispanic	1,497	3.4	102	9.4
White	36,345	83.2	322	29.8
TOTAL	43,668		1,080	

Course Selection and Grading

Tables 2 and 3 showed courses selected and other course and grading characteristics by whether English was the best language. Students whose best language was not English averaged fewer courses (8.1 compared to 8.5 for students whose best language was English), but more advanced courses (7 percent compared to 6 percent), and also more remedial courses (3 percent compared to 1 percent). Among course categories, they took more courses in physical sciences/engineering (22 percent compared to 14 percent), mathematics at the calculus level or higher (15 percent compared to 9 percent), and computer science (4 percent compared to 2 percent); they took fewer courses in the social sciences/humanities (13 percent compared to 19 percent), business/communications (1 percent compared to 3 percent), and history (3 percent compared to 5 percent).

Because of the quantitative nature of the courses these students took more frequently, as shown in Table 3, to predict course grade, more weight was put on the SAT-Mathematical score (32 percent compared to 28 percent) and less weight on the SAT-Verbal score (20 percent compared to 25 percent). These courses were more strictly graded than average (average grade mean residual of –.03).

Compared with students whose best language was English, students whose best language was not English had much lower mean SAT-Verbal scores (by 117 points)

TABLE 21

Course Categories with the Largest Over- (–) and Underpredictions (+) Using SAT Scores and HSR for Students Whose Best Language Is Not English (based on 100 or more course grades)

Students Whose Best Language Is Not English (average +.16)					
Grades	*Underpredictions (+)*		*Grades*	*Overpredictions (–)*	
166	Precalculus	+.41	310	Regular English	–.08
1,097	Calculus	+.34	137	Studio art/music/theater	–.03
1,432	Physical sciences/engineering–lab/majors	+.25			
539	Physical sciences/engineering–nonlab/nonmajors	+.25			

and TSWE scores (by 10 points on a 20 to 60+ scale), by about one and one-half standard deviations. But they had a higher mean SAT-Mathematical score (by 9 points) and a higher mean HSGPA (by .04). They selected courses in which students had a higher mean mathematical score (by 16 points) and HSGPA (by .04), but a lower mean verbal score (by 13 points), than courses selected by students for whom English was their best language. As a result of their course selection, these students were at a competitive disadvantage, not only in terms of their verbal score (by 100 points) and their TSWE score (by 8 points), but also, to a lesser degree, in terms of their mathematical score (by 6 points). They were equally competitive with the other students in their courses in terms of HSGPA.

Predictive Effectiveness

As shown in Table 4, after correction for restriction of range, correlations for students whose best language was English slightly exceeded correlations for students whose best language was not English, but the SAT increment was larger for the latter group.

Over- and Underpredictions

As Table 3 showed, despite much lower SAT-Verbal scores, students whose best language was not English had a higher course grade mean (2.75) than other students in the courses they selected (2.71). As Table 8 showed, the result was underprediction of both FGPA and course grade for students whose best language was not English by .18 (about one-quarter of a standard deviation) using HSGPA, SAT scores, and TSWE. As shown in Table 21, the greatest underprediction using HSGPA and SAT scores to predict course grade was in quantitative courses. There was average overprediction in regular English and studio art/music/theater courses.

Differences Among Colleges

Table 22 compares students whose best language was and was not English at more selective and less selective colleges and at large and small colleges. Students whose best language was not English selected more strictly graded courses at larger and more selective colleges, where underprediction was greatest for them, but not at small and less selective colleges, where underprediction was lower.

Although typically the grades of students whose best language was not English were less predictable, at less selective colleges they were more predictable using HSGPA or SAT scores for FGPA or course grade.

At small colleges, students whose best language was not English had less comparable grades, a low FGPA, a low correlation of HSGPA with course grade, a higher correlation of the SAT-Verbal score than of the SAT-Mathematical score with course grade, and a high SAT increment.

Differences by Ethnic Group

Student Characteristics

Table 23 displays the characteristics of students by ethnic group. Half or more of the black, American Indian, and Hispanic students were in the low academic composite group. Black students were mostly female (63 percent). Of the Asian American students, 16 percent indicated that English was not their best language, compared to 6 percent for Hispanic students and 2 percent or less for the other ethnic groups.

Course Selection and Grading

As shown in Table 2, white students averaged the most courses per student (8.6) and Hispanic students the fewest (7.8). White and Asian American students took more higher-level courses (6 percent advanced, 1 percent remedial) than American Indian students (3 percent advanced, 2 percent remedial) and Black students (4 percent advanced, 3 percent remedial).

TABLE 22

Characteristics of More Selective and Less Selective Colleges and of Large and Small Colleges, by English as Best or Not Best Language

Selectivity					*Size*			
High (1121+)		*Low (985-)*			*Large (900+)*		*Small(484-)*	
English Best Language	*English Not Best Language*	*English Best Language*	*English Not Best Language*		*English Best Language*	*English Not Best Language*	*English Best Language*	*English Not Best Language*
32	13	20	13	Number of courses accounting for half of all credits	29	21	23	11
58	99	65	98	*SD* of course SAT mean	75	100	54	95
.12	.15	.11	.18	Correlation between course grade mean and SAT mean	.12	.22	.10	.04
2.87	2.86	2.32	2.25	Mean FGPA	2.64	2.73	2.64	2.45
.00	−.04	.00	+.01	Average grade mean residual (Z)	.00	−.05	.00	+.01
.11	.12	.13	.16	Z increment*	.10	.08	.14	.22
8.0	7.6	8.4	8.2	Average number of courses	8.4	8.0	7.9	7.7
9%	10%	2%	3%	% of courses advanced	6%	8%	7%	9%
1%	4%	4%	6%	% of courses remedial	1%	3%	1%	3%
				PROPORTIONAL CONTRIBUTIONS TO PREDICT GRADES IN STUDENT-SELECTED COURSES:				
27%	22%	26%	23%	SAT-V	23%	18%	30%	27%
29%	34%	25%	28%	SAT-M	30%	33%	24%	27%
42%	43%	49%	48%	HSGPA	46%	47%	46%	45%
				CORRELATIONS USING SAT AND HSGPA:**				
.70	.60	.58	.58	For FGPA	.63	.58	.65	.49
.62	.56	.54	.58	For course grade	.58	.56	.57	.53
.08	.04	.04	.00	Difference	.05	.02	.08	−.04
				CORRELATIONS USING SAT:**				
.60	.47	.48	.52	For FGPA	.53	.49	.56	.31
.53	.46	.45	.52	For course grade	.50	.46	.48	.49
.07	.01	.03	.00	Difference	.03	.03	.08	−.18
				SAT-V + M MEAN:				
1219	1100	918	795	Group	1075	968	1050	919
1217	1191	916	904	Courses selected	1072	1080	1049	1002
+2	−91	+2	−109	Difference	+3	−112	+1	−83
				HSGPA MEAN:				
3.63	3.64	3.07	3.09	Group	3.46	3.51	3.27	3.21
3.63	3.66	3.07	3.09	Courses selected	3.46	3.50	3.27	3.22
.00	−.02	.00	.00	Difference	.00	.00	.00	−.01
				COURSE GRADE CORRELATIONS:***				
.54	.42	.47	.50	SAT-Verbal	.52	.43	.50	.55
.58	.52	.48	.54	SAT-Mathematical	.55	.53	.50	.47
.65	.55	.55	.64	SAT (V,M)	.62	.57	.59	.60
.62	.54	.53	.53	HSGPA	.59	.55	.55	.43
.14	.15	.13	.18	SAT increment	.12	.14	.15	.22
.76	.69	.66	.71	VMH	.71	.69	.70	.65
				OVER-(-)/UNDERPREDICTIONS (+) FOR COURSE GRADE:				
.00	+.05	.00	−.12	HSGPA	.00	+.09	.00	−.13
.00	+.15	.00	+.07	SAT	−.01	+.21	.00	+.01
.00	+.14	.00	+.04	HSGPA, SAT	−.01	+.19	.00	+.01
.00	+.17	.00	+.08	HSGPA, SAT, TSWE	−.01	+.22	.00	+.04

*Z increment = The difference between "the uncorrected correlation of SAT, HSGPA, and average grade mean residual prediction of FGPA" and "the uncorrected correlation of SAT, HSGPA prediction of FGPA."

**Correlations corrected for restriction of range

***Correlations corrected for restriction of range and criterion unreliability

TABLE 23

Student Characteristics by Ethnic Group

	American Indian		*Asian American*		*Black*		*Hispanic*		*White*	
	N	*Percent*	N	*Percent*	N	*Percent*	N	*Percent*	N	*Percent*
ACADEMIC COMPOSITE										
High	32	17	1,489	39	229	9	319	20	13,539	37
Middle	52	28	1,280	33	559	23	477	30	12,462	34
Low	100	54	1,079	28	1,687	68	803	50	10,742	29
TOTAL	184		3,848		2,475		1,599		36,743	
SEX										
Male	89	48	1,902	49	918	37	812	51	17,941	49
Female	95	52	1,946	51	1,557	63	787	49	18,802	51
TOTAL	184		3,848		2,475		1,599		36,743	
ENGLISH BEST LANGUAGE										
Yes	184	100	3,218	84	2,424	98	1,497	94	36,345	99
No	0	0	614	16	42	2	102	6	322	1
TOTAL	184		3,832		2,466		1,599		36,667	

Compared with the other ethnic groups, American Indian students took more courses in nonstudio art/music/theater (5 percent) and in mathematics below the calculus level (7 percent); they took fewer courses in mathematics at the calculus level or higher (6 percent).

Asian American students took more courses in physical sciences/engineering (23 percent), especially those with a lab or for majors (16 percent), mathematics at the calculus level or higher (16 percent), and computer science (3 percent). They took fewer courses in mathematics below the calculus level (3 percent), social sciences/humanities (15 percent), English (11 percent), history (3 percent), and physical education (2 percent).

Black students took more courses in social sciences/humanities (22 percent), English (15 percent), and remedial mathematics (2 percent). They took fewer courses in physical sciences/engineering (11 percent) and mathematics at the calculus level or higher (6 percent).

Hispanic students took more courses in the biological sciences (7 percent), but fewer in physical education (2 percent).

White students took more courses in business/communications (3 percent), but fewer in the biological sciences (4 percent).

As shown in Table 3, for every ethnic group, SAT scores made a proportional contribution of 53 to 55 percent and HSR made a proportional contribution of 45 to 47 percent to the prediction of the courses selected. For all groups except Asian American and Hispanic, the two groups with most of the students whose best language was not English, 28 percent of the weight was on the mathematical score and 25 or 26 percent was on the verbal score. The proportion of weight on the mathematical score was higher and the proportion on the verbal score was lower for courses selected by Asian American (32 percent versus 20 percent) and Hispanic (30 percent versus 22 percent) students.

Course grades for white students were the most comparable. Using HSGPA and SAT scores, the average difference in the corrected correlation from predicting course grade to predicting FGPA was a relatively high .07 and the correlation between course grade and course SAT mean was a relatively high +.30.

Course grades were the least comparable for Hispanic and black students. They were so lacking in comparability for black students that, using SAT scores with or without HSGPA, the average difference in the corrected correlation from predicting course grade to predicting the eight-course FGPA was .00. For Hispanic students, using SAT scores, the corrected correlation for predicting FGPA was lower by .05 (.40) than the corrected correlation for predicting course grade (.45). Associated with the lack of comparability of grades were very high standard deviations of course SAT means for Hispanic (121) and black (128) students.

The courses selected by Asian American students were the most strictly graded and gave them a "liability" in FGPA of .07 (average grade mean residual of –.07), one-tenth of a standard deviation. All the other ethnic groups selected more leniently graded courses, with positive grade mean residuals. Despite their .07 "liability" on FGPA, Asian American students had by far the highest FGPA (2.80). Despite their .06 "bonus" on FGPA, black students had the lowest FGPA (2.14), about two-thirds of a standard deviation below the overall mean FGPA.

Asian American students selected courses in which

students had the highest test scores (SAT-V, SAT-M, and TSWE), about .6 of a standard deviation of course SAT means above the average course, and the highest HSGPA. Hispanic and white students also selected courses in which other students had relatively high academic credentials. Black students (about .6 of a standard deviation) and American Indian students (about .4 of a standard deviation) selected courses in which students had lower credentials.

As a result of their academic credentials and course selection, white students on average were at a slight advantage compared to other students in their courses (+.02 on predicted FGPA, with higher verbal and mathematical scores, and slightly higher HSGPA). Asian American students had neither a competitive advantage nor disadvantage (with higher mathematical scores and HSGPA, but lower verbal scores). Hispanic, American Indian, and especially black students were at a large competitive disadvantage. Compared with other students in their courses, black students had a lower predicted FGPA (by .26), a lower HSGPA (by .18, about .4 of a standard deviation), a lower mathematical mean (by 71 points, about .8 of a standard deviation), a lower verbal mean (by 52 points, about .6 of a standard deviation), and a lower TSWE mean (by 4 points).

Predictive Effectiveness

As shown in Table 4, the highest predictive effectiveness, after correcting for predictor restriction of range and criterion unreliability, was for Asian American students: the multiple correlation using HSGPA and SAT scores was .69 for FGPA and .76 for course grade. Although the correlations for the mathematical score were typically higher than for the verbal score, for Asian American students the mathematical score was such a good predictor that it had at least as high a correlation with course grade (.59) as did either HSGPA or the multiple of both SAT scores for the other ethnic groups.[28]

The second highest predictive effectiveness of the SAT for both FGPA and course grade was for white students. But the benefit in using the SAT, measured by the SAT increment for predicting course grade, was lower for white students (+.12 for course grade) than for black students.

By far the greatest benefit in using the SAT as a predictor was for black students (an SAT increment of +.18 for course grade). In predicting course grade, not only was the SAT correlation moderately high for black students (.57), but the HSGPA correlation was very low (.46). Black students were the only ethnic group that even for predicting FGPA had a higher correlation with the SAT (.49) than HSGPA (.46). For predicting both FGPA and course grade, the multiple correlation using both HSGPA and SAT scores were the lowest of all ethnic groups.[29] But, as shown in Table 4, because of lack of comparability of course grades, the increment in validity for FGPA using the average grade mean residual as an additional predictor, with HSGPA and SAT scores, was the highest (+.17) for black students. TSWE also had the highest predictive effectiveness for black students.

Correlations of test scores with course grade were lowest for American Indian students, especially for the SAT-Mathematical score. American Indian students were the only group for which the verbal score predicted course grade better than the mathematical score and HSGPA predicted course grade better than SAT scores. Their corrected SAT increment for predicting course grade (+.11) was the lowest of all the ethnic groups, but their corrected SAT increment for predicting FGPA (+.14) was the highest.

Correlations of test scores with FGPA were relatively low for Hispanic students.[30] The corrected SAT increment for FGPA (+.05) was the lowest of all the ethnic groups, but the corrected SAT increment for course grade (+.13) was higher than for white students (+.12), as were corrected course grade correlations. When the students

[28]Sue and Abe (1988) also reported a higher correlation for the SAT and HSR in predicting FGPA for Asian American than for white students, and a higher correlation for the mathematical score than for the verbal score for Asian American students, in the University of California system. Morgan (1990) confirmed a higher correlation for the mathematical score and reported corrected correlations with FGPA similar to those reported here for Asian American students.

[29]Ramist (1984, p.158) described results of validity studies for black and white students at 11 predominantly white colleges and for all students at 11 predominantly black colleges. For black students at both types of colleges, the SAT increment for FGPA was very high and the SAT predicted FGPA better than HSR, consistent with findings here in which the 2,475 black students were at predominantly white colleges. But only at predominantly white colleges was the level of prediction for black students low (.30); at predominantly black colleges, the uncorrected correlation for SAT scores was moderately high (.38), even with very great restriction of range, and the corrected correlation was quite high (.57). For other validity studies on black students, see Morgan (1990) and Breland (1979).

[30]Pennock-Roman (1990) reported uncorrected correlations with FGPA for 1,447 Hispanic students and also for non-Hispanic students at six colleges. The slightly lower levels for Hispanic students correspond with those reported here: for SAT scores, .04 (.29 for non-Hispanic students and .25 for Hispanic students), compared to .05 here (.32 for white students and .27 for Hispanic students); for the multiple of SAT scores and HSR, .01 (.40 for non-Hispanic students and .39 for Hispanic students), compared to .02 here (.45 for white students and .43 for Hispanic students). For other studies of prediction for Hispanic students, see Morgan (1990), Ramist (1984), Duran (1983), and Breland (1979).

TABLE 24

Course Categories with Over (–) and Underpredictions (+) of Course Grade Using HSGPA and SAT Scores, by Ethnic Group

	AMERICAN INDIANS (AVERAGE –.22)				
	Underpredictions (+)			*Overpredictions (–)*	
Grades			Grades		
46	Precalculus	+.12	106	Nonlab. physical sciences/engineering	–.46
32	Business/communications	+.09	86	Beginning foreign language	–.36
			72	Lab./major physical sciences/engineering	–.34
			94	Regular writing	–.30
			96	Calculus	–.28
	ASIAN AMERICANS (AVERAGE +.08)				
	Underpredictions (+)			*Overpredictions (–)*	
Grades			Grades		
488	Precalculus	+.24	59	Advanced reading/literature	–.11
75	Health/nursing	+.24	156	Architecture	–.10
84	Remedial mathematics	+.20	62	Remedial English	–.09
3,999	Calculus	+.18	79	Education	–.07
1,974	Nonlab. physical sciences/engineering	+.15	220	Other	–.06
4,622	Lab./major physical sciences/engineering	+.14	426	Physical education	–.04
1,101	Beginning foreign language	+.14	825	Regular English	–.03
			1,634	Regular writing	–.02
	BLACKS (AVERAGE –.12)				
	Underpredictions (+)			*Overpredictions (–)*	
Grades			Grades		
49	Remedial reading/literature	+.07	1,077	Lab./major physical sciences/engineering	–.19
57	Remedial English	+.05	967	Nonlab. physical sciences/engineering	–.18
52	Advanced physical sciences/engineering	+.01	317	Computer Science	–.18
			438	Regular mathematics	–.18
			380	Business/communications	–.17
			492	Nonlab. biological sciences	–.17
			1,213	Calculus	–.16
			698	History	–.15
	HISPANICS (AVERAGE –.09)				
	Underpredictions (+)			*Overpredictions (–)*	
Grades			Grades		
428	Beginning foreign language	+.05	451	Lab./major biological sciences	–.18
64	Home economics	+.02	1,168	Lab./major physical sciences/engineering	–.16
63	Advanced mathematics	+.01	509	Economics	–.14
213	Regular reading/literature	+.01	154	Studio art/music/theater	–.14
275	Foreign language–beyond entry	+.01	563	Regular writing	–.12
234	Business/communications	+.01			
	WHITES (AVERAGE .00)				
	Underpredictions (+)			*Overpredictions (–)*	
Grades			Grades		
497	Remedial writing	+.04	2,954	Advanced mathematics	–.02
731	Advanced reading/literature	+.02	25,873	Calculus	–.02
18,510	Regular English	+.02	631	Technical/vocational	–.02

in the high academic composite group were compared for each ethnic group, the predictive effectiveness of HSGPA and SAT scores was highest for the 319 (20 percent) of the Hispanic students in the high composite group: the multiple correlation of HSGPA and SAT scores was .90 for predicting FGPA and .91 for predicting course grade.

Over- and Underpredictions

Table 8 revealed, on average, moderate underpredictions for Asian American students and large overpredictions for American Indian students in predicting FGPA or course grade from all combinations of HSGPA and test scores. For example, in predicting course grade from HSGPA and SAT scores, the average underprediction was .08 (.1 of a standard deviation) for Asian American students[31] and the average overprediction was .22 (.3 of a standard deviation) for American Indian students. Black and Hispanic students were also overpredicted: e.g., in predicting course grade from HSGPA and SAT scores, the overprediction was .12 (.2 of a standard deviation) for black students[32] and .09 (.1 of a standard deviation) for Hispanic students.[33]

Although there were, on average, underpredictions for Asian American students (and, to a lesser extent, white students) and overpredictions for American Indian, black, and Hispanic students, there were large differences among the course categories. Each group had over- and underpredictions for some course categories, as displayed in Table 24.

American Indian students were overpredicted in a variety of science, language, English, and mathematics

[31]In predicting FGPA from HSGPA and SAT scores, there was an average underprediction of .04 for Asian American students. In their study at the University of California, Sue and Abe (1988, p. 10) reported approximately the same underprediction (.02) for Asian American students.

courses. Two course categories with at least 25 course grades for which they were underpredicted were precalculus and business/communications.

Asian American students were underpredicted in mathematics, physical sciences/engineering, health/nursing, and foreign language courses. They were overpredicted in English, architecture, education, and physical education courses.

Black students were overpredicted in quantitative and science courses— physical sciences, engineering, computer science, mathematics, and biological sciences—and in business/communications and history. Three course categories with at least 25 grades for which they were underpredicted were remedial reading/literature, remedial English, and advanced physical sciences/engineering.

Hispanic students were overpredicted in biological and physical sciences, economics, studio art/music/theater, and writing courses. They were underpredicted in foreign language courses, especially beginning courses.

Because white students were typically in the majority, the amount of under- or overprediction was small. Nevertheless, there was a pattern of underprediction in English and overprediction in mathematics and technical/vocational courses.

The differences in grading in more strictly graded quantitative and science courses and departments and in more leniently graded nonquantitative courses and departments may have adverse effects on career choice for subgroups of students. Although in both types of courses students with lower test scores and HSR tend to receive grades that are below average for the courses, in quantitative and science courses they receive very low grades. Sabot and Wakeman-Linn (1991) discussed the consequences of grading: low grades result in disappointment, restrictions on participation in sports, academic probation, and parental disapproval; high grades yield intrinsic satisfaction, dean's list, good jobs, and graduate scholarships. Grades are a feedback mechanism that help a student define personal comparative advantage and choose courses on that basis.

Sabot and Wakeman-Linn found that after controlling for other factors, the probability of taking a second course in an academic discipline depended on the grade received in the first course. In economics at Williams College, receiving a B instead of an A reduced the probability of taking a second economics course by 18 percent; a C instead of an A reduced the probability by 28 percent. If the introductory mathematics course at Williams were to adopt the grading distribution of the introductory English course, at least 80 percent more students would take a second mathematics course; if the introductory English course were to adopt the grading distribution of the introductory mathematics course, there would be about a 50 percent decline in the number taking a second English course.

A more uniform grading policy might persuade more students to study science and engineering. As things stand now, students with relatively low test scores and HSR tend to avoid the sciences (Elliott and Strenta 1988) or to migrate out after presumably unrewarding experiences. Willingham (1985, p. 129) reported that students with a low FGPA were three times as likely to migrate to a leniently graded major than were students with a high FGPA. As an even worse consequence, Schurr, Ellen, and Ruble (1987) reported that freshmen who take more strictly graded courses, and thereby receive a lower FGPA, tend to continue to receive relatively low grades in courses beyond their freshman year and are more likely to drop out of college than are freshmen with comparable academic credentials who take more leniently graded courses.

Table 3 indicated that American Indian, black, and Hispanic students, on average, were at a competitive disadvantage in the courses they selected—with lower test scores, HSR, and predicted GPA—and received lower grades than the other students in these courses. Table 7 indicated that their grades on average were substantially lower than were predicted for them. Table 24 showed that the overpredictions were primarily in quantitative and science courses. Because American Indian, black, and Hispanic students tended to receive below-average grades in science and quantitative courses, in which even the average students receive low grades, as Maple and Stage (1991) reported, students in these groups abandon science majors at greater rates than other students. Lack of success in college mathematics and other quantitative courses may be the reason. The result may be artificially forced career choices.

[32]In predicting FGPA for HSGPA and SAT, there was an average overprediction of .16 for black students, which is about one-quarter of a standard deviation. In earlier studies, Ramist (1984, p. 159) reported an average overprediction of .25 for black students, about one-third of a standard deviation. Klitgaard (1985, p. 161) reported a white-black difference in FGPA for students with the same academic aptitude of one-third to two-thirds of a standard deviation, which would translate to overprediction for black students of about one-quarter of a standard deviation. Nettles, Thoeny, and Gosman (1986) found that black students at predominantly white colleges had less well-developed study habits, less academic integration into the college, less satisfaction with the college, and more interfering problems; they hypothesized that reduced feelings of discrimination and improvement in financial aid would lead to better performance by black students. Linn (1983a, 1983b, 1990) suggests that overprediction for black students may be a statistical artifact of affirmative action.

[33]In predicting FGPA from HSGPA and SAT scores, there was an average overprediction of .13 for Hispanic students. Pennock-Roman (1990, p. 75) reported an average of approximately the same overprediction (.15) for Hispanic students. Ramist (1984, p. 159) reported an average overprediction of .05 at four colleges.

TABLE 25

Characteristics of More Selective and Less Selective Colleges by Ethnic Group

More Selective (1121+)						*Less Selective (985–)*				
Amer. Ind.	*Asian Amer.*	*Black*	*Hisp.*	*White*		*Amer. Ind.*	*Asian Amer.*	*Black*	*Hisp.*	*White*
NA	17	16	17	34	Number of courses accounting for half of all credits	NA	18	29	12	19
NA	73	149	142	102	*SD* of course SAT means	NA	100	129	152	104
NA	.13	.28	.23	.27	Correlation between course grade mean and SAT mean	NA	.17	.22	.28	.32
NA	2.92	2.38	2.57	2.92	Mean FGPA	NA	2.49	1.98	2.28	2.33
NA	–.05	+.04	+.03	.00	Average grade mean residual (Z)	NA	.00	+.10	+.09	.00
NA	.11	.18	.12	.12	Z increment*	NA	.13	.17	.14	.13
NA	7.5	7.6	7.2	8.2	Average number of courses	NA	8.3	7.9	7.8	8.5
NA	8%	6%	8%	9%	% of courses advanced	NA	3%	1%	3%	2%
NA	2%	3%	4%	1%	% of courses remedial	NA	3%	8%	3%	4%
					PROPORTIONAL CONTRIBUTIONS TO PREDICT GRADES IN STUDENT-SELECTED COURSES:					
NA	22%	28%	25%	28%	SAT-V	NA	22%	28%	24%	26%
NA	33%	29%	31%	29%	SAT-M	NA	28%	26%	25%	25%
NA	44%	42%	44%	42%	HSGPA	NA	49%	45%	50%	49%
					CORRELATIONS USING SAT AND HSGPA:**					
NA	.67	.54	.54	.69	For FGPA	NA	.54	.51	.49	.59
NA	.63	.50	.49	.60	For course grade	NA	.56	.48	.57	.53
NA	.04	.04	.05	.09	Difference	NA	–.02	.03	–.08	.06
					CORRELATIONS USING SAT:**					
NA	.58	.46	.35	.58	For FGPA	NA	.45	.43	.33	.49
NA	.53	.43	.38	.51	For course grade	NA	.47	.43	.47	.45
NA	.05	.03	–.03	.07	Difference	NA	–.02	.00	–.14	.04
					SAT-V + M MEAN:					
NA	1199	1038	1079	1236	Group	NA	896	769	860	926
NA	1202	1182	1170	1222	Courses selected	NA	943	879	917	917
NA	–3	–144	–91	+14	Difference	NA	–47	–110	–57	+9
					HSGPA MEAN:					
NA	3.73	3.35	3.55	3.64	Group	NA	3.18	2.90	3.08	3.08
NA	3.68	3.61	3.65	3.63	Courses selected	NA	3.16	3.03	3.09	3.07
NA	+.05	–.26	–.10	+.01	Difference	NA	+.02	–.13	–.01	+.01
					COURSE GRADE CORRELATIONS:***					
NA	.50	.39	.38	.53	SAT-Verbal	NA	.44	.43	.46	.47
NA	.60	.46	.38	.55	SAT-Mathematical	NA	.52	.42	.47	.48
NA	.65	.53	.47	.63	SAT (V, M)	NA	.58	.53	.58	.55
NA	.63	.42	.46	.59	HSGPA	NA	.52	.36	.57	.53
NA	.14	.20	.14	.15	SAT Increment	NA	.17	.23	.13	.12
NA	.77	.62	.60	.74	VMH	NA	.69	.59	.70	.65
					OVER-(-)/UNDERPREDICTIONS (+) FOR COURSE GRADE:					
NA	+.04	–.30	–.19	+.03	HSGPA	NA	+.05	–.23	–.09	+.01
NA	+.07	–.19	–.10	+.01	SAT	NA	+.14	–.07	+.01	.00
NA	+.04	–.10	–.06	.00	HSGPA, SAT	NA	+.11	–.05	.00	.00
NA	+.06	–.08	–.06	.00	HSGPA, SAT, TSWE	NA	+.13	–.03	.00	.00

* Z increment = The difference between "the uncorrected correlation of SAT, HSGPA, and average grade mean residual prediction of FGPA" and "the uncorrected correlation of SAT, HSGPA prediction of FGPA."
**Correlations corrected for restriction of range
***Correlations corrected for restriction of range and criter ion unreliability

Differences Among Colleges

Students at more selective and at less selective colleges are compared by ethnic group in Table 25. Because there were only 41 American Indian students at the more selective colleges and 17 at the less selective, data on American Indian students were excluded. The ethnic group distributions are shown in Chart 3.

CHART 3

	More Selective	*Less Selective*
American Indian	41	17
Asian American	1,185	268
Black	576	395
Hispanic	317	134
White	8,467	7,438

In general, the comparisons of ethnic groups at more and less selective colleges[34] were similar to the comparisons over all colleges. But there were some exceptions. At more selective colleges, black students chose courses in which other students had a higher SAT mean by 12 points than did Hispanic students (even though the SAT mean of black students was 41 points lower than that of Hispanic students), thus putting themselves at a greater competitive disadvantage and receiving low grades (the only group closer to a C than a B average). Also, at more selective colleges, the correlations of SAT scores with FGPA and course grade were higher for black students than for Hispanic students. At less selective colleges, the correlations of HSGPA and SAT scores with course grade were highest for the 134 Hispanic students. Also, course grade for Hispanic students was neither over- nor underpredicted at these colleges.

Summary of Student Group Differences

Course Selection and Grading

Academic Composite

Students in the low academic composite group selected courses that were more leniently graded. Course grades for this group were not comparable, from course to course or student to student, especially at less selective colleges, where the validity using SAT scores and HSGPA in predicting FGPA was much lower than the validity in predicting one single course grade. Associated with lack of comparability of course grades was a negative correlation between course grade means and course SAT means. In addition to SAT scores and HSGPA, use of the average grade mean residual of the courses selected as an additional predictor added .10 to the correlation with FGPA, .28 for students in the low composite group at less selective colleges.

Sex

Males selected more quantitative courses with higher predictive weight on the SAT-Mathematical score than on the SAT-Verbal score, more strictly graded courses, and more courses in which other students had higher than average HSGPA and SAT scores than did females. This course selection pattern accounted for .06 of the .09 higher FGPA for females. On average, in terms of predicted FGPA based on HSGPA and SAT scores, neither males nor females had any competitive advantage or disadvantage in their courses; males had higher SAT scores and females had higher HSGPA.

English as Best or Not Best Language

Like males, students whose best language was not English tended to select quantitative, strictly graded, competitive courses, especially physical sciences/engineering and mathematics at the calculus level or higher. Despite higher SAT-Mathematical scores and HSGPA, their lower SAT-Verbal and TSWE scores, combined with their tough course selection, put them at a competitive disadvantage in their courses in terms of predicted FGPA. Nevertheless, they overcame this disadvantage to achieve a higher FGPA than did students whose best language was English.

Ethnic Group

American Indian students selected more courses in nonstudio art/music/theater and in mathematics below the calculus level. They had relatively low test scores and HSGPA, and were at a competitive disadvantage in their courses.

Like males and students whose best language was not English, *Asian American* students tended to select quantitative, strictly graded, competitive courses, especially in physical sciences/engineering and mathematics at the calculus level or higher: with prediction equations having more weight on the SAT-Mathematical score, with grading resulting in a negative average grade mean residual, and with other students who had higher than average test scores and HSGPA. While their high SAT-Mathematical scores took away any competitive disadvantage in their courses (they had no advantage either), their selection of

[34]Large and small colleges were not compared by ethnic group because of insufficient numbers of students at the small colleges.

tough courses made it more difficult for them to obtain higher grades, giving them a potential liability in FGPA. But they were able to overcome this liability to achieve a very high FGPA.

Black students selected courses for which there was a high proportion of predictive weight on the SAT-Verbal score, especially social sciences/humanities and English courses. The courses tended to be nonquantitative and leniently graded, selected by students who had lower than average test scores and HSGPA. Nevertheless, black students were at a competitive disadvantage in their courses and obtained low grades. At more selective colleges, black students chose relatively more difficult and competitive courses than they did at less selective colleges, putting themselves at an even greater competitive disadvantage and receiving low grades. Their course grades were also quite lacking in comparability, and there was a very high standard deviation of course SAT means.

Hispanic students selected more biological sciences courses. They had relatively low test scores and HSGPA, and were at a competitive disadvantage in their courses. Their grades were the least comparable of all the ethnic groups, with better prediction for a single course grade than for the eight-course FGPA. There was a very high standard deviation of course SAT means.

White students took more business/communications courses and fewer biological sciences courses. With high test scores and HSGPA, they were at a competitive advantage in their courses. Their course grades tended to be more comparable, as indicated by a relatively high correlation between course SAT mean and course grade mean.

Predictive Effectiveness

Academic Composite

Students in the low academic composite group had lower correlations with FGPA and course grade for the SAT and HSGPA, especially the latter, but had a very high SAT increment to the correlations over HSGPA. Also, because of lack of comparability of grades, the average grade mean residual provided a large increment to the correlation with FGPA based on the SAT and HSGPA for these students. Students in the high academic composite group had higher correlations with FGPA and course grade for the SAT and HSGPA, and higher correlations between predicted FGPA (based on SAT and/or HSGPA prediction of the courses selected by the student) and actual FGPA, but lower SAT increments to the correlations. Students in the high composite group at more selective colleges had especially high correlations, but also a very high, not low, SAT increment.

Sex

In general, predictive effectiveness of both the SAT and HSGPA, for FGPA and course grade, was higher for females than for males. Females also had a higher SAT increment over HSGPA for both FGPA and course grade. Correlations with course grade for males and females and the correlation with FGPA for females were higher for the SAT than for HSGPA; the correlation with FGPA for males was higher for HSGPA than for the SAT. But there were virtually no sex differences in predictive effectiveness in the high academic composite group or at more selective colleges.

English as Best or Not Best Language

Despite concerns that the SAT may not be appropriate for students whose best language was not English,[35] the SAT increment over HSGPA was high for these students, especially at large colleges, in which the correlation of HSGPA for predicting course grade was especially low.

Ethnic Group

Test score correlations were lowest for *American Indian* students, especially the SAT-Mathematical score; this was the only group for which the verbal score was a better predictor than the mathematical score.

The best prediction was for *Asian American* students, especially for the SAT-Mathematical score; but the verbal score was less effective as a predictor at less selective colleges.

Despite contentions that the SAT is not a useful predictor for *black* students,[36] the SAT was most important for the prediction of their grades: the SAT increment in correlation over HSGPA for this group (for whom the predictive effectiveness of HSGPA was very low) was by far the largest among all the ethnic groups. Because of the lack of comparability of course grades, the increment in the correlation with FGPA over HSGPA and the SAT for the average grade mean residual was also highest for this group.

Compared to other groups, test score correlations were relatively low for *Hispanic* students. But they were relatively high for those Hispanic students in the high academic composite group or at less selective colleges.

Contrary to popular misconceptions that the SAT is primarily useful as a predictor for white males, not only was the SAT increment over HSGPA lower for males than females, but it was also lower for *white* students than for black students.

[35]See Pennock-Roman (1990, p. 12), Hsia (1988), and Sue and Abe (1988).

[36]Crouse and Trusheim (1988).

Over- and Underpredictions

Academic Composite

Because of the nature of the definition of the academic composite groups, using HSGPA or SAT scores separately resulted in underpredictions for the high composite group and overpredictions for the low composite group, but using HSGPA with SAT scores eliminated almost all under- and overpredictions for all composite groups.

Sex

In general, FGPA and course grades were underpredicted for females and overpredicted for males. Course selection, on average, gave females a .03 "bonus" in FGPA and gave males a .03 "liability" in FGPA. Using course grade instead of FGPA eliminated this difference of .06 from the .09 difference in FGPA favoring females. The same could be accomplished with FGPA as the criterion if average grade mean residual were to be used as an additional predictor with HSGPA and SAT scores. With course grade as the criterion, females were, on average, overpredicted by .01 using HSGPA; underpredicted by .06 using the SAT; underpredicted by .03 using HSGPA and the SAT; and underpredicted by .02 using HSGPA, the SAT, and TSWE. The largest underpredictions for females were in English and foreign language courses. Females were on average overpredicted in technical courses other than mathematics: technical/vocational, physical sciences/engineering, economics, and computer science. In more selective colleges, using HSGPA, SAT scores, and TSWE there were no overall average under- or overpredictions for females or males. In less selective colleges, underpredictions for females and over-predictions for males were greater because of a very low FGPA for males.

English as Best or Not Best Language

On average, students whose best language was not English were underpredicted, especially in quantitative courses. They were, however, overpredicted in English courses.

Ethnic Group

American Indian students were overpredicted in a variety of science, language, English, and mathematics courses.

Asian American students were underpredicted, especially in mathematics and science. They were overpredicted in English, architecture/environmental design, and physical education.

Black students were overpredicted, especially in quantitative and science courses.

In general, *Hispanic* students were overpredicted, but not at less selective colleges.

For *white* students, there was a small, discernible pattern of underprediction in English and overprediction in mathematics and technical/vocational courses.

The combination of strict grading in quantitative and science courses, low predicted grades in these courses for American Indian, black, and Hispanic students, and lower grades received by these students than predicted may deter students in these ethnic groups from majoring in quantitative and science disciplines. In effect, receiving such low grades in these courses may amount to artificially forced career choices for these students.

References

American College Testing Program. 1973. *Assessing Students on the Way to College. Vol. 1, Technical Report for the ACT Assessment Program.* Iowa City, Ia.: American College Testing Program.

Breland, Hunter M. 1979. *Population Validity and College Entrance Measures.* Research Monograph No. 8. New York: College Entrance Examination Board.

Breland, Hunter M., and Philip A. Griswold. 1982. "Use of a Performance Test as a Criterion in a Differential Validity Study." *Journal of Educational Psychology* 74:713–21.

Bridgeman, Brent, and Cathy Wendler. 1991. "Gender Difference in Predictors of College Mathematics Performance and in College Mathematics Course Grades." *Journal of Educational Psychology* 83:275–84.

Bridgeman, Brent, and Cathy Wendler. 1989. *Prediction of Grades in College Mathematics Courses as a Component of the Placement Validity of SAT-Mathematics Scores.* College Board Report 89-9. New York: College Entrance Examination Board.

Burton, Nancy W., Charles Lewis, and Nancy Robertson. 1988. *Sex Differences in SAT Scores.* College Board Report 88-9. New York: College Entrance Examination Board.

Burton, Nancy W., Rick Morgan, Charles Lewis, and Nancy J. Robertson. 1989. "The Predictive Validity of SAT and TSWE Item Types for Ethnic and Gender Groups." Paper presented at AERA annual meeting, San Francisco, Calif.

Caldwell, Edward, and Rodney Hartnett. 1967. "Sex Bias in College Grading?" *Journal of Educational Measurement* 4:129–32 .

Clark, M. J., and J. Grandy. 1984. *Sex Differences in the Academic Performance of Scholastic Aptitude Test Takers.* College Board Report No. 84-8 and ETS Research Report

No. 84-43. New York: College Entrance Examination Board.

College Board. 1988. *Guide to the College Board Validity Study Service*. New York: College Entrance Examination Board.

Cronbach, Lee J. 1971. "Test Validation." In *Educational Measurement*, 2nd edition, ed. Robert L. Thorndike. Washington, D.C.: American Council on Education.

Crouse, James, and D. Trusheim. 1988. *The Case Against the SAT*. Chicago: University of Chicago Press.

Donlon, Thomas F., Ed. 1984. *The College Board Technical Handbook for the Scholastic Aptitude Test and Achievement Tests*. New York: College Entrance Examination Board.

Doolittle, Allen E. 1987. *Gender Differences in Performance on Mathematics Achievement Items*. ACT Research Report 87-16. Iowa City, Ia.: American College Testing Program.

Doolittle, Allen E., and Catherine Welch. 1989. *Gender Differences in Performance on a College-Level Achievement Test*. ACT Research Report 89-9. Iowa City, Ia.: American College Testing Program.

Durán, Richard P. 1983. *Hispanics' Education and Background: Predictors of College Achievement*. New York: College Entrance Examination Board.

Elliott, Rogers, and A. C. Strenta. 1988. "Effects of Improving the Reliability of the GPA on Prediction Generally and on Comparative Predictions for Gender and Race Particularly." *Journal of Educational Measurement* 25:333–47.

Gamache, Leann M., and Melvin R. Novick. 1985. "Choice of Variables and Gender Differentiated Prediction Within Selected Academic Programs." *Journal of Educational Measurement* 22:53–70.

Goldman, Roy D., and Barbara Newlin Hewitt. 1975. "Adaptation-Level as an Explanation for Differential Standards in College Grading." *Journal of Educational Measurement* 12:149–61.

Goldman, Roy D., Donald E. Schmidt, Barbara Newlin Hewitt, and Ronald Fisher. 1974. "Grading Practices in Different Major Fields." *American Educational Research Journal* 11:343–57.

Goldman, Roy D., and Robert E. Slaughter. 1976. "Why College Grade Point Average Is Difficult to Predict." *Journal of Educational Psychology* 68:9–14.

Goldman, Roy D., and Mel H. Widawski. 1976. "A Within Subjects Technique for Comparing Grading Standards: Importance in the Validity of the Evaluation of College Achievement." *Educational and Psychological Measurement* 36:381–90.

Gulliksen, Harold. 1950. *Theory of Mental Tests*. New York: John Wiley and Sons.

Hsia, Jayjia. 1988. *Asian Americans in Higher Education and at Work*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Klitgaard, Robert. 1985. *Choosing Elites*. New York: Basic Books.

Lawrence, Ida M., W. Edward Curley, and Frederick J. McHale. 1988. *Differential Item Functioning for Males and Females on SAT-Verbal Reading Subscore Items*. College Board Report 88-4. New York: College Entrance Examination Board.

Linn, Robert L. 1973. "Fair Test Use in Selection." *Review of Educational Research* 43:139–61.

Linn, Robert L. 1983a. "Pearson Selection Formulas: Implications for Studies of Predictive Bias and Estimates of Educational Effects in Selected Samples." *Journal of Educational Measurement* 20:1–15.

Linn, Robert L. 1983b. "Predictive Bias as an Artifact of Selection Procedures." In *Principles of Modern Psychological Measurement: A Festschrift for Frederic M. Lord*, ed. by Howard Wainer and Samuel Messick, pp. 27–40. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Linn Robert L. 1990. "Admissions Testing: Recommended Uses, Validity, Differential Prediction, and Coaching." *Applied Measurement in Education* 3:297–318.

Manski, Charles F., and David A. Wise. 1983. *College Choice in America*. Cambridge, Mass.: Harvard University Press.

Maple, Sue A., and Frances K. Stage. 1991. "Influences on the Choice of Math/Science Major by Gender and Ethnicity." *American Educational Research Journal* 28:37–60 .

McCornack, Robert L. 1983. "Bias in the Validity of Predicted College Grades in Four Ethnic Minority Groups." *Educational and Psychological Measurement* 43:517–22.

McCornack, Robert L., and M. M. McLeod. 1988. "Gender Bias in the Prediction of College Course Performance." *Journal of Educational Measurement* 25:321–31.

Milton, Ohmer, Howard R. Pollio, and James A. Eison. 1986. *Making Sense of College Grades*. San Francisco, Calif.: Jossey-Bass.

Morgan, Rick. 1990. *Predictive Validity Within Categorizations of College Students: 1978, 1981, and 1985*. ETS Research Report 90-141. Princeton, N.J.: Educational Testing Service.

Nettles, Michael T., Ed. 1988. *Toward Black Undergraduate Student Equality in American Higher Education*. New York: Greenwood Press.

Nettles, Michael T., A. Robert Thoeny, and Erica J. Gosman. 1986. "Comparative and Predictive Analyses of Black and White Students' College Achievement and Experiences." *Journal of Higher Education* 57:289–318.

Pennock-Roman, Maria. 1990. *Test Validity and Language Background: A Study of Hispanic-American Students at Six Universities*. New York: College Entrance Examination Board.

Ramist, Leonard. 1978. "Admissions-Yield and Persistence Analysis." Paper presented at the annual forum of the Association for Institutional Research, Houston, Tex.

Ramist, Leonard. 1984. "Predictive Validity of the ATP tests." In *The College Board Technical Handbook for the Scholastic Aptitude Test and Achievement Tests*, ed. T. F. Donlon, pp. 141–70. New York: College Entrance Examination Board.

Ramist, Leonard, Charles Lewis, and Laura McCamley. 1990. "Implications of Using Freshman GPA as the Criterion for the Predictive Validity of the SAT." In Willingham, Warren W., Charles Lewis, Rick Morgan, and Leonard Ramist, *Predicting College Grades: An Analysis of Institutional Trends Over Two Decades*, pp. 253–88. Princeton, N.J.: Educational Testing Service.

Ramist, Leonard, and Gloria Weiss. 1990. "The Predictive Validity of the SAT, 1964 to 1988." In Willingham, Warren W., Charles Lewis, Rick Morgan, and Leonard Ramist, *Predicting College Grades: An Analysis of Institutional Trends Over Two Decades*, pp. 117–40. Princeton, N.J.: Educational Testing Service.

Sabot, Richard, and John Wakeman-Linn. 1991. "Grade Inflation and Course Choice." *Journal of Economic Perspectives* 5:159–70.

Sawyer, Richard. 1986. "Using Demographic Subgroup and Dummy Variable Equations to Predict College Freshman Grade Average." *Journal of Educational Measurement* 23:131–45.

Schurr, K. Terry, Arthur S. Ellen, and Virgil E. Ruble. 1987. "Actual Course Difficulty as a Factor in Accounting for the Achievement and Attrition of College Students." *Educational and Psychological Measurement* 47:1049–54.

Strenta, A. Christopher, and Rogers Elliott. 1987. "Differential Grading Standards Revisited." *Journal of Educational Measurement* 24:281–91.

Stricker, Lawrence J., Donald A. Rock, and Nancy W. Burton. 1991. *Sex Differences in SAT Predictions of College Grades*. College Board Report 91-2. New York: College Entrance Examination Board.

Sue, Stanley, and Jennifer Abe. 1988. *Predictors of Academic Achievement among Asian American and White Students*. College Board Report No. 88-11. New York: College Entrance Examination Board.

Ware, Norma C., and Valerie E. Lee. 1988. "Sex Differences in Choice of College Science Majors." *American Educational Research Journal* 25:593–614.

Wild, Cheryl L. 1977. "Statistical Issues Raised by Title IX Requirements on Admission Procedures." *Journal of the National Association of Women Deans, Administrators, and Counselors* 40 (Winter): 53–6.

Wilder, Gita Z., and K. Powell. 1989. *Sex Differences in Test Performance: A Survey of the Literature*. College Board Report No. 89-3 and ETS Research Report 89-4. New York: College Entrance Examination Board.

Willingham, Warren W. 1985. *Success in College: The Role of Personal Qualities and Academic Ability*. New York: College Entrance Examination Board.

Willingham, Warren W., and Hunter M. Breland. 1982. *Personal Qualities and College Admissions*. New York: College Entrance Examination Board.

Willingham, Warren W., Charles Lewis, Rick Morgan, and Leonard Ramist. 1990. *Predicting College Grades: An Analysis of Institutional Trends Over Two Decades*. Princeton, N.J.: Educational Testing Service.

Young, John W. 1990. "Are Validity Coefficients Understated Due to Correctable Defects in the GPA?" *Research in Higher Education* 31:319–25.

Young, John W. 1991. "Gender Bias in Predicting College Academic Performance: A New Approach Using Item Response Theory." *Journal of Educational Measurement* 28:37–47.

Appendix A

Colleges Participating in the Study

Total Students	English Not Best Language	College	Ethnic Group: American Indian	Asian American	Black	Hispanic	White	Other/ Missing
1,163	16	Arizona State University	11	41	25	48	1,007	31
1,027	7	Auburn University	3	7	19	7	981	10
355	9	Augusta College*	3	10	44	1	286	11
1,543	44	Boston College	5	67	39	25	1,350	57
572	4	Bryant College	0	7	3	2	551	9
679	6	Bucknell University	1	13	11	6	638	10
764	19	California State University, Sacramento	5	81	34	41	569	34
414	3	Carleton College	3	19	7	9	364	12
322	1	Colby College	0	2	8	0	298	14
537	10	Colgate University	3	28	10	3	465	28
533	10	Columbia University*	0	54	43	19	373	44
679	6	Dartmouth College*	13	30	47	4	567	18
485	5	Dickinson College	1	9	4	4	455	12
1,175	13	Duke University	3	60	68	6	995	43
400	8	Franklin and Marshall College*	0	12	15	2	360	11
782	23	George Washington University	2	39	13	7	691	30
1,161	24	Harvard University	4	138	82	51	844	42
546	7	Kutztown University*	1	3	16	3	511	12
627	12	La Salle University	1	24	20	6	573	15
823	20	Lehigh University	0	21	8	8	764	22
817	23	Marquette University	1	19	11	27	745	14
599	4	Mary Washington College	2	7	9	1	567	13
208	3	Marywood College	1	0	1	3	200	3
361	3	Mount Holyoke College	2	9	25	3	312	10
297	5	New Hampshire College	2	3	3	1	281	7
2,091	26	Ohio State University*	8	69	73	12	1,886	43
381	4	St. Michael's College	1	1	3	1	363	12
427	5	Slippery Rock University	0	2	6	1	404	14
430	16	Suffolk University	2	9	13	5	383	18
297	3	Susquehanna University	1	0	3	1	288	4
254	4	Swarthmore College	0	7	18	2	222	5
2,330	133	University of California, Berkeley	10	702	182	188	1,111	137
3,429	162	University of California, Los Angeles	24	773	304	350	1,758	220
498	13	University of Central Florida	0	16	12	11	445	14
1,124	36	University of Maryland, Baltimore County*	4	89	152	9	851	29
3,599	124	University of Maryland, College Park	7	304	393	37	2,727	131
2,518	24	University of North Carolina, Chapel Hill	14	43	222	9	2,204	26
2,261	104	University of Southern California	9	428	113	119	1,476	116
5,358	133	University of Texas, Austin	16	333	258	535	4,086	130
1,643	41	University of Washington	9	283	29	12	1,270	40
907	8	Vanderbilt University	2	14	19	4	855	13
407	13	Wellesley College	0	46	24	7	307	23
481	6	Wesleyan University	0	24	30	5	402	20
802	15	Western Carolina University	10	4	54	2	710	22
273	2	Whitman College	0	20	2	2	248	1
46,379	1,156	TOTAL	184	3,848	2,475	1,599	36,743	1,530

*Supplied data for 1985 only.

Appendix B

Course Categories

1. Advanced mathematics (Postcalculus)
2. Calculus
3. Precalculus
4. Remedial mathematics
5. Regular mathematics (other than 1–4)

6. English-advanced
7. English-regular
8. English-remedial

9. Reading/literature-advanced
10. Reading/literature-regular
11. Reading/literature-remedial

12. Writing/composition-advanced
13. Writing/composition-regular
14. Writing/composition-remedial

15. Biological sciences-advanced
16. Biological sciences-introductory with laboratory or for majors
17. Biological sciences-introductory with no laboratory and for nonmajors

18. Physical sciences/engineering-advanced
19. Physical sciences/engineering-introductory with laboratory or for majors
20. Physical sciences/engineering-introductory with no laboratory and for nonmajors

21. Foreign languages-beyond entry level
22. Foreign languages-entry level

23. History

24. Social sciences/humanities–political science, sociology, psychology, philosophy, religion, anthropology, archaeology, geography, law, criminal justice, social work, library science, public affairs, area studies, ethnic studies

25. Economics
26. Business/communications
27. Art/music/theater–studio
28. Art/music/theater–nonstudio
29. Computer science
30. Health/nursing
31. Education
32. Physical education
33. Military science
34. Home economics
35. Architecture/environmental design
36. Technical/vocational
37. Other