RESEARCH REPORT 2012-1

CollegeBoard

The Validity of SAT® Scores in Predicting First-Year Mathematics and English Grades

By Krista D. Mattern, Brian F. Patterson, and Jennifer L. Kobrin

Krista D. Mattern is an associate research scientist at the College Board.

Brian F. Patterson is an assistant research scientist at the College Board.

Jennifer L. Kobrin is a research scientist at the College Board.

Mission Statement

The College Board's mission is to connect students to college success and opportunity. We are a not-for-profit membership organization committed to excellence and equity in education.

About the College Board

The College Board is a mission-driven not-for-profit organization that connects students to college success and opportunity. Founded in 1900, the College Board was created to expand access to higher education. Today, the membership association is made up of over 6,000 of the world's leading educational institutions and is dedicated to promoting excellence and equity in education. Each year, the College Board helps more than seven million students prepare for a successful transition to college through programs and services in college readiness and college success — including the SAT® and the Advanced Placement Program®. The organization also serves the education community through research and advocacy on behalf of students, educators and schools.

For further information, visit www.collegeboard.org.

Visit the College Board on the Web: www.collegeboard.org.

For more information on College Board research and data, visit www.collegeboard.org/research.

VALIDITY

Contents

Figures

Executive Summary

This study examined the validity of the SAT® for predicting performance in first-year English and mathematics courses. Results reveal a significant positive relationship between SAT scores and course grades, with slightly higher correlations for mathematics courses compared to English courses. Correlations were estimated by student characteristics (gender, ethnicity, and best language), institutional characteristics (size, selectivity, and control, i.e., private or public), and course content (e.g., calculus, algebra). The findings suggest that performance on the SAT is predictive of performance in specific college courses. Furthermore, stronger relationships were found between test scores and grades when the content of the two were aligned (such as the SAT mathematics section and mathematics course grades, or the SAT writing section and English course grades).

Introduction

The main purpose of the SAT is to measure individuals' potential for academic success in postsecondary institutions. As such, the SAT, in conjunction with other measures, has been used in college admission to determine which applicants would likely succeed at a particular institution. According to the *Standards for Educational and Psychological Testing* (American Educational Research Association/American Psychological Association/National Council on Measurement in Education, 1999), the College Board is "responsible for furnishing relevant evidence and a rationale in support of the intended test use" (p. 11). In order to validate this use, hundreds of studies — conducted by College Board researchers as well as independent researchers — have examined the predictive validity of the SAT in terms of first-year college grade point average (FYGPA). The results of these studies were summarized in a meta-analysis by Hezlett, Kuncel, Vey, Ahart, Ones, Campbell, and Camara (2001), which found a strong relationship between SAT scores and FYGPA. Recently, Kobrin, Patterson, Shaw, Mattern, and Barbuti (2008) examined the predictive validity of the most recent version of the SAT, which includes a writing section, in terms of FYGPA. They found that after correcting for restriction of range in the predictor variables, the multiple correlation for the three sections of the SAT with FYGPA was .53 (uncorrected r = .35). Validity research on the SAT, and on all tests, should be continually conducted to understand how performance on one test relates to the relevant outcomes. Such data comprise evidence that is necessary in order to ensure the proper use of scores.

> … when validating the SAT for use in college admission, FYGPA should not be the only factor in determining college success.

Additionally, when validating the SAT for use in college admission, FYGPA should not be the only factor in determining college success. This is particularly true because students have much autonomy in the type and difficulty of courses in which they enroll, and the correlations are likely to differ for courses of varying content and difficulty. Individual course grades, which are not contaminated by the construct-irrelevant variance of course-taking patterns associated with FYGPA, should be examined in order to make more accurate and precise inferences about the extent to which the SAT is related to the breadth of courses taken.

Relationship Between SAT® Scores and Course Grades

Though most studies investigating the predictive validity of the SAT have focused on FYGPA as the academic outcome of interest, there has been some research examining the relationship between SAT scores and specific course grades. For example, a study by Ramist, Lewis, and McCamley-Jenkins (1994) analyzed student-level data for 38 colleges and universities for the incoming freshman classes of 1982 and 1985, and examined the relationship between SAT scores and course grades in 37 categories (such as advanced mathematics and remedial English literature). After correcting for restriction of range in the predictor variables, with the national cohort of SAT test-takers serving as the population, the average correlation across course categories was .49, compared to .53 for FYGPA. This value increased to .60 when correlations were corrected for attenuation (criterion unreliability) in addition to restriction of range.

As for specific course categories, Ramist et al. (1994) found that correlations were highest for science and quantitative courses, and that the strongest relationship was between biological science course grades and SAT scores, with correlations ranging from .58 to .61. The lowest correlations were for nonquantitative courses such as physical education, which had the weakest relationship with SAT scores, with a correlation of .21. Not surprisingly, the SAT mathematics section (SAT-M) was a better predictor of mathematics and science courses, with corrected correlations ranging from .43 to .55 across content areas, than the SAT verbal[1] section (SAT-V), which had corrected correlations ranging from .26 to .43. Additionally, SAT-V was a slightly better predictor of English and history courses, with corrected correlations ranging from .43 to .48, than SAT-M, which had corrected correlations ranging from .37 to .42. The results indicate that there was a medium to large effect[2] (Cohen, 1988) of SAT scores on course grades and that the magnitude of the relationship varies with the alignment of test and course content.

Finally, Ramist et al. (1994) conducted additional analyses by gender, ethnicity, and best language. They found that SAT scores were related to course grades for all demographic categories, though they observed a stronger correlation for females, Asian American, and white students, as well as students whose best language was English. The pattern of differential validity parallels what these authors found for FYGPA; however, the magnitude of differences tended to be smaller for course grades.

A more focused investigation of the validity of SAT mathematics scores for predicting college grades in mathematics courses was conducted by Bridgeman and Wendler (1989). Based on course-level data from 10 colleges and universities, uncorrected correlations between SAT-M scores and grades in algebra, precalculus, and calculus courses were computed. Unlike the conclusions from the Ramist et al. (1994) study, the findings suggested that SAT-M scores were only moderately related (mid-.30s) to college mathematics grades. The authors state that this is due to the disconnect between the content of the prior version of the SAT, which only required computational knowledge up to algebra, compared to the content of specific first-year mathematics courses, which often require higher levels of knowledge in areas such as precalculus and calculus. However, recent revisions to the mathematics section of the SAT — particularly the inclusion of third-year mathematics content, such as exponential growth, absolute value, functional notation, negative and fractional exponents, and the removal of quantitative comparisons — may result in different conclusions (see Lawrence, Rigol, Van Essen, and Jackson, [2003] for a more detailed description of revisions made to the SAT). Furthermore, as mentioned previously, Bridgeman and Wendler (1989) provided uncorrected correlations, whereas Ramist et al. (1994) corrected correlations for range restriction, which explains some of the differences in results between the two studies.

In preparation for the planned changes to the 1994 version of the SAT, Bridgeman, Hale, Lewis, Pollack, and Wang (1992) examined the validity of a prototype SAT verbal section, which emphasized more passage-based reading questions, and an SAT Subject Test in

1 In March 2005, substantial revisions were made to the SAT, most notably the addition of a writing section. The verbal section was renamed the critical reading section, and the changes to that section included the elimination of analogies and the addition of shorter reading passages.

2 Cohen's rule of thumb (1988) provides general guidelines in terms of what effect size is considered "small," "medium," and "large." In terms of correlations (*r*), these values are .10, .30, and .50, respectively.

Writing, consisting of an essay and multiple-choice items. Specifically, the study examined the validity of the proposed SAT verbal section and the new SAT Subject Test in Writing compared to that of the current verbal section in terms of English course grades. Participating colleges and universities administered the current and prototype SAT verbal section along with the SAT Subject Test in Writing to incoming freshmen and provided their subsequent English course grades. The results indicated that performance on the prototype Writing Subject Test and verbal section was positively correlated to course grades; however, Writing Subject Test scores (uncorrected correlations in the low .30s) were more strongly correlated to performance in English courses compared to verbal section scores (correlations in the high .10s to low .20s). Furthermore, the Writing Subject Test resulted in less underprediction for females than did the verbal section. Whether these results would replicate when students take these examinations in a high-stakes situation needs to be investigated.

Because substantial revisions to the SAT, most notably the addition of the writing section, were made in 2005, the relationship between SAT scores and course performance should be re-evaluated. Furthermore, it needs to be assessed whether the results from Bridgeman et al. (1992) would generalize to the current version of the SAT using data on a national sample from a live administration. This study addresses these research questions.

Method

Sample

The current study used data collected for the national SAT Validity Study, which included college performance data on 196,364 first-year, first-time students from 110 institutions in the entering cohort of 2006. The sample of participating institutions was diverse with regard to U.S. region, selectivity, size, and control (see Kobrin et al. [2008] for details). Data from each participating institution included students' course work and grades, FYGPA, and whether they had returned for the second year. These data were matched to College Board databases that included SAT scores, self-reported high school grade point average (HSGPA), and other self-reported demographic information.

Students in the sample who did not have SAT scores, did not report their HSGPA, or did not complete a mathematics or English course in their first year of college were excluded from the analyses. To identify first-year mathematics and English courses, participating institutions were asked to complete a short online questionnaire that specifically asked for a list of all courses that were considered first-year mathematics and English courses at their institution. Additionally, all courses in the database were coded by an independent firm into 14 higher-level categories, including mathematics and English categories. Having two independent

Because substantial revisions to the SAT, most notably the addition of the writing section, were made in 2005, the relationship between SAT scores and course performance should be re-evaluated.

course coding schemes allowed for a check of the accuracy of the survey responses[3] and also facilitated the addition of courses that had been omitted from the survey responses but had appeared in the database. Furthermore, it allowed for the identification and removal of courses that institutions had listed but that fell outside the content domain. For example, one institution included a college physics course because it fulfilled a student's quantitative reasoning requirement; however, the authors felt that its content was only tangentially related to mathematics and hence excluded it from the mathematics analyses.

To examine the validity of SAT scores for predicting mathematics and English grades, the analyses only included a student's first course in mathematics and English. For example, if a student took Calculus I in his or her first semester and Calculus II in his or her second semester, only the grade for Calculus I was included in the analyses. This was done because performance in the Calculus II course would be a function of not only a student's academic preparation (as measured by SAT scores) but also a function of what he or she had learned in the Calculus I course, and we wanted to isolate the relationship between SAT scores and course performance, irrespective of college course-taking behavior. A small percentage of students (fewer than 1% of students in the mathematics analyses and approximately 4% of students in the English analyses) were enrolled in more than one first-year mathematics or English course in the semester in which they took their first course in the respective content area (for example, a student took no English courses in the first semester but took both a literature and composition course in the second semester). Because analyses were conducted at the level of the course, not the institution, both course grades were included in the analyses for these students. Finally, courses that had no variability in grades earned (e.g., all students earned an A) were excluded from analyses. Across the 110 institutions, the final sample included 222 first-year English courses completed by 96,589 students and 378 first-year mathematics courses completed by 70,840 students.

Measures

SAT scores. Official SAT scores obtained from the 2006 College-Bound Seniors cohort database were used in the analyses. This database consists of the students who participated in the SAT program and reported plans to graduate from high school in 2006. The student's most recent score was used in the analyses. The SAT is composed of three sections: critical reading, mathematics, and writing, and the score scale range for each section is 200 to 800.

SAT Questionnaire Responses. Self-reported gender, race/ethnicity, and best language were obtained from the SAT Questionnaire, which students completed during registration for the SAT. Self-reported HSGPA was also obtained from the SAT Questionnaire. HSGPA is on a 12-point scale with the following response options: A+ (97–100), A (93–96), A– (90–92), B+ (87–89), B (83–86), B– (80–82), C+ (77–79), C (73–76), C– (70–72), D+ (67–69), D (65–66), and E or F (below 65). This scale was recoded to a conventional 12-point numeric GPA scale ranging from .00 to 4.33.

Course grades. Participating institutions supplied course-level data. Specifically, the course name/label, course term, number of credits earned, and grade earned for each course in which a student was enrolled during his or her first year of college were provided. The grade scale ranged from .00 to 4.33, with most institutions truncating at 4.0. Score scales varied across institutions, with some schools reporting only integer data (4 for A, 3 for B, etc.), whereas other institutions reported to one or two decimal places (B+ = 3.67). Differences in

3 Of the courses identified by the participating institutions as first-year mathematics and English courses, 94% of the mathematics courses and 92% of the English courses were coded similarly by the independent firm.

grading scales across institutions were not a problem because analyses were conducted at the level of the course within an institution.

Data Analysis

Descriptive Statistics. Means and standard deviations for each of the SAT sections, HSGPA, and mathematics and English course grades were computed for the total sample. Means and standard deviations were also provided by student, institutional, and course subgroups.

Predictive Validity. To assess the predictive validity of the SAT in terms of mathematics and English course grades, within-course correlations were computed between SAT scores and HSGPA with course grades separately for each course. Specifically, raw covariance matrices including HSGPA, the relevant SAT sections, and first-year course grades were estimated for each of the 222 English and 378 mathematics courses, which had at least 10 students, non-missing data on study variables as described above, and some variability in numeric course grade. If any of the covariance matrices were not of full rank (i.e., the matrix was singular), then that course was excluded from the analyses.

All correlations were corrected for restriction of range in the predictor variables at the course level using the Pearson–Lawley multivariate correction, with the 2006 College-Bound Seniors cohort serving as the population (Gulliksen, 1950; Lawley, 1943). At this point, the corrected covariance matrices for each course were averaged, weighting by the number of students included for analysis of that course. To estimate the multiple correlations, the study employed the procedure described in Powers (2004), in which the corrected sample size–weighted average correlation matrix was used to compute multiple correlations. Corrected correlations reported in the tables and raw correlations reported in the appendixes include both bivariate and multiple correlations, all of which are derived from the pooled correlation matrices.

Differential Validity. To examine the extent to which the SAT exhibits differential validity across varying types of institutions for predicting first-year mathematics and English course grades, correlations were averaged to the level of their respective group (for example, English course correlations at private institutions were averaged together). If the correlation coefficient varies by subgroup, then a test is said to exhibit differential validity. For the analyses by specific content areas within English (e.g., literature) and mathematics courses (e.g., algebra), correlations were estimated using a method similar to that in which the correlations by institutional characteristics had been estimated, but they were reported only if there were at least five courses from at least three unique institutions to be analyzed. Courses that failed to meet this minimum requirement were labeled and grouped as “Other Mathematics” or “Other English” for the two content areas.

For differential validity by student characteristics (such as gender, ethnicity/race, and best language), correlations were computed for each subgroup within courses. Courses that had fewer than 10 students of a particular subgroup were excluded from that group’s results. For example, if a course had fewer than 10 females, that course was excluded from the female analyses. However, if the same course had at least 10 males, it was included in the analyses for males. These raw covariance matrices were corrected separately using the Pearson–Lawley multivariate correction for restriction of range. They were then transformed into correlations and, after weighting by the subgroup-by-course sample size, the average of the corrected correlation matrices was treated as the pooled, subgroup-specific correlation matrix. On the basis of this matrix, multiple correlations were computed. As with the overall analyses, all correlations were corrected for restriction of range in the predictor variables at the level of the course using the Pearson–Lawley multivariate correction, with the 2006 College Bound Seniors cohort serving as the population (Gulliksen, 1950; Lawley, 1943).

Differential Prediction. Differential prediction occurs when a test over- or underpredicts the criterion (e.g., English and mathematics course grades) by student subgroups. This is calculated by subtracting the predicted course grade derived from a regression analysis from the earned course grade (i.e., residual = course grade$_{earned}$ – course grade$_{predicted}$). Negative values (residuals) indicate overprediction, and positive values indicate underprediction. For example, if a specific subgroup (such as females) tends to earn higher grades in English courses than is predicted by a regression equation using SAT scores, then the SAT exhibits differential prediction by gender, namely underprediction for females.

To assess the extent to which the SAT, as well as HSGPA, exhibits differential prediction, regression equations within each course were estimated. The average residual of course grades by various student subgroups was computed across the entire sample. Note that the issue of minimum subgroup size described above did not prevent us from doing this analysis, because regression equations were estimated for the entire course and the residuals were simply aggregated by subgroup. In other words, if a course had at least one student from a given subgroup, that student and course were included in the analysis.

Results

Mathematics Courses

Descriptive Statistics. Table 1 provides the means and standard deviations of the study variables for first-year mathematics courses for the total sample; by students' gender, race/ethnicity, and best spoken language; and by institutions' control, size, and selectivity. Of all courses from the 110 institutions, 378 were identified as introductory mathematics courses. The average grade in these courses was a B- (mean = 2.71; SD = 1.18). In terms of students' high school academic performance, the average HSGPA was A- (mean = 3.63; SD = .49), and the average SAT scores were 560 (SD = 93), 586 (SD = 94), and 554 (SD = 91) for SAT-CR, SAT-M, and SAT-W, respectively. Observations concerning performance by demographic characteristics follow.

- Females, on average, earned higher grades in their first-year mathematics courses (mean = 2.79 versus 2.63 for males) and also had slightly higher SAT-W scores and HSGPAs, whereas males tended to have higher SAT-CR and SAT-M scores.
- Asian American and white students earned higher SAT scores and HSGPAs, and also earned higher first-year mathematics course grades, than the other racial/ethnic groups. Hispanic students had similar HSGPAs to white students but earned lower grades in their first-year mathematics courses.
- Students who reported that a language other than English was their best spoken language earned the highest first-year mathematics course grades (mean = 3.08) and the highest SAT-M scores (mean = 611).

Mean performance on the precollege academic indicators also varied by institutional characteristics.

- Students at private institutions tended to have higher SAT scores as well as earn higher grades in their first-year mathematics courses compared to students at public institutions (2.81 versus 2.68, respectively). These groups were comparable in terms of HSGPA.

- Students attending small institutions (750 to 1,999 undergraduates) performed the least well on the SAT and HSGPA, whereas students at large institutions (7,500 to 14,999 undergraduates) earned the lowest grades in their first-year mathematics courses. Students at medium to large institutions (2,000 to 7,499 undergraduates) or very large institutions (15,000 or more undergraduates) performed the best on all four precollege academic measures.

- Students attending institutions that admitted fewer than 50% of their applicants had higher HSGPAs and SAT scores, and earned higher grades, in their first-year mathematics courses compared to students at less selective institutions.

Table 1.

Descriptive Statistics of Study Variables for First-Year Mathematics Courses

			Course Grade		SAT- CR		SAT-M		SAT-W		HSGPA	
Variable		*n*	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD
Gender	Male	34,572	2.63	1.22	566	92	610	92	553	92	3.59	0.51
	Female	36,268	2.79	1.14	554	93	562	90	555	90	3.67	0.47
Race/ Ethnicity	American Indian	376	2.53	1.25	548	82	567	85	530	78	3.56	0.51
	Asian American	6,978	3.00	1.13	572	104	637	95	572	99	3.70	0.46
	Black/African American	5,418	2.26	1.26	510	85	514	84	503	84	3.43	0.54
	Hispanic	5,793	2.50	1.26	527	89	551	91	523	86	3.65	0.49
	White	47,338	2.74	1.16	567	89	591	90	560	88	3.64	0.48
	Other	1,981	2.71	1.15	554	96	576	95	551	93	3.60	0.49
	No Response	2,956	2.76	1.15	583	99	596	97	573	99	3.64	0.51
Best Language	English Only	65,535	2.71	1.18	563	91	586	93	556	90	3.63	0.49
	English & Another Language	3,783	2.73	1.22	534	98	580	107	537	98	3.65	0.48
	Another Language	815	3.08	1.08	468	101	611	112	484	98	3.63	0.50
	No Response	707	2.69	1.16	546	104	570	108	541	104	3.57	0.56
Control	Public	52,567	2.68	1.22	553	91	582	93	546	88	3.63	0.49
	Private	18,273	2.81	1.06	579	97	596	98	575	96	3.65	0.49
Size	Small	2,913	2.64	1.10	539	98	548	97	534	95	3.49	0.55
	Medium to Large	13,180	2.75	1.11	565	97	577	98	560	96	3.57	0.50
	Large	17,372	2.59	1.18	551	94	574	97	544	92	3.55	0.51
	Very Large	37,375	2.76	1.21	564	90	597	90	558	88	3.70	0.46
Selectivity	Under 50%	13,045	2.84	1.07	596	94	618	93	593	93	3.70	0.46
	50% to 75%	40,088	2.78	1.17	560	92	588	95	554	89	3.64	0.49
	Over 75%	17,707	2.46	1.24	532	84	557	85	524	82	3.55	0.51
Total		70,840	2.71	1.18	560	93	586	94	554	91	3.63	0.49

Note: *n* = number of students. With regard to institution size, small = 750 to 1,999 undergraduates; medium to large = 2,000 to 7,499 undergraduates; large = 7,500 to 14,999 undergraduates; and very large = 15,000 or more undergraduates.

Predictive Validity. Table 2 provides the overall correlations, corrected for restriction of range using the 2006 College-Bound Seniors cohort, as well as the correlations by student and institutional characteristics for the precollege academic measures with mathematics course grades (refer to Appendix A for uncorrected correlations). Not surprisingly, of the three SAT sections, SAT-M correlated most highly with mathematics course grades (.52). In fact, SAT-CR and SAT-W scores did not provide any incremental validity over SAT-M scores. SAT-M was also more highly correlated with mathematics course grades than HSGPA, which correlated .48 with mathematics course grades. The multiple correlation of all three SAT sections and HSGPA with mathematics grades was .58. In sum, the corrected correlation between SAT mathematics scores and mathematics course grades is considered a large effect (Cohen, 1988); furthermore, it is higher than the correlation between SAT mathematics scores and FYGPA (.47, as reported in Kobrin et al., 2008).

In Figure 1, the average mathematics course grade and the percentage of students earning a B or higher by SAT mathematics score band are provided to illustrate the positive relationship graphically. Specifically, the mean mathematics course grade increased as the SAT-M score band increased. For example, students in the highest SAT-M score band (700–800) averaged a B+ (3.31) in their first-year mathematics course compared to students in the two lowest score bands (200–390), who averaged less than a C (1.92). Presented another way, less than a third of students in the two lowest SAT-M score bands earned a B, compared to 78.3% of students in the highest score band.

… students in the highest SAT-M score band (700–800) averaged a B+ (3.31) in their first-year mathematics course compared to students in the two lowest score bands (200–390), who averaged less than a C (1.92).

Table 2.
Corrected Correlations of SAT Scores and HSGPA with First-Year Mathematics Course Grades Overall and by Subgroups

Variable		*k*	*n*	SAT-CR	SAT-M	SAT-W	SAT	HSGPA	HSGPA, SAT
Gender	Male	306	34,242	.36	.48	.40	.49	.45	.54
	Female	321	35,953	.43	.57	.46	.57	.48	.62
Race/Ethnicity	American Indian	5	63	.14	.34	.17	.37	.44	.51
	Asian American	137	6,378	.33	.45	.37	.46	.40	.50
	Black/African American	129	4,719	.37	.49	.40	.50	.42	.54
	Hispanic	122	5,120	.29	.42	.34	.42	.36	.46
	White	333	47,101	.41	.52	.44	.53	.51	.60
	Other	65	1,190	.26	.37	.30	.37	.32	.40
	No Response	102	2,120	.32	.40	.35	.41	.38	.46
Best Language	English Only	371	65,490	.41	.52	.45	.53	.49	.59
	English & Another Language	106	3,130	.31	.43	.34	.44	.36	.47
	Another Language	17	373	.23	.33	.28	.34	.30	.37
	No Response	7	114	.26	.37	.38	.43	.36	.47
Control	Public	191	52,567	.39	.51	.43	.52	.47	.57
	Private	187	18,273	.42	.54	.46	.55	.50	.61
Size	Small	50	2,913	.36	.49	.41	.50	.45	.55
	Medium to Large	134	13,180	.41	.53	.45	.54	.51	.60
	Large	77	17,372	.41	.53	.45	.54	.51	.60
	Very Large	117	37,375	.39	.51	.43	.52	.46	.56
Selectivity	Under 50%	98	13,045	.43	.54	.47	.55	.50	.61
	50% to 75%	191	40,088	.38	.51	.42	.52	.46	.57
	Over 75%	89	17,707	.40	.52	.44	.53	.50	.59
Total		378	70,840	.40	.52	.44	.52	.48	.58

Note: *k* = number of courses; *n* = number of students. Pooled within-course correlations are presented. Correlations are corrected for restriction of range using the institution's enrolled first-year class as the population. SAT is the multiple correlation for all three sections. HSGPA, SAT is the multiple correlation of all three SAT sections and HSGPA. With regard to institution size, small = 750 to 1,999 undergraduates; medium to large = 2,000 to 7,499 undergraduates; large = 7,500 to 14,999 undergraduates; and very large = 15,000 or more undergraduates.

Examining differential validity by student characteristics, the following observations may be made (refer to Table 2):

- The correlations between the academic measures and mathematics course grades were consistently higher for females, ranging from .03 higher for HSGPA to .09 higher for SAT-M. The correlation between SAT-M scores and mathematics course grades was higher than the correlation between HSGPA and mathematics course grades for both females and males.

- The correlation between SAT-M and mathematics grades was highest for Asian American, African American, and white students. The correlation was lowest for American Indian students; however, the results for this group were based on a small number of courses (*k* = 5) and students (*n* = 63), and should be interpreted with caution. Compared to HSGPA, SAT-M is more strongly related to mathematics course grades for all racial/ethnic groups except for American Indian students.

- The correlation between SAT-M and mathematics grades was highest for students who reported that their best language was English, followed by students who reported that their best language was English and another language, and lowest for students who reported that their best language was another language. Interestingly, students whose best language was another language had the highest mean SAT-M score but also had the highest standard deviation, indicating that group members varied greatly in their mathematics ability. Compared to HSGPA, SAT-M is more strongly related to mathematics course grades for all best language groups.

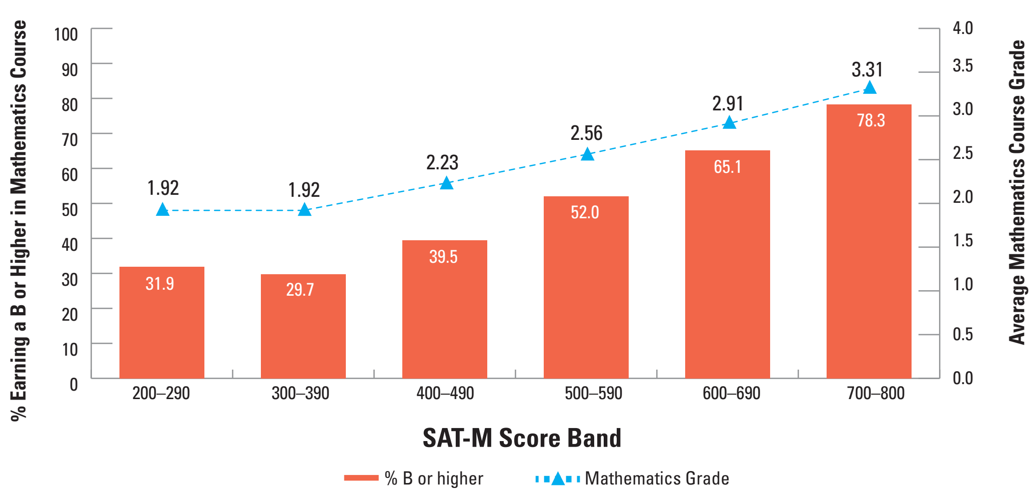
Differential validity by institutional characteristics was also examined. As shown in Table 2, across all institutional groups the correlation between SAT-M and mathematics course grades was higher compared to HSGPA.

- The correlations of SAT-M with mathematics course grades varied only slightly across institutional groups, with correlations ranging from .49 to .54. The correlation was slightly higher at private institutions compared to public institutions and the most selective institutions (those admitting under 50% of applicants).

- Compared to HSGPA, SAT-M was more strongly related to mathematics course grades for all institutional groups.

Figure 1.

The relationship between SAT-M scores and first-year mathematics grades



Note: The sample sizes by SAT score band are as follows: 200–290 = 91; 300–390 = 1,216; 400–490 = 11,215; 500–590 = 25,544; 600–690 = 23,481; and 700–800 = 9,293.

- The data were also analyzed by math content area, such as algebra, precalculus, and calculus. Content areas with fewer than five courses and based on fewer than three institutions were grouped together as "Other Mathematics." Descriptive statistics by mathematics content areas are provided in Table 3. The most common first-year mathematics course was Calculus I, which accounted for 28% of courses. Students enrolled in Calculus III courses had the highest mathematics course grades (mean = 3.09) and the highest SAT-M scores (mean = 730). Students enrolled in business mathematics courses had the lowest mathematics grades (mean = 2.36) but only slightly below-average SAT-M scores (mean = 577). Students enrolled in courses covering algebra also had low average course grades and SAT-M scores compared to the total group.

Table 3.

Descriptive Statistics of Study Variables for First-Year Mathematics Courses by Course Content

			Course Grade		SAT-CR		SAT-M		SAT-W		HSGPA	
Variable		*n*	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD
Course Content	Algebra	11,850	2.49	1.23	507	81	515	74	501	78	3.41	0.51
	Algebra & Trigonometry	2,237	2.42	1.12	522	75	536	64	511	72	3.45	0.48
	Business Mathematics	3,101	2.36	1.29	545	75	577	70	544	74	3.76	0.41
	Calculus I	20,812	2.84	1.14	593	86	636	74	586	85	3.77	0.43
	Calculus II	4,541	2.94	1.05	636	83	697	62	628	84	3.93	0.36
	Calculus III	426	3.09	0.94	648	80	730	55	641	76	3.96	0.36
	Calculus Other	381	2.71	1.18	602	92	631	75	588	95	3.65	0.48
	Discrete/Finite Mathematics	1,892	2.79	1.08	557	98	547	84	552	94	3.45	0.50
	Precalculus	17,647	2.71	1.22	547	88	563	90	541	86	3.59	0.50
	Statistics/ Probability	3,647	2.51	1.16	531	78	547	72	525	76	3.48	0.46
	Other Mathematics	4,306	2.97	1.07	566	92	578	86	562	91	3.60	0.48
Total		70,840	2.71	1.18	560	93	586	94	554	91	3.63	0.49

Note: *n* = number of students; SD = standard deviation.

Correlations by mathematics content areas are provided in Table 4 (refer to Appendix B for uncorrected correlations). SAT-M was most highly correlated with grades in business mathematics (.59), Calculus I (.54), statistics/probability (.54), discrete/finite mathematics (.52), and "other" calculus (.51) courses. On the other hand, SAT-M had the lowest correlation with grades in Calculus III (.29) courses, which may be due to higher-level content that does not overlap with the knowledge and skills measured by SAT-M. In general, SAT-M was equally or more strongly related to course grades than HSGPA across content areas except for Calculus II, Calculus III, and precalculus.

Table 4.

Corrected Correlations of SAT Scores and HSGPA with First-Year Mathematics Course Grades by Course Type and Content

Variable		*k*	*n*	SAT-CR	SAT-M	SAT-W	SAT	HSGPA	HSGPA, SAT
Course Content	Algebra	46	11,850	.35	.47	.39	.48	.46	.54
	Algebra & Trigonometry	6	2,237	.35	.49	.40	.50	.49	.57
	Business Mathematics	8	3,101	.45	.59	.51	.60	.50	.64
	Calculus I	104	20,812	.42	.54	.46	.55	.48	.60
	Calculus II	32	4,541	.39	.49	.43	.50	.50	.57
	Calculus III	5	426	.25	.29	.31	.33	.41	.44
	Calculus Other	7	381	.38	.51	.42	.51	.49	.58
	Discrete/Finite Mathematics	15	1,892	.38	.52	.43	.53	.51	.60
	Precalculus	29	3,647	.35	.46	.39	.46	.47	.54
	Statistics/ Probability	36	4,306	.43	.54	.46	.55	.52	.62
	Other Mathematics	90	17,647	.41	.53	.44	.54	.47	.58
Total		378	70,840	.40	.52	.44	.52	.48	.58

Note: *k* = number of courses; *n* = number of students. Pooled within-course correlations are presented. Correlations are corrected for restriction of range using the institution's enrolled first-year class as the population. SAT is the multiple correlation for all three sections. HSGPA, SAT is the multiple correlation of all three SAT sections and HSGPA.

Differential prediction. The extent to which the SAT and HSGPA resulted in differential prediction for mathematics course grades was examined by student characteristics. Similar to differential prediction results for FYGPA (Mattern, Patterson, Shaw, Kobrin, & Barbuti, 2008), SAT-M scores and to a lesser extent HSGPA underpredicted math course grades for females (see Table 5). For example, the SAT underpredicted mathematics course grades for females by 0.11. In other words, females had an average mathematics course grade of 2.79 (Table 1), but the model predicted a value of 2.68. Regarding ethnicity, SAT-M scores and to a larger extent HSGPA overpredicted mathematics course grades for African American and Hispanic students. For Asian American students, mathematics course grades were underpredicted by SAT-M scores and, to a larger extent, HSGPA. SAT-M scores and HSGPA both underpredicted mathematics course grades for students who reported that their best language was another language but accurately predicted mathematics course grades for English-speaking students and students who reported that their best language was English and another language. Differential prediction results for the other SAT sections and the combination of SAT sections and HSGPA are also provided in Table 5.

Table 5.

Average Over- (-) and Underprediction (+) of First-Year Mathematics Course Grades for SAT Scores and HSGPA by Subgroups

Variable		*k*	*n*	SAT-CR	SAT-M	SAT-W	SAT	HSGPA	HSGPA, SAT
Gender	Male	362	34,572	-0.09	-.014	-0.07	-.012	-0.04	-0.08
	Female	375	36,268	0.08	0.13	0.06	0.11	0.03	0.07
Race/ Ethnicity	American Indian	157	376	-0.16	-0.14	-0.12	-0.12	-0.14	-.011
	Asian American	304	6,978	0.13	0.05	0.13	0.06	0.15	0.08
	Black/African American	314	5,418	-0.25	-0.13	-0.22	-.012	-0.25	-0.08
	Hispanic	310	5,793	-0.18	-0.10	-0.16	-0.09	-0.24	-0.09
	White	369	47,338	0.03	0.02	0.03	0.02	0.04	0.01
	Other	306	1,981	-0.01	0.02	0.00	0.02	0.00	0.03
	No Response	332	2,956	-0.03	-0.01	-0.03	-0.02	0.01	-0.01
Best Language	English Only	376	65,535	-0.01	0.00	-0.01	-0.01	0.00	-0.01
	English & Another Language	309	3,783	0.00	0.02	0.00	0.03	-0.05	0.03
	Another Language	196	815	0.47	0.29	0.50	0.34	0.33	0.33
	No Response	244	707	-0.02	0.00	-0.02	0.00	-0.02	0.00
Total		375	70,840	0.00	0.00	0.00	0.00	0.00	0.00

Note: *k* = number of courses; *n* = number of students. Mean residuals are provided. Negative values indicate overprediction. Positive values indicate underprediction. Values are computed by subtracting the predicted course grade from the actual course grade. Course grade prediction equations are calculated for each course separately.

English Courses

Analyses conducted for English courses parallel the analyses reported above for mathematics courses and are summarized below. Results are based on 222 first-year English courses completed by 96,589 students.

Descriptive Statistics. Table 6 provides the means and standard deviations of the study variables for first-year English courses for the total sample; by student characteristics of gender, race/ethnicity, and best language; and by institutional characteristics of control, size, and selectivity. Of the course-level data from the 110 institutions, 222 courses were identified as introductory English courses. The average grade in these courses was a B (mean = 3.11; SD = .90). In terms of students' high school academic performance, the average self-reported HSGPA was close to an A- (mean = 3.54; SD = .51). The average SAT scores were 544 (SD = 90), 565 (SD = 96), and 539 (SD = 90) for SAT-CR, SAT-M, and SAT-W, respectively.

Table 6.
Descriptive Statistics of Study Variables for First-Year English Courses

Variable		n	Course Grade mean	Course Grade SD	SAT-CR mean	SAT-CR SD	SAT-M mean	SAT-M SD	SAT-W mean	SAT-W SD	HSGPA mean	HSGPA SD
Gender	Male	44,657	3.00	0.96	550	90	589	95	537	91	3.49	0.52
	Female	51,932	3.21	0.83	539	90	543	91	540	89	3.58	0.49
Race/ Ethnicity	American Indian	513	2.93	1.01	534	83	544	87	519	83	3.47	0.54
	Asian American	8,485	3.18	0.82	554	100	617	99	556	99	3.63	0.47
	Black/ African American	7,711	2.77	1.05	499	84	497	85	490	82	3.36	0.54
	Hispanic	6,708	2.86	1.01	515	89	528	93	511	88	3.52	0.51
	White	66,283	3.17	0.86	550	87	569	91	544	87	3.55	0.50
	Other	2,895	3.09	0.89	544	93	560	98	541	93	3.53	0.50
	No Response	3,994	3.15	0.88	566	96	573	99	557	97	3.56	0.51
Best Language	English Only	89,848	3.12	0.90	546	89	565	94	540	89	3.54	0.51
	English & Another Language	4,682	2.97	0.95	520	96	558	109	525	99	3.56	0.49
	Another Language	1,000	3.11	0.87	465	94	602	119	481	98	3.62	0.51
	No Response	1,059	3.00	0.93	528	101	542	110	520	102	3.45	0.56
Control	Public	66,807	3.10	0.94	533	86	555	93	526	84	3.52	0.51
	Private	29,782	3.14	0.79	569	95	585	99	567	96	3.58	0.50
Size	Small	4,357	2.93	0.91	534	96	539	97	531	93	3.43	0.54
	Medium to Large	20,846	3.07	0.90	542	94	554	97	538	93	3.48	0.52
	Large	28,851	3.14	0.93	532	88	553	98	527	87	3.46	0.52
	Very Large	42,535	3.14	0.88	554	88	580	91	548	89	3.63	0.47
Selectivity	Under 50%	16,561	3.19	0.77	587	91	613	98	588	93	3.67	0.46
	50% to 75%	53,895	3.11	0.88	544	89	563	92	538	87	3.55	0.50
	Over 75%	26,133	3.07	1.00	517	82	538	90	510	81	3.44	0.53
Total		96,589	3.11	0.90	544	90	565	96	539	90	3.54	0.51

Note: *n* = number of students. With regard to institution size, small = 750 to 1,999 undergraduates; medium to large = 2,000 to 7,499 undergraduates; large = 7,500 to 14,999 undergraduates; and very large = 15,000 or more undergraduates.

In terms of performance by demographic variables:

- Females, on average, earned higher grades in their first-year English courses (mean = 3.21 versus 3.00 for males) and had slightly higher HSGPAs, whereas males tended to have higher SAT-CR and SAT-M scores. Performance on SAT-W was similar for males and females.

- Asian American and white students had higher SAT scores and HSGPAs, and earned higher first-year English course grades than other racial/ethnic groups (other than "no response").

- Students who reported that English was their best language, along with students who reported that another language was their best language, earned higher first-year English course grades (mean = 3.12 and 3.11, respectively) than students who reported both English and another language as their best language (mean = 2.97). On the other hand, students reporting another language as their best had the lowest mean SAT-CR and SAT-W score (465 and 481, respectively) and students reporting English only as their best language had the highest SAT-CR and SAT-W score (546 and 540, respectively).

Mean performance on these academic indicators also varied by institutional characteristics:

- Students at private institutions had higher SAT scores than students at public institutions but earned only slightly higher grades in high school and in their first-year English courses.

- Students at very large institutions (15,000 or more undergraduates) scored the highest on all academic measures.

- Students attending institutions that admit fewer than 50% of their applicants had higher HSGPAs and SAT scores, and earned higher grades in their first-year English courses, compared to students at less selective institutions. Refer to Table 6 for more detailed information.

Predictive Validity. Next, the predictive validity of the SAT in terms of English course grades was examined. Specifically, for each of the 222 English courses included in the current study, within-course correlations were estimated between SAT scores and HSGPA with course grades. Table 7 provides the overall correlations between SAT scores and HSGPA with English course grades, as well as the correlations by student and institutional characteristics, corrected for restriction of range (refer to Appendix C for the uncorrected correlations). Of the three individual SAT sections, SAT-W correlated most highly with English grades (.37), followed by SAT-CR (.33) and then SAT-M (.29). This pattern was true not just for the overall sample but also for each demographic and institutional subgroup of students. Likewise, this pattern of results is consistent with the Bridgeman et al. study (1992), which found that the SAT Subject Test in Writing was more highly correlated with English course grades than the SAT critical reading section. The correlation between HSGPA and English grades was .39 for the overall sample, a figure that is slightly higher than the correlation between SAT-W and English grades. Overall, the magnitudes of the correlations constitute medium to large effects (Cohen, 1988). Such magnitudes may be biased downward due to the fact that there are relatively few students earning low grades, thereby restricting the variance in the outcome variable. In fact, over 75% of students earned a B (3.00) or higher, suggesting either that English courses may be leniently graded or that there is a ceiling effect. When compared to FYGPA — which has much more variability than English grades — the predictive validity of SAT-W was higher for FYGPA ($r = .51$, as reported in Kobrin et al., 2008).

Table 7.

Corrected Correlations of SAT Scores and HSGPA with First-Year English Course Grades Overall and by Subgroups

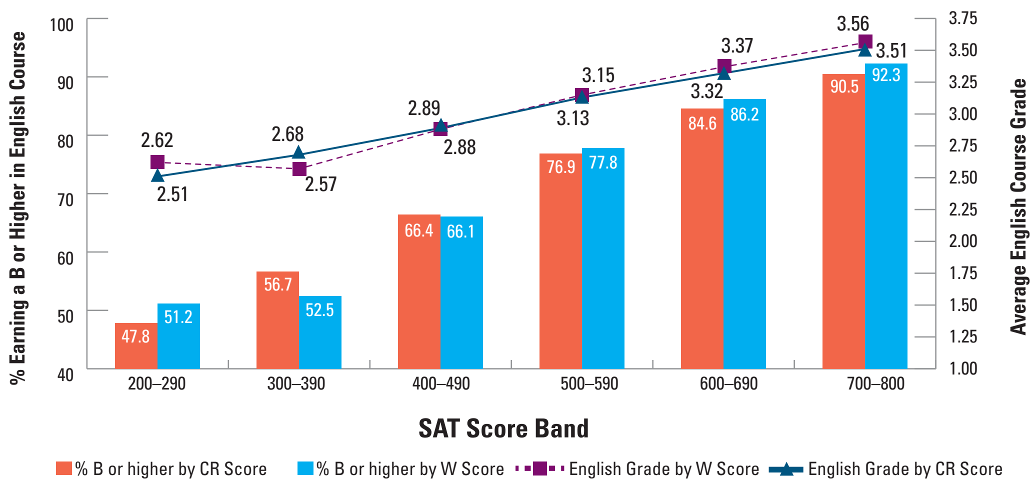
Variable		*k*	*n*	SAT-CR	SAT-M	SAT-W	SAT	HSGPA	HSGPA, SAT
Gender	Male	194	44,514	.30	.27	.33	.34	.36	.41
	Female	202	51,797	.38	.36	.41	.42	.40	.47
Race/ Ethnicity	American Indian	13	181	.35	.30	.36	.37	.26	.38
	Asian American	103	8,110	.27	.21	.29	.30	.30	.35
	Black/African American	106	7,403	.30	.25	.33	.33	.33	.38
	Hispanic	107	6,367	.31	.27	.34	.34	.33	.39
	White	208	66,181	.31	.26	.35	.35	.40	.44
	Other	76	2,460	.36	.31	.39	.39	.37	.44
	No Response	96	3,580	.27	.23	.32	.32	.32	.37
Best Language	English Only	221	89,840	.33	.29	.37	.37	.39	.45
	English & Another Language	95	4,338	.28	.25	.31	.32	.31	.36
	Another Language	22	622	.13	.13	.19	.19	.26	.28
	No Response	40	694	.28	.23	.33	.33	.33	.39
Control	Public	106	66,807	.30	.27	.34	.34	.38	.42
	Private	116	29,782	.39	.33	.42	.43	.42	.49
Size	Small	36	4,357	.38	.33	.43	.44	.42	.50
	Medium to Large	84	20,846	.35	.30	.39	.39	.40	.46
	Large	42	28,851	.29	.25	.32	.33	.37	.41
	Very Large	60	42,535	.34	.30	.38	.38	.39	.45
Selectivity	Under 50%	50	16,561	.36	.30	.40	.40	.40	.46
	50% to 75%	125	53,895	.33	.29	.37	.37	.39	.44
	Over 75%	47	26,133	.30	.27	.34	.34	.39	.43
Total		222	96,589	.33	.29	.37	.37	.39	.44

Note: *k* = number of courses; *n* = number of students. Pooled within-course correlations are presented. Correlations are corrected for restriction of range using the institution's enrolled first-year class as the population. SAT is the multiple correlation for all three sections. HSGPA, SAT is the multiple correlation of all three SAT sections and HSGPA. With regard to institution size, small = 750 to 1,999 undergraduates; medium to large = 2,000 to 7,499 undergraduates; large = 7,500 to 14,999 undergraduates; and very large = 15,000 or more undergraduates.

To provide a graphical alternative to the correlations discussed above, Figure 2 presents the average English course grade along with the percentage of students earning a B or higher by SAT score band for the critical reading and writing sections. For both SAT sections, the mean English course grade increased as the SAT score band increased. Specifically, students in the highest SAT-CR and W score band (i.e., 700–800) earned roughly a whole letter grade higher than students in the lowest SAT score band. Furthermore, roughly half of students in the lowest score band earned a B compared to over 90% in the highest score band, clearly illustrating the strong association between SAT-CR and SAT-W scores and first-year English grades.

Figure 2.

The relationship between SAT-CR and SAT-W with first-year English grades



Note: The sample sizes by SAT-CR score band are as follows: 200–290 = 232; 300–390 = 3,821; 400–490 = 24,998; 500–590 = 39,213; 600–690 = 23,600; and 700–800 = 4,725. The sample sizes by SAT-W score band are as follows: 200–290 = 207; 300–390 = 4,477; 400–490 = 26,460, 500–590 = 39,732; 600–690 = 21,155; and 700–800 = 4,558.

With regard to the differential validity results by student characteristics, the following observations may be made (refer to Table 7):

- Consistent with previous research, correlations between the academic measures and English course grades were consistently higher for females than for males, ranging from .04 higher for HSGPA to .08 to .09 higher for each of the SAT sections. Interestingly, although the differences are not large, SAT-W was a slightly better predictor of English grades for females than HSGPA, whereas HSGPA was a better predictor for males than SAT-W.

- The correlations between SAT sections and English grades were lower for Asian American students and higher for American Indian students relative to the rest of the racial/ethnic groups. For the other racial/ethnic groups, there was little variability in the magnitude of the correlations. The correlation between English grades and HSGPA was highest for white students and lowest for American Indian students, although it should be noted that the sample size for the latter group is relatively small (k = 13; n = 181).

- The correlations were consistently higher across all academic measures for students who reported that their best language was English, whereas the correlations were lowest for students who reported that their best language was another language.

Differential validity by institutional characteristics was also examined.

- As shown in Table 7, SAT scores tended to be more strongly correlated with English course grades at private institutions and at more selective institutions.

- Regarding institutional size, correlations between SAT scores and English course grades were smallest in magnitude for large institutions and largest for small institutions.

- SAT-W was more predictive of English grades than SAT-CR across all student and institutional subgroups examined, and HSGPA tended to be at least as predictive of English grades as SAT-W, with some exceptions for particular subgroups (such as female students).
- One consistent finding across all subgroups was that the best predictor of course grades was the combination of SAT scores and HSGPA.

Results were also examined separately by English content area. Specifically, based on an examination of course titles, courses were coded as either composition/writing or literature. The remaining 15 courses were combined into an "Other English" group, as no other single content category was discovered to have at least five courses across three institutions. Table 8 provides the descriptive statistics of the study variables by the English content areas. Students enrolled in literature courses earned higher grades and higher SAT scores. English course correlations by content area are provided in Table 9 (refer to Appendix D for the uncorrected correlations). For the SAT, the writing section had the highest correlation with English grades across the three content categories. Furthermore, SAT-CR and SAT-W scores were more strongly correlated with course grades in literature courses relative to composition/writing courses. In comparison to HSGPA, SAT-W was more predictive of grades in literature courses but relatively less predictive for composition/writing courses.

Table 8.

Descriptive Statistics of Study Variables for First-Year English Courses by Course Content

			Course Grade		SAT-CR		SAT-M		SAT-W		HSGPA	
Variable		*n*	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD
Course Content	Composition/ Writing	88,596	3.10	0.90	542	89	563	95	536	89	3.53	0.51
	Literature	6,222	3.27	0.73	587	96	597	96	581	94	3.67	0.47
	Other English	1,771	3.03	1.10	523	101	529	98	510	101	3.39	0.55
Total		96,589	3.11	0.90	544	90	565	96	539	90	3.54	0.51

Note: *n* = number of students; SD = standard deviation.

In sum, SAT scores, specifically SAT-CR and SAT-W scores, were moderately to strongly predictive of first-year English course grades. In many instances, HSGPA was more strongly related to first-year English courses grades compared to SAT scores alone, but in all cases the best predictor was the combination of HSGPA and SAT scores. Finally, it appears that SAT-CR does not add much to the prediction of first-year English course grades above SAT-W (compare SAT-W and SAT correlations). As mentioned above, these findings are partly attributable to lack of variance in English course grades. Additional research should be conducted to determine other possible reasons why SAT-CR and SAT-W scores were not more strongly related to English course performance. One potential explanation could be the subjective nature of grading assignments in English courses. Unlike math examinations, which have only one correct answer, most English college examinations, especially writing/essay tests, are subjectively scored. This potentially reduces the reliability of English course grades. Such an explanation would also help to account for the differential validity results by English content area, given the assumption that literature courses are more objectively graded than composition/writing courses.

Table 9.

Corrected Correlations of SAT Scores and HSGPA with First-Year English Course Grades by Course Content

Variable		*k*	*n*	SAT-CR	SAT-M	SAT-W	SAT	HSGPA	HSGPA, SAT
Course Content	Composition/ Writing	165	88,596	.32	.28	.36	.36	.39	.44
	Literature	42	6,222	.42	.35	.45	.45	.43	.51
	Other English	15	1,771	.26	.25	.30	.31	.37	.40
Total		222	96,589	.33	.29	.37	.37	.39	.44

Note: *k* = number of courses; *n* = number of students. Pooled within-course correlations are presented. Correlations are corrected for restriction of range using the institution's enrolled first year class as the population. SAT is the multiple correlation for all three sections. HSGPA, SAT is the multiple correlation of all three SAT sections and HSGPA.

Differential prediction. Next, the extent to which the SAT and HSGPA resulted in differential prediction for English course grades was examined by student characteristics. Similar to the mathematics results and the results of an earlier report examining the differential prediction for FYGPA (Mattern et al., 2008), SAT scores and HSGPA underpredicted English course grades for females (refer to Table 10). For example, the SAT underpredicted English courses grades for females by 0.10. In other words, females had an average English course grade of 3.21 (Table 6), but the model predicted a value of 3.11. Similarly, SAT scores and HSGPA overpredicted English course grades for males, as well as American Indian, African American, and Hispanic students. In the category of best language, SAT-W and SAT-CR scores underpredicted course grades for students who reported that their best language was another language, whereas HSGPA overpredicted their course grades. English course grades for students who reported that their best language was English and another language were overpredicted by both SAT scores and HSGPA.

Table 10.
Average Over- (-) and Underprediction (+) of First-Year English Course Grades by SAT Scores and HSGPA by Subgroups

Variable		*k*	*n*	SAT-CR	SAT-M	SAT-W	SAT	HSGPA	HSGPA, SAT
Gender	Male	218	44,657	-0.13	-0.15	-0.12	-0.12	-0.10	-0.09
	Female	222	51,932	0.12	0.13	0.10	0.10	0.08	0.08
Race/ Ethnicity	American Indian	126	513	-0.17	-0.17	-0.14	-0.14	-0.16	-0.13
	Asian American	199	8,485	0.01	-0.04	0.00	0.00	-0.01	0.01
	Black/African American	193	7,711	-0.17	-0.17	-0.14	-0.14	-0.17	-0.11
	Hispanic	196	6,708	-0.11	-0.12	-0.10	-0.09	-0.16	-0.10
	White	222	66,283	0.03	0.04	0.03	0.03	0.04	0.02
	Other	189	2,895	-0.01	-0.01	-0.01	-0.01	-0.01	0.00
	No Response	205	3,994	-0.01	0.01	-0.01	-0.01	0.01	-0.01
Best Language	English Only	222	89,848	0.00	0.01	0.00	0.00	0.01	0.00
	English & Another Language	192	4,682	-0.06	-0.10	-0.07	-0.06	-0.12	-0.07
	Another Language	134	1,000	0.10	-0.07	0.10	0.11	-0.07	0.08
	No Response	150	1,059	-0.06	-0.06	-0.05	-0.04	-0.06	-0.04
Total		222	96,589	0.00	0.00	0.00	0.00	0.00	0.00

Note; *k* = number of courses; *n* = number of students. Mean residuals are provided. Negative values indicate overprediction. Positive values indicate underprediction. Values are computed by subtracting the predicted course grade from the actual course grade. Course grade prediction equations are calculated for each course separately.

Future Research

Future research should examine issues related to the predictive validity of SAT scores for English as a second language (ESL) students, who have increasingly comprised a larger percentage of the college-going population recently, and their course-taking patterns, especially given the results from this study. Specifically, this study — whose sample consisted overwhelmingly of non-ESL students — found that correlations were consistently higher across all academic measures for students who reported that their best language was English, whereas the correlations were lowest for students who reported that their best language was another language. The lower correlations for this group may be partly due to the fact that students whose best language was a language other than English scored lower than English-speaking students on the critical reading and writing sections of the SAT but performed comparably in first-year English courses. Future research should examine whether the pattern of college course taking is similar across language groups. In other words, do students reporting a best language other than English first enroll in English courses geared toward ESL students or remedial courses, both of which may be less rigorous than standard introductory English courses?

Similarly, future research should examine course-taking patterns for Hispanic students, in order to understand the differential validity results for mathematics courses. Specifically, Asian American and white students had higher SAT scores and HSGPAs, and earned higher first-year mathematics course grades than the other racial/ethnic groups. Hispanic students had lower SAT scores than white students but similar HSGPAs; however, Hispanic students earned lower grades in their first-year mathematics courses. Why did Hispanic students have similar HSGPAs to white students yet perform worse on the SAT and in college-level mathematics courses? Do Hispanic students take less rigorous mathematics courses in high school, thereby leaving them less prepared for the more challenging workload of college courses?

Future research should also examine college course-taking patterns in terms of course difficulty. What types of students take more rigorous college courses? Do course-taking patterns moderate the relationship between SAT scores and college performance? For example, this study found that the SAT was not as predictive of grades in Calculus III compared to those of Calculus I; however, SAT scores were still a good indicator of which students would enroll in Calculus III as their first mathematics college-level course. This point is indexed by the large differences in mean SAT-M score for Calculus III as compared to the other mathematics content areas.

Conclusion

The current study examined the validity of the SAT for predicting first-year college mathematics and English course grades, and the results revealed that performance on the three sections on the SAT was a meaningful predictor of grades in first-year mathematics and English courses. Furthermore, stronger relationships were found between test scores and grades when the content of the two were aligned. That is, SAT-M scores were even more predictive of first-year mathematics course grades than SAT-CR or SAT-W scores. Likewise, SAT-CR and SAT-W scores were even more predictive of first-year English course grades than SAT-M scores. These results suggest that the content knowledge, skills, and abilities required to perform well on the SAT are also those required to perform well in first-year college English and mathematics courses. Therefore, the study provides further evidence of the validity of SAT scores for use in college admission.

. . . results revealed that performance on the three sections on the SAT was a meaningful predictor of grades in first-year mathematics and English courses.

References

American Educational Research Association/American Psychological Association/National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Bridgeman, B., Hale, G. A., Lewis, C., Pollack, J., & Wang, M. (1992). *Placement validity of a prototype SAT with an essay* (ETS Research Report RR-92-28). Princeton, NJ: Educational Testing Service.

Bridgeman, B., & Wendler, C. (1989). *Prediction of grades and college mathematics courses as a component of the placement validity of SAT mathematics scores* (College Board Research Report No. 1989-9). New York: The College Board.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Gulliksen, H. (1950). *Theory of mental tests*. New York: John Wiley and Sons.

Hezlett, S.A., Kuncel, N., Vey, M.A., Ahart, A.M., Ones, D.S., Campbell, J.P., & Camara, W.J. (2001, April). *The effectiveness of the SAT in predicting success early and late in college: A comprehensive meta-analysis*. Paper presented at the annual meeting of the National Council on Measurement in Education, Seattle.

Kobrin, J. L., Patterson, B. F., Shaw, E. J., Mattern, K. D., & Barbuti, S. M. (2008). *Validity of the SAT for predicting first-year college grade point average* (College Board Research Report No. 2008-5). New York: The College Board.

Lawley, D. N. (1943). A note on Karl Pearson's selection formula. *Royal Society of Edinburgh, Proceedings, Section A, 62*, 28–3.

Lawrence, I. M., Rigol, G. W., Van Essen, T., & Jackson, C.A. (2003). *A historical perspective on the content of the SAT* (College Board Research Report No. 2003-3). New York: The College Board.

Mattern, K. D., Patterson, B. F., Shaw, E. J., Kobrin, J. L., & Barbuti, S. M. (2008). *Differential validity and prediction of the SAT* (College Board Research Report No. 2008-4). New York: The College Board.

Powers, D. E. (2004). Validity of Graduate Record Examinations (GRE) General Test scores for admissions to colleges of veterinary medicine. *Journal of Applied Psychology*, 89 (2), 208–219.

Ramist, L., Lewis, C., & McCamley-Jenkins, L. (1994). *Student group differences in predicting college grades: Sex, language, and ethnic groups* (College Board Research Report No. 93-1). New York: The College Board.

Appendix A

Table A1.

Uncorrected Correlations of SAT Scores and HSGPA with First-Year Mathematics Course Grades Overall and by Subgroups

Variable		*k*	*n*	SAT-CR	SAT-M	SAT-W	SAT	HSGPA	HSGPA, SAT
Gender	Male	306	34,242	.12	.26	.17	.27	.25	.34
	Female	321	35,953	.17	.34	.22	.35	.25	.40
Race/ Ethnicity	American Indian	5	63	.07	.11	.07	.17	.26	.34
	Asian American	137	6,378	.10	.27	.16	.28	.20	.33
	Black/African American	129	4,719	.13	.27	.16	.27	.23	.35
	Hispanic	122	5,120	.11	.25	.16	.27	.19	.31
	White	333	47,101	.13	.26	.18	.27	.28	.37
	Other	65	1,190	.13	.24	.19	.26	.17	.30
	No Response	102	2,120	.13	.24	.18	.26	.25	.33
Best Language	English Only	371	65,490	.16	.28	.21	.30	.27	.38
	English & Another Language	106	3,130	.13	.30	.19	.31	.17	.34
	Another Language	17	373	.02	.23	.11	.25	.13	.27
	No Response	7	114	.20	.45	.35	.48	.33	.52
Control	Public	191	52,567	.14	.28	.20	.29	.25	.36
	Private	187	18,273	.16	.29	.20	.31	.28	.39
Size	Small	50	2,913	.11	.27	.19	.29	.27	.37
	Medium to Large	134	13,180	.15	.29	.20	.30	.31	.40
	Large	77	17,372	.13	.27	.19	.29	.29	.38
	Very Large	117	37,375	.15	.28	.20	.30	.23	.36
Selectivity	Under 50%	98	13,045	.17	.28	.21	.30	.27	.38
	50% to 75%	191	40,088	.14	.28	.19	.29	.24	.36
	Over 75%	89	17,707	.16	.29	.21	.30	.30	.40
Total		378	70,840	.15	.28	.20	.29	.26	.37

Note: *k* = number of courses; *n* = number of students. Pooled within-course correlations are presented. SAT is the multiple correlation for all three sections. HSGPA, SAT is the multiple correlation of all three SAT sections and HSGPA. With regard to institution size, small = 750 to 1,999 undergraduates; medium to large = 2,000 to 7,499 undergraduates; large = 7,500 to 14,999 undergraduates; and very large = 15,000 or more undergraduates.

Appendix B

Table B1.

Uncorrected Correlations of SAT Scores and HSGPA with First-Year Mathematics Course Grades by Course Type and Content

Variable		*k*	*n*	SAT-CR	SAT-M	SAT-W	SAT	HSGPA	HSGPA, SAT
Course Content	Algebra	46	11,850	.11	.25	.16	.26	.28	.36
	Algebra & Trigonometry	6	2,237	.09	.25	.14	.26	.31	.39
	Business Mathematics	8	3,101	.17	.33	.25	.36	.22	.41
	Calculus I	104	20,812	.16	.29	.21	.30	.24	.37
	Calculus II	32	4,541	.12	.21	.17	.24	.25	.32
	Calculus III	5	426	.04	.06	.14	.16	.22	.25
	Calculus Other	7	381	.03	.19	.09	.21	.28	.34
	Discrete/ Finite Mathematics	15	1,892	.14	.30	.21	.31	.32	.40
	Precalculus	29	3,647	.11	.21	.15	.22	.27	.34
	Statistics/ Probability	36	4,306	.20	.33	.24	.35	.32	.43
	Other Mathematics	90	17,647	.17	.31	.22	.32	.25	.39
Total		378	70,840	.15	.28	.20	.29	.26	.37

Note: *k* = number of courses; *n* = number of students. Pooled within-course correlations are presented. SAT is the multiple correlation for all three sections. HSGPA, SAT is the multiple correlation of all three SAT sections and HSGPA.

Appendix C

Table C1.

Uncorrected Correlations of SAT Scores and HSGPA with First-Year English Course Grades Overall and by Subgroups

Variable		*k*	*n*	SAT-CR	SAT-M	SAT-W	SAT	HSGPA	HSGPA, SAT
Gender	Male	194	44,514	.14	.11	.19	.20	.23	.28
	Female	202	51,797	.20	.17	.23	.25	.23	.31
Race/ Ethnicity	American Indian	13	181	.24	.16	.23	.26	.19	.30
	Asian American	103	8,110	.16	.07	.20	.20	.19	.26
	Black or African American	106	7,403	.13	.08	.18	.18	.21	.26
	Hispanic	107	6,367	.15	.10	.19	.20	.20	.26
	White	208	66,181	.13	.07	.20	.20	.26	.30
	Other	76	2,460	.20	.12	.24	.25	.23	.31
	No Response	96	3,580	.15	.09	.21	.21	.21	.28
Best Language	English Only	221	89,840	.16	.10	.22	.22	.25	.31
	English & Another Language	95	4,338	.15	.11	.21	.21	.20	.26
	Another Language	22	622	.05	.06	.11	.11	.15	.18
	No Response	40	694	.19	.14	.26	.26	.27	.34
Control	Public	106	66,807	.14	.10	.20	.20	.24	.29
	Private	116	29,782	.21	.12	.26	.27	.26	.34
Size	Small	36	4,357	.22	.15	.29	.30	.30	.38
	Medium to Large	84	20,846	.17	.09	.23	.23	.26	.32
	Large	42	28,851	.13	.08	.19	.19	.25	.29
	Very Large	60	42,535	.17	.12	.23	.23	.24	.30
Selectivity	Under 50%	50	16,561	.20	.11	.25	.26	.25	.32
	50% to 75%	125	53,895	.16	.09	.21	.21	.23	.29
	Over 75%	47	26,133	.14	.12	.21	.21	.28	.32
Total		222	96,589	.16	.10	.22	.22	.25	.30

Note: *k* = number of courses; *n* = number of students. Pooled within-course correlations are presented. SAT is the multiple correlation for all three sections. HSGPA, SAT is the multiple correlation of all three SAT sections and HSGPA. With regard to institution size, small = 750 to 1,999 undergraduates; medium to large = 2,000 to 7,499 undergraduates; large = 7,500 to 14,999 undergraduates; and very large = 15,000 or more undergraduates.

Appendix D

Table D1.

Uncorrected Correlations of SAT Scores and HSGPA with First-Year English Course Grades by Course Content

Variable		*k*	*n*	SAT-CR	SAT-M	SAT-W	SAT	HSGPA	HSGPA, SAT
Course Content	Composition/ Writing	165	88,596	.15	.10	.21	.21	.25	.30
	Literature	42	6,222	.25	.14	.28	.30	.25	.35
	Other English	15	1,771	.12	.07	.19	.19	.28	.31
Total		222	96,589	.16	.10	.22	.22	.25	.30

Note: *k* = number of courses; *n* = number of students. Pooled within-course correlations are presented. SAT is the multiple correlation for all three sections. HSGPA, SAT is the multiple correlation of all three SAT sections and HSGPA.

The Research & Development department actively supports the College Board's mission by:

- Providing data-based solutions to important educational problems and questions
- Applying scientific procedures and research to inform our work
- Designing and evaluating improvements to current assessments and developing new assessments as well as educational tools to ensure the highest technical standards
- Analyzing and resolving critical issues for all programs, including AP®, SAT®, PSAT/NMSQT®
- Developing standards and conducting college and career readiness alignment studies
- Publishing findings and presenting our work at key scientific and education conferences
- Generating new knowledge and forward-thinking ideas with a highly trained and credentialed staff

Our work focuses on the following areas

Admission	Measurement
Alignment	Research
Evaluation	Trends
Fairness	Validity

Follow us online: collegeboard.org/research

CollegeBoard

CollegeBoard

11b-4950