Validity of the SAT® for Predicting First-Year Grades and Retention to the Second Year

PAUL A. WESTRICK, JESSICA P. MARINI, LINDA YOUNG, HELEN NG, DORON SHMUELI, AND EMILY J. SHAW

Contents

Abstract

This report represents the first national operational SAT® validity study since the SAT was redesigned and launched in March 2016. This is among the largest SAT validity studies ever conducted and is based on data from more than 223,000 students across 171 four-year colleges and universities. Results show that the SAT is essentially as effective as high school grades in predicting students' college performance and, when these two measures are combined, offers the most accurate understanding of student performance than either measure used alone. Analyses also demonstrate that SAT scores are useful for understanding whether students will be retained to the second year of college. In particular, findings show:

- SAT scores are strongly predictive of college performance—students with higher SAT scores are more likely to have higher grades in college.
- SAT scores are predictive of student retention to their second year—students with higher SAT scores are more likely to return for their sophomore year.
- SAT scores and HSGPA are both related to academic performance in college but tend to measure slightly different aspects of academic preparation. Using SAT scores in conjunction with HSGPA is the most powerful way to predict future academic performance.
    - On average, SAT scores add 15% more predictive power above grades alone for understanding how students will perform in college.
    - SAT scores help to further differentiate student performance in college within narrow HSGPA ranges.
- Colleges can use SAT scores to identify students who may be in need of academic support before they start college and throughout their college education by monitoring predicted versus actual performance and by helping to position these students for success.

Having a more accurate understanding of students' future performance helps ensure that colleges and universities will not overlook or miss students who will be successful on campus. Also, the combination of SAT score information with HSGPA information helps institutions identify enrolling students who may benefit from additional support and monitoring to ensure that they are academically successful in a way that HSGPA alone cannot do. Future research will examine SAT validity by institutional and student subgroups and will continue to analyze and document the relationship between SAT scores and other college outcomes, including course grades, later college performance, and degree completion.

Introduction

College Board launched the redesigned SAT in 2016 to better reflect the work that students do in high school, focusing on the core knowledge and skills that research has shown to be critical for students to be ready for college and career. This study represents the first national operational SAT validity research study to examine the utility of SAT scores for admission decisions, focusing on the outcomes of first-year grade point average (FYGPA) and retention to the second year. This report is divided into two sections with the first section focusing on SAT score relationships with FYGPA and the second section focusing on SAT score relationships with retention outcomes.

As noted above, College Board aimed to make the SAT an assessment that reflects the work that students need to do to be ready for and successful in college.[1] Scholarly research and empirical data derived from secondary and postsecondary curriculum surveys conducted by College Board and other organizations formed the evidentiary foundation for specifying the test content and domains of interest. The SAT has three sections: The Evidence-Based Reading and Writing section, the Math section, and an optional Essay section. The Evidence-Based Reading and Writing section and the Essay section incorporate key design elements supported by evidence, including:

- The use of a range of text complexity aligned to college- and career-ready reading levels;
- An emphasis on the use of evidence and source analysis;
- The incorporation of data and informational graphics that students will analyze along with text;
- A focus on relevant words in context and on word choice for rhetorical effect;
- Attention to a core set of important English language conventions and to effective written expression; and
- The requirement that students interact with texts across a broad range of disciplines.

The key evidence-based design elements incorporated into the SAT Math section include:

- A focus on the content that matters most for college and career readiness (rather than broad coverage of a vast array of content and skills);
- An emphasis on problem solving and data analysis; and
- The inclusion of "Calculator: Permitted" questions as well as "Calculator: Not Permitted" questions and attention to the use of the calculator as a tool.

The SAT Evidence-Based Reading and Writing section and the Math section each report scores on a 200–800 scale. The Evidence-Based Reading and Writing section has two test scores, the Reading Test score and the Writing and Language Test score, each reported on a 10–40 scale. The Math section also produces a Math Test score that is reported on a 10–40 scale. Several new subscores and cross-test scores provide richer information to students, schools, and institutions on student performance.

---

[1] More information on the development of the SAT can be found in *Test Specifications for the Redesigned SAT®* (College Board, 2015) and *SAT® Suite of Assessments Technical Manual: Characteristics of the SAT* (College Board, 2017).

Students earn points for correct answers to questions, but, unlike the previous SAT, they do not lose points for incorrect answers. For the optional Essay, students now have 50 minutes instead of 25 minutes to write a response to a reading passage.

To gather early SAT validity evidence, a pilot predictive validity study was conducted in 2014-2015 to study the relationships between SAT scores and college performance across a sample of 2,050 students at 15 four-year institutions, using a pilot form of the redesigned SAT (Shaw, Marini, Beard, Shmueli, Young, & Ng, 2016). That study found that the redesigned SAT was as predictive of college success as the previous SAT, that redesigned SAT scores improved the ability to predict college performance above HSGPA alone, and that there was a strong, positive relationship between redesigned SAT scores and grades in matching college course domains, suggesting that the redesigned SAT is sensitive to instruction in English language arts, math, science, and history/social studies. The entering class of fall 2017 was the first cohort of students to have primarily taken the redesigned SAT and completed the first year of college in 2017-2018 so that their corresponding college outcome information could be examined for validity research. Therefore, this first validity study will focus on the relationships between SAT section scores and FYGPA and retention to the second year for that cohort.

SAT Score Relationships with First Year Grade-Point-Average
Methodology

Sample
College Board broadly recruited four-year institutions with at least 250 first-year students (at least 75 of those students had to have SAT scores) to participate in this initial SAT validity study. These institutions provided data through College Board's secure online Admitted Class Evaluation Service™ (ACES™) system. Ultimately, 171 institutions provided the complete student-level information needed for the analyses that follow in this section of the report. See Appendix A for the list of participating institutions.

Table 1 includes the characteristics of the institutions in the sample and shows that the sample is quite diverse with regard to region of the U.S., control (public/private), selectivity, and size. Compared to the population[2] of four-year institutions for this study, the institutional study sample included more public institutions, more selective institutions, and more "large" and "very large" institutions than the reference population. This is to be expected as there was a sample size minimum to participate in the study and more-selective institutions rather than less selective-institutions would be more apt to use the SAT and examine the relationship between the SAT and college outcomes.

---

[2] The population included four-year public or private nonprofit institutions that accepted 90% or fewer applicants for admission.

Table 1: Institutional Characteristics of the 2017 SAT Validity Study Sample and Population of Four-Year Institutions

	Variable	FYGPA Sample (*k*=171)	Reference Population of Institutions (*k*=1,230)
U. S. Region	Midwest	35 (20%)	343 (28%)
	Mid-Atlantic	31 (18%)	246 (20%)
	New England	22 (13%)	119 (10%)
	South	28 (16%)	277 (23%)
	Southwest	20 (12%)	90 (7%)
	West	35 (20%)	155 (13%)
Control	Public	82 (48%)	417 (34%)
	Private	89 (52%)	813 (66%)
Admittance Rate	Under 25%	20 (12%)	57 (5%)
	25% to 50%	30 (18%)	211 (17%)
	51% to 75%	73 (43%)	651 (53%)
	Over 75%	48 (28%)	311 (25%)
Undergraduate Enrollment	Small	67 (39%)	761 (62%)
	Medium	30 (18%)	202 (16%)
	Large	30 (18%)	136 (11%)
	Very Large	44 (26%)	131 (11%)

Note. *k* = number of institutions. Percentages may not sum to 100 due to rounding. Undergraduate enrollment was categorized as follows: small: 4,999 or less; medium: 5,000 to 9,999; large: 10,000 to 19,999; and very large: 20,000 or more.

Inclusion in the study sample required students to have redesigned SAT scores, a valid self-reported HSGPA, and a valid FYGPA supplied by the institution. This resulted in a sample size of 223,858 students. See Table 2 for more information about the characteristics of the student sample and the population of 2017 graduating seniors who took the SAT. Compared to the population, the study sample, which included students who were enrolled in college, tended to have slightly more female students, slightly more white students and fewer black or African American students, and more students whose highest parental education level was a bachelor's degree or higher than the overall SAT-taking population.

Table 2: Student Characteristics of the 2017 SAT Validity Study Sample and 2017 Graduating Seniors with SAT Scores

	Variable	FYGPA Sample (*n* = 223,858)	2017 Graduating Seniors Who Took the SAT (*N* = 1,715,481)
Gender	Male	97,080 (43%)	809,462 (47%)
	Female	126,778 (57%)	906,019 (53%)
Race/Ethnicity	American Indian or Alaska Native	658 (<1%)	7,782 (<1%)
	Asian	25,209 (11%)	158,031 (9%)
	Black or African American	16,004 (7%)	225,860 (13%)
	Hispanic or Latino	47,175 (21%)	408,067 (24%)
	Native Hawaiian or Other Pacific Islander	319 (<1%)	4,131 (<1%)
	White	122,750 (55%)	760,362 (44%)
	Two or More Races	8,548 (4%)	57,049 (3%)
	Not Stated	3,195 (1%)	94,199 (5%)
Highest Parental Education Level	No High School Diploma	12,850 (6%)	137,437 (8%)
	High School Diploma	48,127 (21%)	482,194 (28%)
	Associate Degree	15,659 (7%)	134,451 (8%)
	Bachelor's Degree	80,465 (36%)	473,103 (28%)
	Graduate Degree	63,539 (28%)	339,743 (20%)
	Not Stated	3,218 (1%)	148,553 (9%)

Measures

High School GPA (HSGPA). Students' self-reported HSGPA was obtained from the SAT Questionnaire when they registered for the SAT and is reported on a 12-point interval scale, ranging from 0.00 (F) to 4.33 (A+). Institutional HSGPA could not be used in this national study because it is reported on so many different scales across institutions. Note that the inclusion of self-reported HSGPA is consistent with previous admission test validity studies (e.g. Mattern and Patterson, 2014; Sawyer, 2013) and studies have found self-reported HSGPA to be highly correlated with actual HSGPA (Kuncel, Credé, & Thomas, 2005; Shaw & Mattern, 2009). In the class of 2017, 93% of the SAT-taking population reported their HSGPA. The HSGPA measure in this study had a sample mean of 3.67 (SD=0.47).

SAT Scores. SAT scores were obtained from College Board's database and matched to each student provided in the institution files. The SAT scores included in this study are:

SAT Total Score (400 to 1600 scale)—increments of 10, sample mean of 1187 (SD=163) for the FYGPA analyses.

SAT Evidence-Based Reading and Writing (ERW) Section Score (200 to 800 scale) —increments of 10, sample mean of 596 (SD=83) for the FYGPA analyses.

SAT Math Section Score (200 to 800 scale)—increments of 10, sample mean of 591 (SD=93) for the FYGPA analyses.

College Grades. Each institution provided FYGPA values for their 2017 first-time, first-year students. The FYGPAs across the 171 institutions in this sample ranged from 0.00 to 4.30. FYGPA had a sample mean of 3.03 (SD=0.81).

Descriptive Statistics

Table 3 includes descriptive statistics for all measures of interest in the sample and for the 2017 SAT-tested graduating seniors. As the sample includes students enrolled in college, it is not surprising that these students are academically stronger than the total SAT test-taking population across all measures. Descriptive statistics are reported for all SAT scores utilized in the study analyses: SAT ERW section, SAT Math section, and SAT Total scores, as well as HSGPA and FYGPA. Additional SAT scores will be examined in follow-up studies related to course placement and subscore validity.

Table 3: Descriptive Statistics for Measures of Interest

	FYGPA Sample (*n* = 223,858)				2017 Graduating Seniors Who Took the SAT (*N* = 1,715,481)			
	M	*SD*	*Min*	*Max*	*M*	*SD*	*Min*	*Max*
HSGPA	3.67	0.47	0.00	4.33	3.33	0.65	0.00	4.33
SAT Total	1187	163	400	1600	1060	195	400	1600
SAT ERW	596	83	200	800	533	100	200	800
SAT Math	591	93	200	800	527	107	200	800
FYGPA	3.03	0.81	0.00	4.30				
Note. Not all 2017 graduating seniors who took the SAT reported their HSGPA (*n* = 1,594,136).								

Table 4 displays the intercorrelations between the predictors in the current study – SAT ERW section scores, SAT Math section scores, and HSGPA. Similar to previous research (Kobrin, Patterson, Shaw, Mattern, & Barbuti, 2008; Shaw et al., 2016), the correlation between each of the SAT section scores and HSGPA is .50, indicating a strong relationship but also demonstrating that the two measures — SAT and HSGPA — are not identical constructs.

Table 4: Corrected (Raw) Correlation Matrix of SAT Scores and HSGPA

	SAT ERW	SAT Math	HSGPA
SAT ERW			
SAT Math	.78 (.58)		
HSGPA	.50 (.24)	.50 (.23)	
Note. The correlation between SAT Total score and HSGPA was .53 (.27).			

Methods

Analyses consisted of correlations between the predictors — SAT scores and HSGPA — with FYGPA, and logistic regression analyses for predicting students' probability of earning a FYGPA of 2.50 or higher. This FYGPA criterion was selected as a reasonable threshold for indicating that a student is managing to navigate college-level work and can remain enrolled and progress through college with that GPA or higher.

Raw and adjusted correlations (predictive strength) were calculated between predictors and FYGPA at the institution-level and weighted by the number of students in each institutional analysis. The weighted correlations were summed and then divided by the total number of students across institutions. Correlations were adjusted to account for the selectivity of the student sample. It is a widely accepted practice to statistically correct correlation coefficients in admission validity research for restriction of range because the raw correlation tends to underestimate the true relationship between the test scores and the college outcome (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 2014). Without information on how students who were not admitted or those who did not enroll would have performed at an institution, there is only a small glimpse into how the tests work for selection. This restricts the variability or range in test scores available for analysis since the test scores available tend to be the higher scores of students who were admitted, minimizing the test score–criterion relationship. Correlations in this study were corrected for multivariate range restriction (Lawley, 1943) using the 2017 graduating seniors who took the SAT as the reference population.

To estimate the probability of earning a FYGPA of 2.5 or higher, logistic regression analyses were conducted at each institution; the institution-level coefficients were weighted by the number of students in the institutional study; and then mean coefficients from the aggregated weights were calculated.

Results

Table 5 shows the correlations of the singular predictors and combinations of predictors with FYGPA. The adjusted correlations of the different predictors with FYGPA ranged from .47 (SAT Math) to .61 (SAT and HSGPA); the correlation between SAT and FYGPA was .51. Cohen (1988) defined correlations with absolute values of .50 or higher as large, correlations with absolute values less than .50 and greater than

or equal to .30 as medium, and correlations with absolute values less than .30 but greater than or equal to .10 as small. To contextualize the strength of these values, as well as other correlation coefficients discussed in this analysis, it can be helpful to understand the relative strength of other relationships that are believed to co-occur. For example, research has found that the correlation between taking aspirin and reduced risk of death by heart attack is .02, smoking and subsequent incidence of lung cancer within 25 years is .08, and the validity of employment interviews for predicting job success is .20 (Meyer et al., 2001).

Despite the strength of each variable analyzed individually to predict student collegiate success, the use of these predictors in combination provides institutions with the greatest benefits. The correlations between HSGPA and the SAT with FYGPA are .53 and .51, respectively. When HSGPA and SAT are combined, the correlation with FYGPA jumps to .61, an increase of .08 and a 15% boost in the correlation calculated when using HSGPA alone. Jointly using HSGPA and SAT scores to predict first-year academic performance surpasses the predictive strength of either predictor used alone. These findings are consistent with findings from the most recent national SAT validity study examining the *previous* version of the SAT, which found nearly identical relationships between HSGPA, SAT, and FYGPA, including the incremental validity of the SAT above HSGPA to predict FYGPA of .08 (Beard & Marini, 2018).

Table 5: Corrected (Raw) Correlations of Predictors with FYGPA

Predictor(s)	Correlation
SAT, HSGPA	.61 (.42)
HSGPA	.53 (.33)
SAT	.51 (.32)
SAT ERW	.49 (.29)
SAT Math	.47 (.27)
Note. n = 223,858. References to "SAT" on its own include SAT ERW and SAT Math sections.	

The correlations in Table 5 represent relationships that can sometimes be better understood visually. As such, the following graphs demonstrate the strength and value of using multiple measures in university contexts to understand and predict students' future academic performance.

Figure 1 graphically depicts the mean FYGPA by SAT score band. As SAT scores increase, so do the average FYGPAs. For example, students with SAT Total scores between 800 and 990 had a mean FYGPA of 2.50. In contrast, students with SAT Total scores between 1400 and 1600 had a mean FYGPA of 3.51, a full letter grade higher than that for the students previously mentioned.
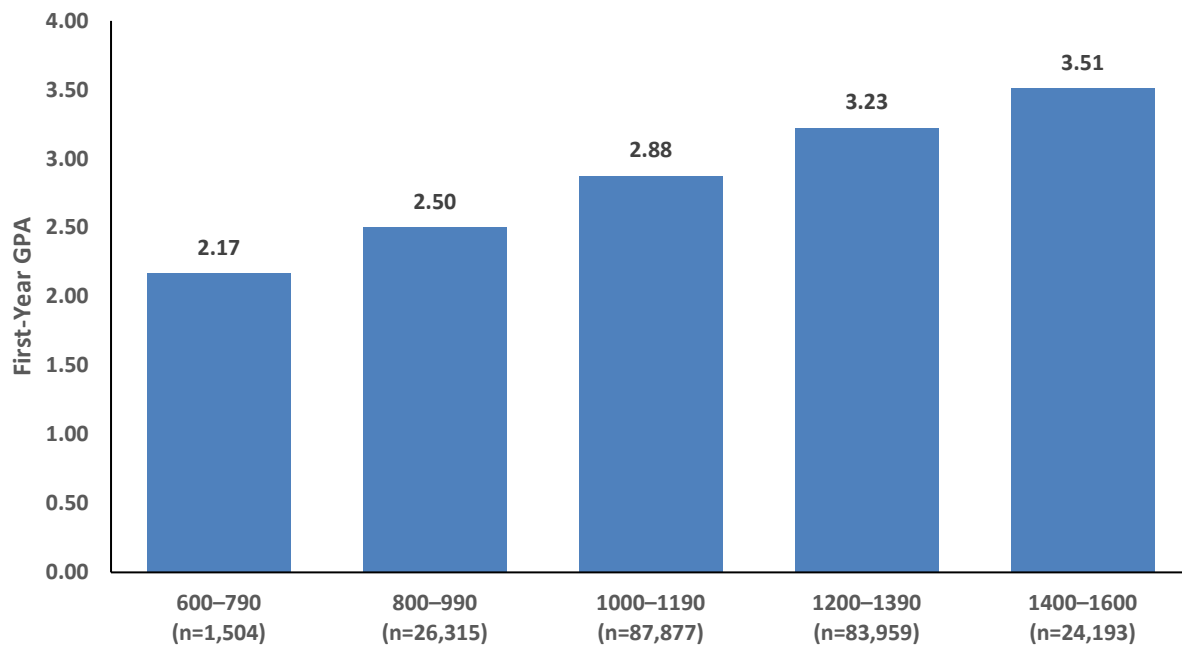
Note. Results based on fewer than 15 students are not reported (e.g., score band 400–590, n = 10).

Figure 2 graphically communicates the validity of the SAT for predicting FYGPA after controlling for HSGPA, complementing the information presented in Table 5 that showed the SAT added a 15% predictive boost above HSGPA to predict FYGPA. Based on SAT Total score bands within each HSGPA category, it is evident that the relationship between SAT scores and FYGPA remains positive and increases by SAT score. If SAT scores did not add information above HSGPA, each SAT score band within a HSGPA category would have the same or very similar mean FYGPAs. That is clearly not the case. Combining HSGPA and SAT information reveals additional insights regarding student performance and allows institutions to more accurately predict differences in the future academic performances of students with similar HSGPAs. Note that as HSGPA increases from C+ or lower to A+, the gaps between students within the same HSGPA category, but within different SAT score bands, increase. This is especially true for students in the A-, A, and A+ categories, which contain more than two-thirds of the students. For example, among students with an "A" HSGPA, students with SAT Total scores between 600 and 790 had a mean FYGPA of 2.44, but students with SAT Total scores between 1400 and 1600 had a mean FYGPA of 3.54, more than a full letter grade higher than the students with the lower scores but in the same HSGPA group.

Figure 2: Mean FYGPA by HSGPA and SAT Total Score Bands
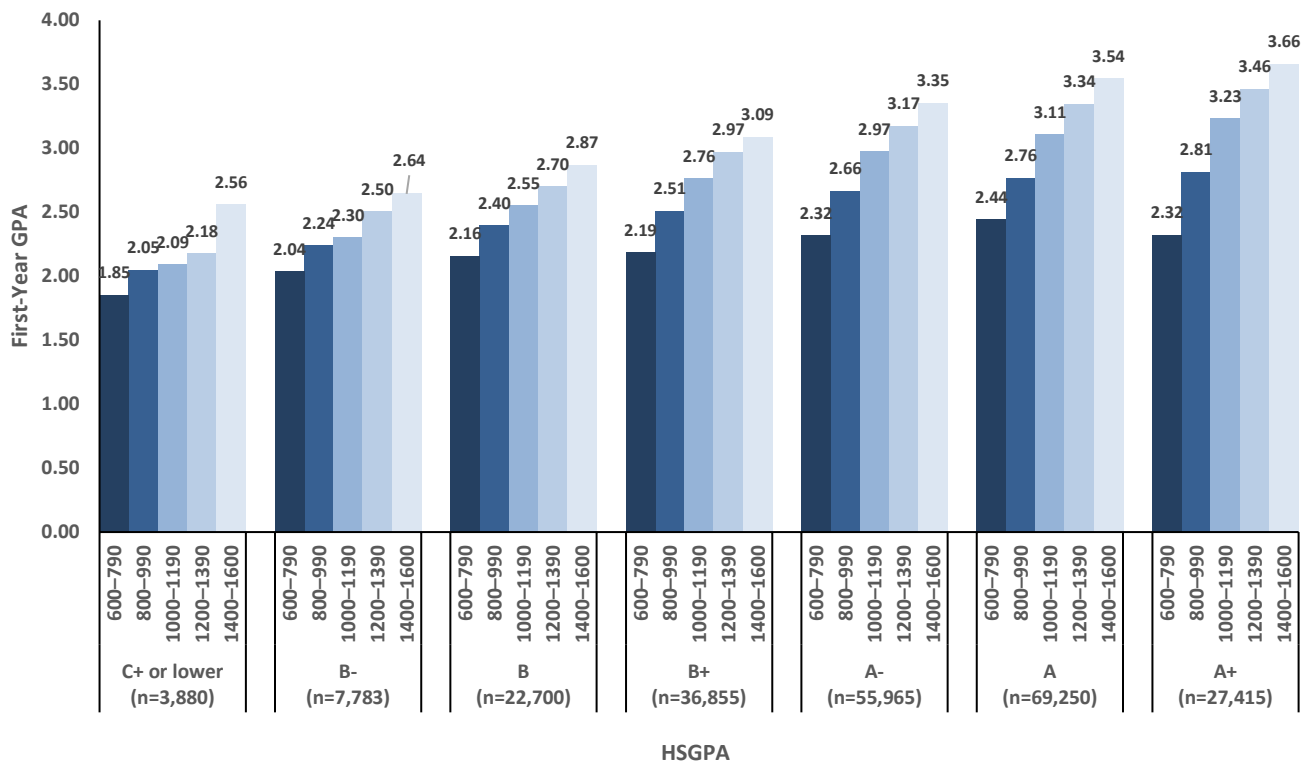
Note. Results are reported for categories with at least 15 students.

Figure 3 further demonstrates the value of using SAT scores with HSGPA to predict future academic success. This graph shows students' probability of earning a FYGPA of 2.50 or higher in college given their HSGPA and selected SAT Total score. For example, a student with a HSGPA of 3.00 and an SAT Total score of 1000, has approximately a 57% chance of earning a FYGPA of 2.50 or higher, while a student with the same HSGPA (3.00) and an SAT Total score of 1400 has approximately an 82% chance of earning a FYGPA of 2.50 or higher. Even among students with higher HSGPAs, we see the added SAT value in understanding student success in college. The SAT scores provide meaningful information in predicting a student's probability of earning a 2.50 or higher FYGPA in college at every point on the HSGPA scale.

Figure 3: Probability of a 2.50 or Higher FYGPA Given HSGPA and SAT Total Score
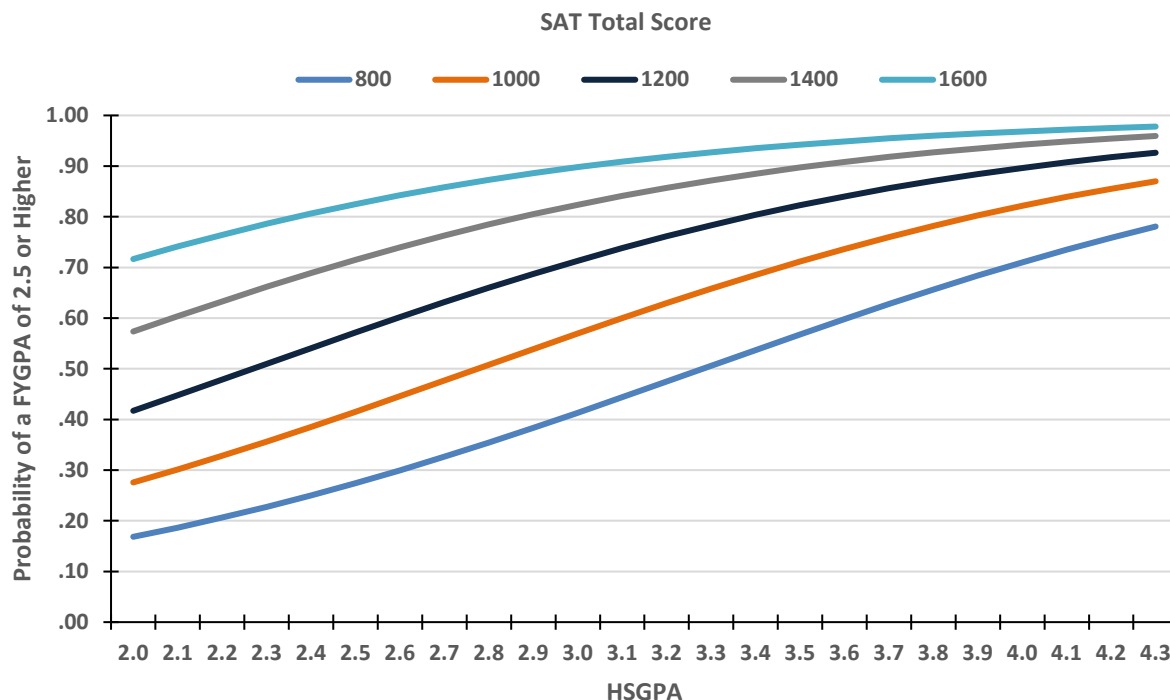
Using SAT scores in conjunction with HSGPA in a compensatory model like the one illustrated above helps institutions predict a student's likelihood of succeeding in college despite having a low level of performance on either of the two predictors. Ignoring SAT scores and using HSGPA alone reduces an institution's ability to identify applicants who may excel despite having low high school grades and applicants who may struggle despite stellar high school grades. The ability to identify students who may struggle academically allows institutions to target these students for academic support, which likely benefits both the student and the institution with regard to retention outcomes.

SAT Score Relationships with Retention to Second Year
Methodology

Sample
Similar to the FYGPA analyses above, inclusion in the retention analysis sample required students to have SAT scores on the redesigned SAT, a valid self-reported HSGPA, and a valid FYGPA supplied by the institution. However, not all participating institutions reported retention data to the National Student Clearinghouse (NSC), resulting in a sample size of 156 institutions and 204,504 students. Retention to the same institution required students to be enrolled in the same institution in the Fall 2017 and Fall 2018 semesters. Appendices B through F provide more detail on the retention analysis student and

institutional samples, which are largely similar to the student and institutional samples examined in the FYGPA analyses.

Methods

Retention analyses consisted of 1) calculating the percentage of students retained to the same institution, 2) using logistic regression to estimate students' probability of returning to the same institution for a second year, and 3) analyzing the relationships between academic over- and under-performance and retention to the second year. As in the FYGPA analyses, logistic regression analyses were conducted at each institution, institution-level coefficients were weighted by the number of students in the institutional study, and the mean coefficients from the aggregated weights were calculated. For the over- and under-performance analyses, students were classified into two categories "Performing as Expected or Overperforming" or "Underperforming." In the first part of this analysis, each student's FYGPA was predicted using both HSGPA and SAT section scores within an institution. That predicted FYGPA was compared to the student's actual FYGPA. A student was categorized as "Performing as Expected or Overperforming" when their actual FYGPA was no more than 1.5 standard deviations below their predicted FYGPA. A student was categorized as "Underperforming" when their actual FYGPA was more than 1.5 standard deviations below their predicted FYGPA.

Results

Table 6 provides a summary of HSGPA, SAT scores, and other performance indicators for students who returned to the same institution and students who did not. On average, students who returned to the same institution had higher SAT scores, HSGPAs, and FYGPAs in college.

*Table 6: Summaries of Study Measures for Students Returning to the Same Institution for the Second Year*

	Returned					Did not Return				
	n	M	SD	Min	Max	n	M	SD	Min	Max
HSGPA	170,247	3.71	0.45	0.00	4.33	34,257	3.47	0.52	0.00	4.33
SAT Total	170,247	1198	160	400	1600	34,257	1122	158	470	1600
SAT ERW	170,247	602	81	200	800	34,257	565	82	200	800
SAT Math	170,247	597	91	200	800	34,257	557	90	200	800
FYGPA	170,247	3.18	0.62	0.00	4.30	34,257	2.28	1.17	0.00	4.21

Note: Students within an institution that participated in NSC with missing retention information were classified as not returning for the second year.

Figure 4 shows the average second-year retention rate by SAT Total score bands for students retained at the same institution. As SAT scores increase, retention rates also increase, showing the positive relationship between SAT scores and retention to the second year. For example, students with SAT Total scores between 800 and 990 had a mean retention rate of 72%. In contrast, students with SAT Total scores between 1400 and 1600 had a mean retention rate of 92%.
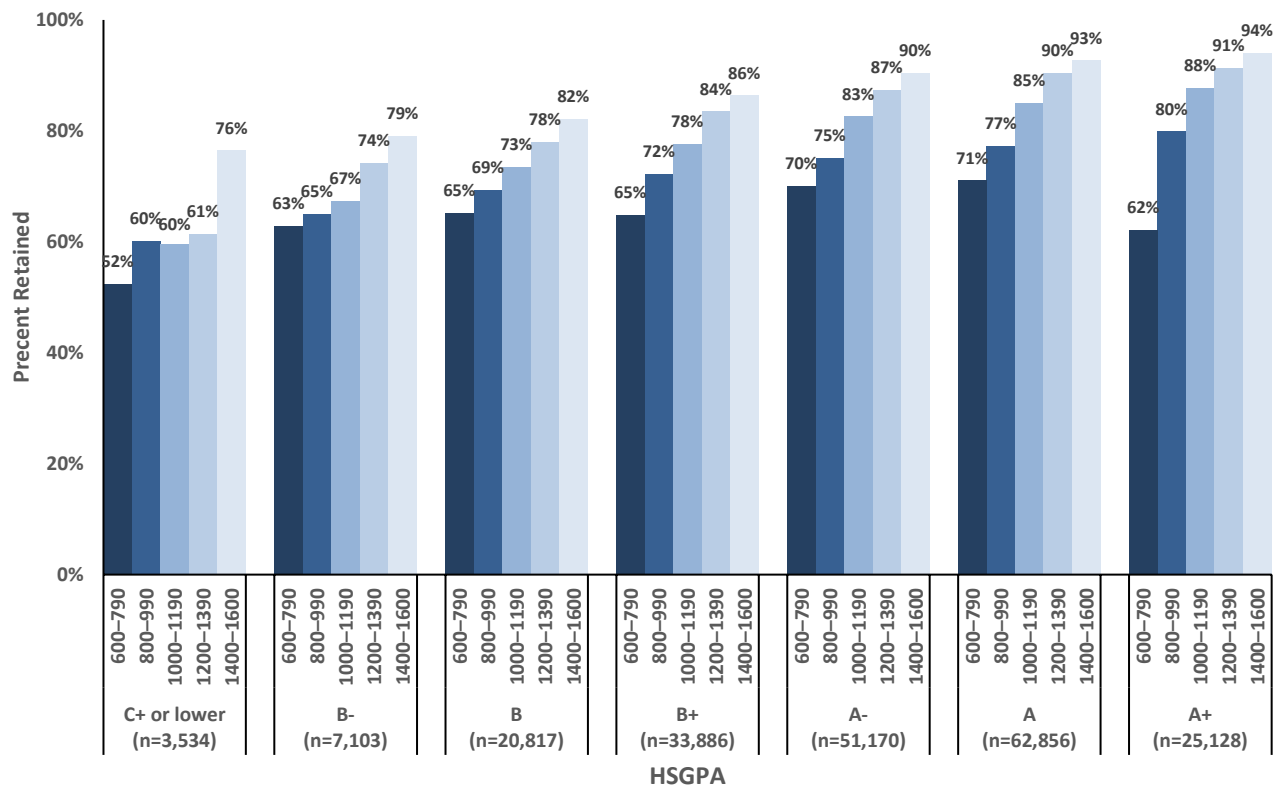
Figure 4: Mean Second-Year Retention Rate by SAT Total Score Bands
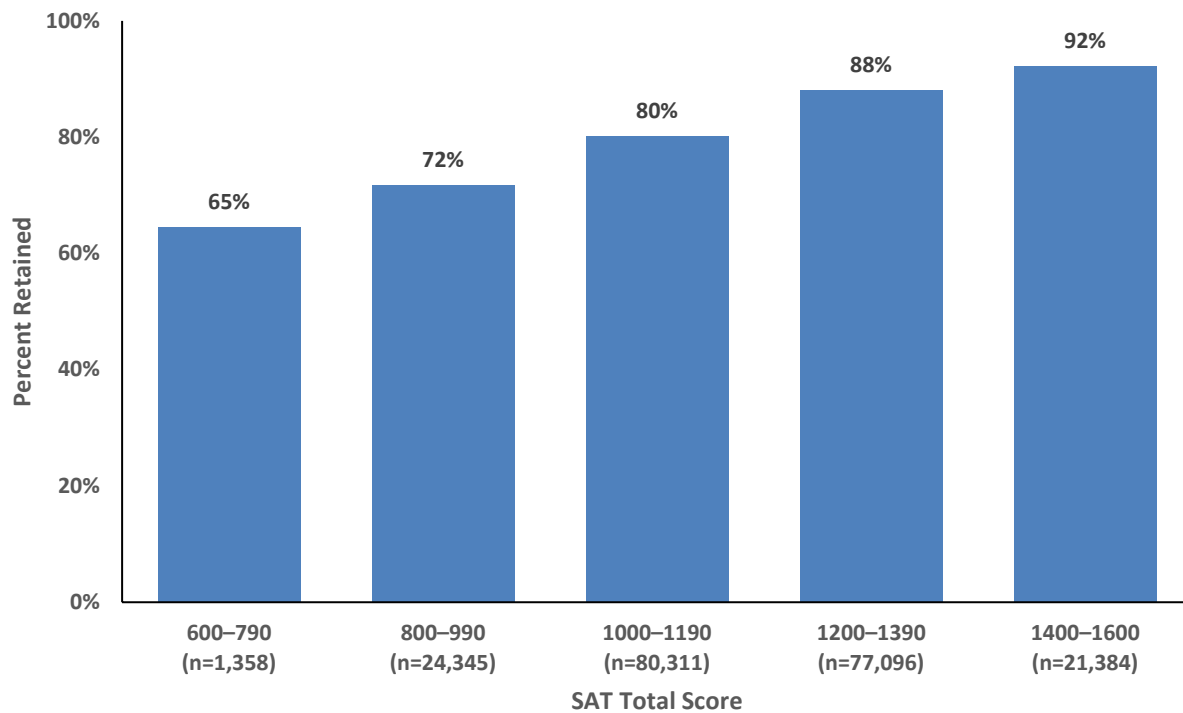
Note. Results are not reported for categories with less than 15 students (e.g., 400–590).

Figure 5 depicts second-year retention rates to the same institution when using HSGPA and SAT scores jointly. Students were first categorized by their HSGPAs, from C+ or lower to A+. Within each HSGPA category, students were further categorized by SAT score bands. The figure shows a positive relationship between SAT scores and retention across all HSGPA categories, especially for students within the A and B HSGPA categories, students who represented more than 98% of the study sample. For example, for those students with a HSGPA of A but with an SAT score between 800–990, they have an average second-year retention rate of 77%, while the same A students with an SAT score between 1400–1600 have a 93% retention rate. Combining HSGPA and SAT information reveals additional insights about student performance in college that is not evident to an institution when using either measure alone.

Figure 5: Mean Second-Year Retention Rate by HSGPA and SAT Total Score Bands

Note. Results are not reported for categories with less than 15 students (e.g., 400–590).

Whereas the retention tables and figures thus far have shown actual outcomes, SAT scores also provide information that helps institutions estimate the probability of students returning for a second year. Table 7 shows the probability of retention to the second year by SAT Total score deciles and HSGPA quintiles for students retained to the same institution. The probability of a student being retained to the same institution increases as SAT Total score deciles and HSGPA quintiles increase, showing the positive and complementary relationship between SAT scores and HSGPA with retention to the second year.

Table 7: Probability of Retention to the Same Institution by SAT Deciles and HSGPA Categories

HSGPA Category	HSGPA Quintiles	SAT Deciles									
		400–970	980–1040	1050–1100	1110–1140	1150–1190	1200–1230	1240–1280	1290–1330	1340–1400	1410–1600
A+	4.01–4.33	86%	87%	88%	89%	89%	90%	90%	91%	91%	92%
A	3.67–4.00	84%	85%	86%	87%	87%	88%	88%	89%	90%	90%
A-	3.34–3.66	82%	83%	84%	85%	85%	86%	87%	87%	88%	89%
B+	3.01–3.33	79%	81%	82%	83%	83%	84%	84%	85%	86%	87%
B or lower	0.00–3.00	74%	76%	78%	79%	79%	80%	81%	82%	83%	84%

Figure 6 illustrates the probability of a student returning to the same institution, based on HSGPA and SAT Total scores. At every HSGPA point on the scale, results indicate that SAT scores provide useful information to help institutions determine a students' probability of returning for the second year. As an example, a student with a HSGPA of 3.00 and an SAT Total score of 1000 has an approximately 78% chance of returning for the second year, while a student with the same HSGPA of 3.00 but who earned an SAT Total score of 1400 has approximately an 84% chance of returning for the second year. Understanding even these relatively small differences in retention rates alerts faculty and staff to students who may benefit from academic intervention and, as a result, stand a better chance of completing their educational goals.

Even among students with higher HSGPAs, we see the added SAT value in understanding student retention in college. For example, among students with a 4.00 HSGPA, those with an SAT Total score of 1000 have approximately an 85% probability of returning for the second year, while those with the same HSGPA but an SAT Total score of 1400 have approximately a 90% probability of returning for the second year. Although these differences are relatively small, they represent important information for colleges and universities that wish to increase student retention and completion rates.

Figure 6: Probability of Returning to the Same Institution in the Second Year given HSGPA and SAT Total Scores
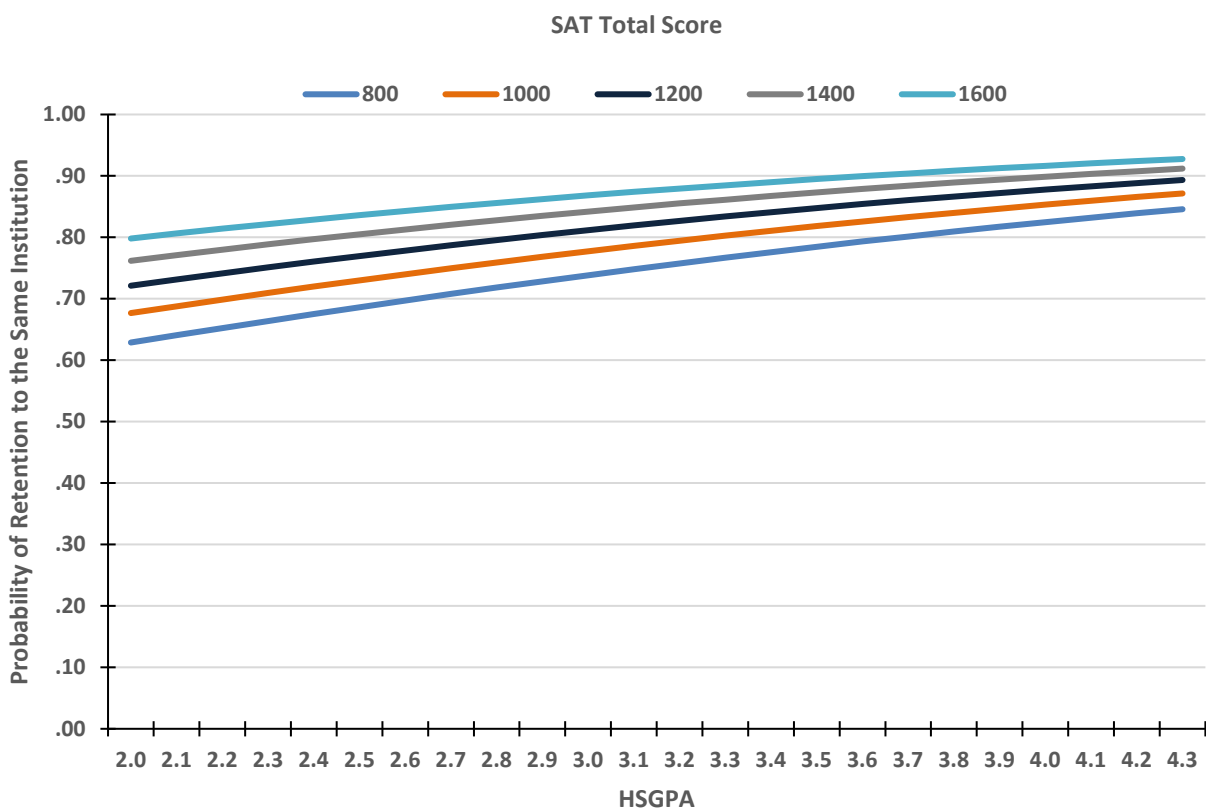
To better understand how student performance in college relates to predicted performance, student FYGPA was predicted within each institution using their HSGPA, SAT ERW section score, and SAT Math section score. The difference between a student's actual and predicted performance during the first year can be used to categorize students into two groups— those who performed as well as predicted or better, and those who underperformed from what was expected. Students who greatly underperform are students who earn grades in college that are much lower than predicted by their high school performance, and these students depart college at higher rates than do other students (Shaw & Mattern, 2013).

It is important to note that not all students classified as underperforming, and therefore at risk for departure, have a low FYGPA. In this sample, 24% of the students classified as underperforming had a FYGPA of 2.00 *or higher*, a FYGPA that many consider an acceptable minimum for avoiding academic probation. For example, a student predicted to earn a FYGPA of 3.80 but who actually earned a FYGPA of 2.25 in college would be otherwise considered in good academic standing. However, by taking account of their predicted performance (based on SAT scores and HSGPA) the institution would have information to proactively flag this student as being at risk for departure.

Figure 7 shows retention rates for students who underperform and for those who perform as well as expected or above. Accurate prediction is essential in promoting student retention at an institution. This chart shows higher retention rates when students perform as expected or better. Eighty-seven percent of students who performed as expected or above returned for the second year, while only 40% of students who underperformed returned for the second year. Accepting students and helping to position them for success based on the information one has about them (e.g. determining possible supports needed) is critical. Arriving at a predicted FYGPA for students using both HSGPA and SAT scores and storing this value in the student information system to monitor the predicted FYGPA value against the actual FYGPA can be a useful tool to identify students who may be at risk for not returning. This is especially true at less selective institutions, as the likelihood of underperforming students returning for the second year dramatically decreases as institutional admission selectivity decreases.

Figure 7: Retention Rates of Students Underperforming and Performing as Expected or Better, Total Sample and by Institutional Admittance Rate
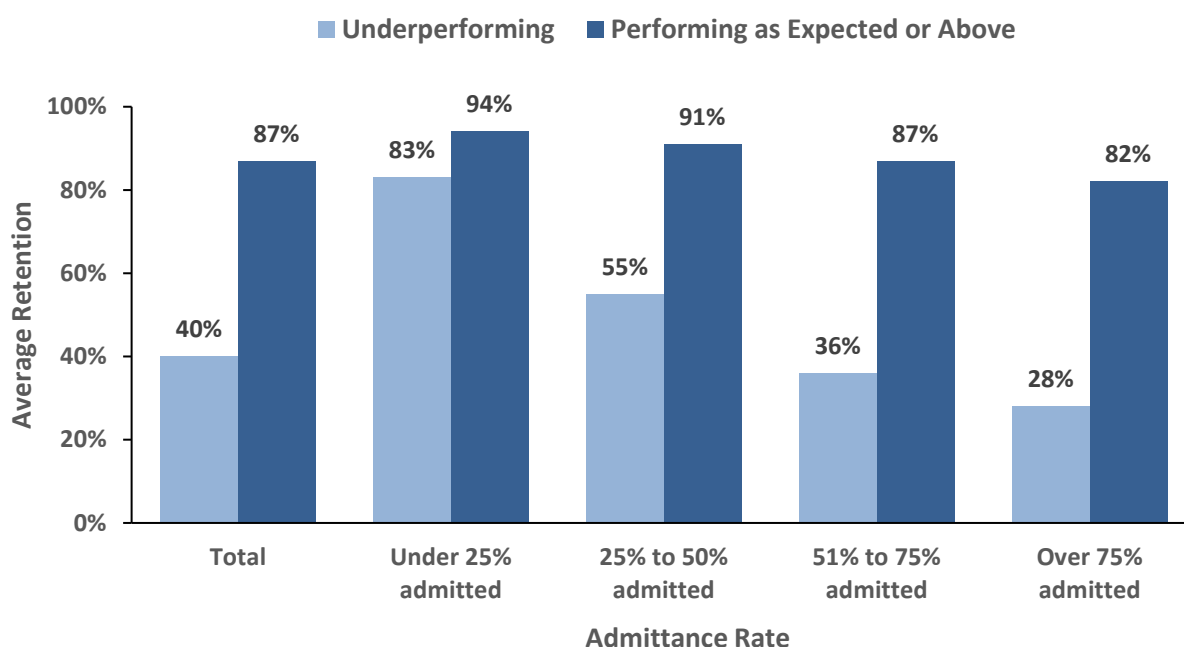
Conclusion

Findings from the current study affirm the value and effectiveness of the SAT as a tool for institutions to use to inform decisions related to admission and retention and to assist institutions in targeting instructional supports and interventions for students who may need them to be successful in their academic endeavors. This study finds that:

- SAT scores are strongly predictive of college performance—students with higher SAT scores are more likely to have higher grades in college.
- SAT scores are predictive of student retention to their second year—students with higher SAT scores are more likely to return for their sophomore year.
- SAT scores and HSGPA are both related to academic performance in college but tend to measure slightly different aspects of academic preparation. Using SAT scores in conjunction with HSGPA is the most powerful way to predict future academic performance.
  - On average, SAT scores add 15% more predictive power above grades alone for understanding how students will perform in college.
  - SAT scores help to further differentiate student performance in college within narrow HSGPA ranges.
- Colleges can use SAT scores to identify students who may be in need of academic support before they start college and throughout their college education by monitoring predicted versus actual performance and by helping to position these students for success.

20

College Board will continue to maintain a robust and ongoing national SAT validity research agenda, which will include the study of SAT score relationships with performance in particular college courses and academic domains, as well as with longer-term outcomes, including degree completion. College Board also provides a free online service for higher education institutions and systems (Admitted Class Evaluation Service, ACES) to conduct campus or system-specific validity studies (with outcomes such as FYGPA, course grades, retention, and completion) that meet their specific institutional needs.

References

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (2014). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.

Beard, J. & Marini, J. (2018). *Validity of the SAT® for Predicting First-Year Grades: 2013 SAT Validity Sample* (College Board Statistical Report). New York: College Board.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.) Hillsdale, NJ: Erlbaum.

College Board. (2015). *Test Specifications for the Redesigned SAT®*. New York: College Board.

College Board. (2017). *SAT® Suite of Assessments Technical Manual: Characteristics of the SAT*. New York: College Board.

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* (Research Report 2008–5). New York: College Board.

Kuncel, N. R., Credé, M., & Thomas, L. L. (2005). The validity of self-reported grade point average, class ranks, and test scores: A meta-analysis and review of the literature. *Review of Educational Research, 75*, 63–82.

Lawley D. N. (1943). A note on Karl Pearson's selection formulae. *Proceedings of the Royal Society of Edinburgh.*, 62(Section A, Pt. 1), 28–30.

Mattern, K. D., & Patterson, B. F. (2014). *Synthesis of recent SAT validity findings: Trend data over time and cohorts* (College Board Research in Review 2014-1). New York: The College Board.

Meyer, G. J., Finn, S. E., Eyde, L. D., Kay, G. G., Moreland, K. L., Dies, R. R., Eisman, E. J., Kubiszyn, T. W., & Reed, G. M. (2001). Psychological Testing and Psychological Assessment: A Review of Evidence and Issues. *American Psychologist, 56*(2), 128–165.

Sawyer, R. (2013). Beyond correlations: Usefulness of High School GPA and Test Scores in Making College Admissions Decisions. *Applied Measurement in Education, 26*(2), 89–112.

Shaw, E. J., Marini, J. P., Beard, J., Shmueli, D., Young, Y., & Ng, H. (2016). *The Redesigned SAT® Pilot Predictive Validity Study: A First Look* (Research Report 2016–1). New York: College Board.

Shaw, E. J., & Mattern, K. D. (2009). *Examining the Accuracy of Self-Reported High School Grade Point Average* (Research Report 2009–5). New York: College Board.

Shaw, E. J., & Mattern, K. D. (2013). Examining student under- and overperformance in college to identify risk of attrition. *Educational Assessment, 18*(4), 251–268.

Appendix A: Participating Institutions

Abilene Christian University
Appalachian State University
Austin College
Ball State University
Baylor University
Boston College
Boston University
Bucknell University
Cal Poly
Caldwell University
California State University, Fresno
Capital University
Carleton College
Carlow University
Carnegie Mellon University
Chapman University
Claremont McKenna College
Clemson University
Coastal Carolina University
Colorado College
Colorado State University
DePauw University
Earlham College
Elms College
Florida Institute of Technology
Florida State University
Fordham University
Framingham State University
Franklin & Marshall College
Furman University
Georgia College
Georgia Southern University
Gonzaga University
Grace College
Grand Valley State University
Grinnell College
Harvey Mudd College
Indiana University Bloomington
Indiana University East
Indiana University Kokomo
Indiana University Northwest
Indiana University South Bend
Indiana University Southeast

Indiana University–Purdue University
Indianapolis
Jefferson (Philadelphia University + Thomas
Jefferson University)
Lasell College
Lawrence University
Lewis & Clark College
Linfield College
Lock Haven University
Loyola University Maryland
Lycoming College
Macalester College
Meredith College
Miami University
Missouri State University
Moravian College
Mount Aloysius College
Niagara University
North Carolina State University
Northwestern University
Oakland University
Occidental College
Portland State University
Presbyterian College
Purdue University
Queens University of Charlotte
Quinnipiac University
Radford University
Rhodes College
Rider University
Roger Williams University
Rollins College
Saint Anselm College
Saint Michael's College
Seton Hall University
Shepherd University
Skidmore College
Southeastern University
Spelman College
St. Edward's University
St. Joseph's College Long Island
St. Mary's University
St. Olaf College
Stanford University

Stetson University
Stony Brook University
SUNY New Paltz
Susquehanna University
Taylor University
Texas A&M International University
Texas A&M University
Texas State University
Texas Tech University
Texas Woman's University
The Ohio State University
University of Alaska Fairbanks
University of Arkansas
University of California, Berkeley
University of California, Davis
University of California, Irvine
University of California, Los Angeles
University of California, Riverside
University of California, San Diego
University of California, Santa Barbara
University of California, Santa Cruz
University of Cincinnati
University of Colorado, Colorado Springs
University of Denver
University of Georgia
University of Houston
University of Illinois at Chicago
University of Maine
University of Massachusetts Dartmouth
University of New Hampshire
University of North Carolina at Greensboro
University of North Texas
University of Notre Dame
University of Pittsburgh
University of Pittsburgh at Bradford
University of Rhode Island
University of San Diego
University of San Francisco
University of South Carolina
University of Southern California
University of Southern Indiana
University of Southern Maine
University of Texas at Austin

University of Texas at San Antonio
University of Texas Rio Grande Valley
University of Vermont
Ursinus College
Vanderbilt University
Washington University in St. Louis
Wesleyan University
Western Washington University
Wheaton College
Wilkes University
York College of Pennsylvania
Institution A
Institution B
Institution C
Institution D
Institution E
Institution F
Institution G
Institution H
Institution I
Institution J
Institution K
Institution L
Institution M
Institution N
Institution O
Institution P
Institution Q
Institution R
Institution S
Institution T
Institution U
Institution V
Institution W
Institution X
Institution Y
Institution Z
Institution AA
Institution AB
Institution AC
Institution AD
Institution AE
Institution AF

Note. There were 32 institutions that wished to remain anonymous, hence the list of Institutions A through AF.

Appendix B: Retention Analysis Sample and Population Institutional Characteristics

	Variable	Retention Sample (*k*=156)	Reference Population of Institutions (*k*=1,230)
U. S. Region	Midwest	34 (22%)	343 (28%)
	Mid-Atlantic	28 (18%)	246 (20%)
	New England	19 (12%)	119 (10%)
	South	26 (17%)	277 (23%)
	Southwest	18 (12%)	90 (7%)
	West	31 (20%)	155 (13%)
Control	Public	75 (48%)	417 (34%)
	Private	81 (52%)	813 (66%)
Admittance Rate	Under 25%	16 (10%)	57 (5%)
	25% to 50%	29 (19%)	211 (17%)
	51% to 75%	65 (42%)	651 (53%)
	Over 75%	46 (29%)	311 (25%)
Undergraduate Enrollment	Small	61 (39%)	761 (62%)
	Medium	28 (18%)	202 (16%)
	Large	27 (17%)	136 (11%)
	Very Large	40 (26%)	131 (11%)

Note. Percentages may not sum to 100 due to rounding. Undergraduate enrollment was categorized as follows: small: 4,999 or less; medium: 5,000 to 9,999; large: 10,000 to 19,999; and very large: 20,000 or more.

Appendix C: Retention Analysis Sample and Population Student Characteristics

		Retention Sample (n = 204,504)	2017 Graduating Seniors Who Took the SAT (N = 1,715,481)
Gender	Male	88,428 (43%)	809,462 (47%)
	Female	116,076 (57%)	906,019 (53%)
Race/Ethnicity	American Indian or Alaska Native	629 (<1%)	7,782 (<1%)
	Asian	21,193 (10%)	158,031 (9%)
	Black or African American	14,851 (7%)	225,860 (13%)
	Hispanic or Latino	42,815 (21%)	408,067 (24%)
	Native Hawaiian or Other Pacific Islander	286 (<1%)	4,131 (<1%)
	White	114,051 (56%)	760,362 (44%)
	Two or More Races	7,733 (4%)	57,049 (3%)
	Not Stated	2,946 (1%)	94,199 (5%)
Highest Parental Education Level	No High School Diploma	11,291 (6%)	137,437 (8%)
	High School Diploma	43,479 (21%)	482,194 (28%)
	Associate Degree	14,361 (7%)	134,451 (8%)
	Bachelor's Degree	74,218 (36%)	473,103 (28%)
	Graduate Degree	58,220 (28%)	339,743 (20%)
	Not Stated	2,935 (1%)	148,553 (9%)

Appendix D: Retention Analysis Sample and Population Descriptive Statistics for Measures of Interest

	Retention Sample (n=204,504)				2017 Graduating Seniors Who Took the SAT (N=1,715,481)			
	M	SD	Min	Max	M	SD	Min	Max
HSGPA	3.67	0.47	0.00	4.33	3.33	0.65	0.00	4.33
SAT Total	1186	162	400	1600	1060	195	400	1600
SAT ERW	596	82	200	800	533	100	200	800
SAT Math	590	92	200	800	527	107	200	800
FYGPA	3.03	0.81	0.00	4.30				
Note. Not all 2017 graduating seniors who took the SAT reported their HSGPA (n = 1,594,136).								

Appendix E: Percentage of Students Retained to the Same Institution, Institutional Characteristics

	Variable	Percentage
U. S. Region	Midwest	81%
	Mid-Atlantic	84%
	New England	83%
	South	85%
	Southwest	81%
	West	85%
Control	Public	83%
	Private	85%
Admittance Rate	Under 25%	93%
	25% to 50%	88%
	51% to 75%	83%
	Over 75%	78%
Undergraduate Enrollment	Small	80%
	Medium	79%
	Large	82%
	Very Large	85%
Overall		83%
Note: Percentages may not sum to 100 due to rounding. Undergraduate enrollment was categorized as follows: small: 4,999 or less; medium: 5,000 to 9,999; large: 10,000 to 19,999; and very large: 20,000 or more.		

Appendix F: Percentage of Students Retained to the Same Institution, Student Sample Characteristics

	Variable	Percentage
Gender	Male	82%
	Female	84%
Race/Ethnicity	American Indian or Alaska Native	76%
	Asian	87%
	Black or African American	79%
	Hispanic or Latino	81%
	Native Hawaiian or Other Pacific Islander	77%
	White	84%
	Two or More Races	85%
	Not Stated	78%
Highest Parental Education Level	No High School Diploma	79%
	High School Diploma	78%
	Associate Degree	79%
	Bachelor's Degree	85%
	Graduate Degree	87%
	Not Stated	76%
Overall		83%
Note. Retention information came from the National Student Clearinghouse (NSC) and could differ from individual institution retention rates by a small amount as students are not always captured in every year of NSC data. Percentages are not shown where there are fewer than 15 students in the cell.		