JULY 2019

# Differential Validity and Prediction of the SAT®

Examining First-Year Grades and Retention to the Second Year

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

# Contents

# Abstract

This study examines the validity of the current SAT® as a predictor of first-year academic performance and retention to the second year by student and institutional subgroups across more than 223,000 students from 171 four-year institutions. Results show that institutions can feel confident using SAT scores and HSGPA for admission, scholarship, and advising/retention decisions across various student and institutional subgroups. Similar to previous research, we found that the SAT and HSGPA tend to have slightly stronger predictive relationships with FYGPA for female students, Asian and white students, students with higher parental education levels, and students whose best language is English Only. Across institutional subgroups, SAT and HSGPA tended to have slightly stronger predictive relationships with FYGPA at private institutions and small institutions. Also, compared to HSGPA, SAT scores tended to have stronger relationships with FYGPA at more selective institutions. However, the reverse was true at less selective institutions. As previously found, SAT scores and HSGPA tended to overpredict FYGPA for underrepresented minority students, male students, and students with lower parental education levels; SAT and HSGPA slightly underpredict FYGPA for female students. SAT scores showed clear positive relationships with retention to the second year across all student and institutional subgroups examined. Additional retention analyses indicated that monitoring student underperformance (calculated using the difference between the actual FYGPA and a predicted FYGPA based on SAT scores and HSGPA) can be a useful approach in identifying which students may be less likely to return, across all student and institutional subgroups. In general, we find that the utility of the SAT, and its added informational value above HSGPA to predict FYGPA and retention, holds across the student and institutional subgroups examined in this study.

# Introduction

This study examines the validity of current SAT scores as a predictor of first-year academic performance and retention to the second year by relevant student and institutional subgroups. This study follows the first national operational validity study of the current SAT, which found that SAT scores strongly predict first-year grade point average (FYGPA) and retention to the second year across more than 223,000 students at 171 four-year institutions (Westrick, Marini, Young, Ng, Shmueli, & Shaw, 2019). This initial study also found that using SAT scores with high school grade point average (HSGPA) to predict FYGPA and retention proved better than using either the SAT or HSGPA alone. These findings are consistent with previous research on the SAT and other standardized tests (Mattern & Patterson, 2014; Zwick, 2006). Analyzing validity results by relevant subgroups is an important part of understanding whether the relationships between test scores and the outcomes they are intended to predict hold across groups or if there are differences that may require further investigation to understand. The analyses that follow represent a key contribution to the national SAT validity research agenda.

## Previous Research

For the prediction of FYGPA, validity evidence usually takes two forms: validity coefficients – the correlation coefficients between predictors and FYGPA – and regression analyses.

Correlation coefficients are statistical indices of the linear relationship between two variables, the predictor and the criterion. Correlations range from -1 to +1, with a correlation of ±1 indicating a perfect linear relationship and a correlation of zero indicating no relationship. Correlations with absolute values greater than or equal to .1 and less than .3 are considered "small"; correlations with absolute values greater than or equal to .3 and less than .5 are considered "medium"; and correlations of .5 or higher are considered "large" (Cohen, 1988).

The term "differential validity" is used when examining whether there are significant differences in the validity coefficients calculated for different groups of examinees (Young, 2001). Previous research on differential validity has produced consistent findings, at least for gender and race/ethnicity. Regarding gender, the relationship between test scores and subsequent academic performance tend to be slightly higher for female students than they are for male students (Breland, 1979; Mattern & Patterson, 2014; Ramist, Lewis, & McCamley-Jenkins, 1994; Young, 2001; Zwick, 2006). Research on differential validity for different racial/ethnic groups has found that correlations between test scores and academic performance tend to be lower for African American and Hispanic or Latino students than they are for Asian and white students (Mattern & Patterson, 2014; Ramist et al., 1994; Young, 2001; Zwick, 2006, 2019). Previous research on differential validity for students whose first language is not English is inconclusive, with results varying depending on how language groups have been defined (Zwick, 2006).

"Differential prediction" refers to systematic differences in the accuracy of predicting a criterion (college performance) across student groups using a common prediction equation. The focus has traditionally been on the overprediction and underprediction of a group's performance when using the regression equation calculated based on the overall sample (Young, 2001). Differences between a group's actual performance and the group's predicted performance (e.g., actual FYGPA minus predicted FYGPA), also

known as the residuals, may be negative or positive. Negative residuals indicate overprediction—the group's actual FYGPA was lower than their predicted FYGPA that was calculated using the overall regression equation for the pooled sample. Positive residuals indicate under-prediction—the group's actual FYGPA was higher than the group's predicted FYGPA that was calculated using the overall regression equation for the pooled sample.

As with the findings from research on differential validity, the findings from previous research on differential prediction have also been consistent, especially regarding gender and race/ethnicity. Test scores tend to overpredict male students' later academic performance and underpredict female students' later academic performance (Breland, 1979; Linn, 1973; Mattern & Patterson, 2014; Ramist et al., 1994; Sawyer, 1986; Young, 2001; Zwick, 2006, 2019), and test scores tend to overpredict academic performance for African Americans and Hispanics/Latinos and to underpredict academic performance for Asian and white students (Breland, 1979; Linn, 1973; Mattern & Patterson, 2014; Maxey & Sawyer, 1981; Ramist et al., 1994; Young, 2001; Zwick, 2006). Overprediction for African Americans and Hispanic/Latinos is often worse when HSGPA is used alone to predict future academic performance. The joint use of test scores and HSGPA usually lessens the overprediction (Ramist et al., 1994; Zwick, 2006; Zwick & Sklar, 2005; Zwick & Schlemer, 2004).

The possible reasons and theories for seeing differences in the test score criterion relationship by subgroup have been hypothesized and studied by many, but typically result in an overarching conclusion that there is a combination of factors that can impact these relationships and that it's difficult to isolate the precise source(s). Some studies have noted that differences can be due to an unmeasured difference or omitted variable (e.g., quality of early schooling or home environment) that can differentially impact results, differences in the collegiate experiences (e.g., adjustment, fit) of students in subgroups that differentially impact their college grades, differences in coursework pursued by subgroups that may have more lenient or stringent grading practices associated, and many other theories ranging from measurement error to affirmative action admission policies (Zwick, 2002, 2019).

Research on retention rates for different student subgroups has not received as much attention as the research on the prediction of FYGPA. Overall analyses have found that most students return for their second year of college, but students with higher SAT scores and HSGPAs are more likely to return for their second year (Mattern & Patterson, 2014).[1] Even among students within the same HSGPA category, students with higher SAT scores were more likely to return (Mattern & Patterson, 2014). However, students are less likely to return to their institution when their FYGPA is significantly higher or lower than predicted (Shaw & Mattern, 2013).

Regarding retention for specific subgroups, results have been stable and not sample specific. Retention analyses have found differences between retention rates for race/ethnicity groupings, students whose

---

[1] Research has also shown that first-year academic performance is the strongest predictor of second-year retention (Pascarella & Terenzini, 2005; Westrick, Le, Robbins, Radunzel, & Schmidt, 2015), highlighting the importance of first-year academic performance, which is best predicted by the joint use of admission test scores and high school grades.

parental education levels differ, and retention rates disaggregated by institution size. Specifically, retention rates tend to be higher for white and Asian students and for students whose parents have bachelor's or graduate degrees. However, these differences largely disappear or reverse when controlling for SAT score bands (Mattern & Patterson, 2014), meaning that differences in student performance drive these relationships or patterns more than membership in particular groups. Similar patterns have been observed when disaggregating institutional results by institution size and students' SAT scores.

Small, but consistent results have been found when examining results broken out by gender, institutional control (public/private), and institutional admission selectivity. Overall and across all SAT score ranges, retention rates tend to be slightly higher for female students than those for male students, and retention rates are slightly higher for private institutions than they are for public institutions (Mattern & Patterson, 2014). Across institutional admission selectivity levels, students with higher SAT scores are more likely to return for their second year. More selective institutions tend to have higher retention rates than less selective institutions across all SAT score ranges (Mattern & Patterson, 2014).

The results in this report are divided into two sections. The first section focuses on SAT score relationships with FYGPA and second-year retention for student subgroups broken out by gender, race/ethnicity, best language, and parental education. The second section concentrates on the same relationships broken out by institutional subgroups—control (public/private), size, admission selectivity, and control by admission selectivity. The student and institutional subgroups examined in this study are those considered to be of key relevance and of historical interest in the differential validity and prediction literature (see Mattern & Patterson, 2014; Ramist et al., 1994; Young, 2001).

## 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 the national SAT validity research study. Participating 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.

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 1 for more information about the student sample and the population of 2017 graduating seniors who took the redesigned SAT. Compared to the population of SAT test takers, 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.

Table 1: Comparison of the 2017 SAT Validity Study Sample and 2017 Graduating Seniors with SAT Scores

| | Subgroup | FYGPA Sample | 2017 Graduating Seniors Who Took the SAT |
|---|---|---|---|
| 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%) |
| Overall | | 223,858 | 1,715,481 |

The sample was diverse with regard to the region of the U.S., institutional control (public/private), selectivity, and size. Comparing to the population of four-year institutions, the study sample included more public institutions, more selective institutions, and more "large" and "very large" institutions. See Table 2 for these comparisons.

Table 2: Comparison of Institutional Sample to Population

| | Subgroup | FYGPA Sample of Institutions ($k$=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 Size | Small | 67 (39%) | 131 (11%) |
| | Medium | 30 (18%) | 136 (11%) |
| | Large | 30 (18%) | 202 (16%) |
| | Very Large | 44 (26%) | 761 (62%) |

Note. $k$ = number of institutions. Percentages may not sum to 100 due to rounding. Undergraduate enrollment size 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.

## 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+). Note that the inclusion of self-reported HSGPA is consistent with previous admission test validity studies (e.g., Mattern & 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). 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).

*SAT Questionnaire Responses.* Self-reported gender, race/ethnicity, language they know best, and highest parental education level were obtained from the SAT Questionnaire that each student completed during registration for the SAT.

## Subgroups of Interest

### Student Level Subgroups

The student level subgroups that were explored in this study were gender, race/ethnicity, best language, and highest parental education level. These data were self-reported and collected from the SAT Questionnaire completed by each student as part of the registration process for the SAT. Highest Parental Education Level was identified based on the highest level of education that any parent had completed, as reported by the student.

### Institutional Level Subgroups

The institutional level subgroups that were explored in this study were control (public/private), admittance rate (selectivity), size, and the intersection of control and admittance rate (e.g., a highly selective public institution).

## Analyses

### Differential Validity

Differential validity was evaluated by computing the correlations between SAT scores and HSGPA with FYGPA. If correlations vary by subgroup, then a test is said to show differential validity since the validity differs by group; however, small differences in validity are not typically of concern.[2] Correlations were computed between SAT scores (both section scores) and FYGPA as well as HSGPA and FYGPA for all student and institutional subgroups. The incremental validity of the SAT over HSGPA alone was also evaluated for each subgroup. All correlational analyses were conducted at the institution level and then weighted by the institutional sample size and pooled together for the entire sample. For correlations to be run for a subgroup at an institution, there had to be at least 15 students within that subgroup. If that was not the case, the institution was removed from that specific subgroup analysis. Correlations were corrected for multivariate range restriction (Lawley, 1943) using the 2017 graduating seniors who took the SAT as the reference population. Raw correlations can be found in the Appendices.

---

[2] While we value correlations as useful indicators of the strength of the relationships between predictors and outcomes, differential prediction is generally considered more important for subgroup-related analyses as it is more directly related to the selection of applicants (see Linn, 1982; Young, 2004). Also, groups can have differing correlation coefficients but the same prediction system, simply because one group has greater variability than the other (Linn, 1978).

## Differential Prediction

Differential prediction by student subgroup was evaluated based on regression equations to predict FYGPA using SAT scores, HSGPA, and SAT scores and HSGPA together. First, overall regressions were run within institutions and the residuals—actual FYGPA minus predicted FYGPA—for individual students were calculated. Next, average residuals were calculated by subgroup (if n≥15) across all institutions. A negative mean residual indicates that the predictors overestimate FYGPA for students within the subgroup, on average. A positive mean residual indicates that the predictors underestimate FYGPA for students within the subgroup, on average. Note that differential prediction analyses are not meaningfully analyzed by institutional subgroups (e.g., public and private) since regressions are run at the institution level and the mean residual for a given institution is zero by definition; therefore, only student subgroups were analyzed in the differential prediction analyses.

## Retention Analyses

Retention analyses examined the relationships between SAT score bands and average rate of retention to the second year at the same institution by student and institutional subgroups. Additional analyses examined the relationships between overperformance and underperformance and retention to the second year for each of the subgroups. Students were classified into two categories: "Performing as Expected or Overperforming" or "Underperforming." Each student's FYGPA was predicted using both HSGPA and SAT section scores within an institution. That predicted FYGPA was compared to a 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

## Student Subgroups

### Descriptive Statistics

Table 3 shows descriptive statistics of the study variables by student subgroups. On average for this sample, males scored higher on the SAT ERW and Math section than females. Females in this sample, on average, earned higher HSGPAs and FYGPAs and are slightly more likely to be retained to the second year. In the *2017 SAT Suite of Assessments Annual Report*, females had higher average SAT ERW section scores compared to males (534 compared to 532, respectively) and males had higher SAT Math section scores (538 compared to 516, respectively), and SAT Total scores (1070 compared to 1050, respectively) than females (College Board, 2017). However, unlike the *2017 SAT Suite of Assessments Annual Report*, which details data on all SAT takers, the sample in this report represents students who were accepted to

and enrolled in the first year of a four-year institution, resulting in a higher scoring sample with slightly different demographic characteristics[3].

In terms of race/ethnicity, Asian, white, and Two or More races students tended to score higher on both sections of the SAT and also achieve higher HSGPAs, FYGPAs, and retention rates than students of other race/ethnicities. This follows trends in previous research (Beard & Marini, 2018; Mattern, Patterson, Shaw, Kobrin, & Barbuti, 2008) and is consistent with the *2017 SAT Suite of Assessments Annual Report* (College Board, 2017). It should be noted that the SAT race/ethnicity reporting changed for the class of 2016 and now includes a separate category for "Native Hawaiian or Other Pacific Islander" and a category for "Two or More Races," which can make comparisons to older studies inappropriate as the makeup of all groups has changed.

In terms of best language, students in this study who reported "English Only" as their best language tended to have higher SAT ERW section scores, FYGPA, and rates of retention to the second year than the other groups. Those reporting "Another Language" as their best language have higher SAT Math section scores and HSGPAs than the other groups. In terms of the highest parental education level subgroups, in general, as education level increases, so do average scores on both SAT sections, HSGPA, FYGPA, and second-year retention rates for the students in this study. These findings are generally consistent with previous research (Beard & Marini, 2018).

---

[3] According to the National Center for Education Statistics, since 1976, college enrollment has increased for white, black, and Hispanic students with females increasing more than males within each group (Aud, Fox, & KewalRamani, 2010). The current sample is made up of more females than males, and within each gender, the percentage of underrepresented minorities is higher for females than for males.

## Table 3: Descriptive Statistics of Study Variables by Student Subgroups

| | Student Subgroup | k | n | SAT ERW M | SAT ERW SD | SAT Math M | SAT Math SD | HSGPA M | HSGPA SD | FYGPA M | FYGPA SD | Retention M | Retention SD |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Gender | Male | 169 | 97,080 | 602 | 83 | 613 | 92 | 3.62 | 0.50 | 2.91 | 0.86 | .82 | .38 |
| | Female | 171 | 126,778 | 591 | 82 | 574 | 89 | 3.71 | 0.45 | 3.11 | 0.76 | .84 | .36 |
| Race/Ethnicity | American Indian or Alaska Native | 129 | 658 | 559 | 77 | 550 | 85 | 3.56 | 0.51 | 2.72 | 0.92 | .76 | .43 |
| | Asian | 167 | 25,209 | 620 | 82 | 658 | 93 | 3.77 | 0.41 | 3.17 | 0.68 | .87 | .34 |
| | Black or African American | 171 | 16,004 | 548 | 82 | 528 | 85 | 3.46 | 0.54 | 2.68 | 0.88 | .79 | .41 |
| | Hispanic or Latino | 171 | 47,175 | 568 | 81 | 558 | 86 | 3.64 | 0.48 | 2.84 | 0.83 | .81 | .40 |
| | Native Hawaiian or Other Pacific Islander | 90 | 319 | 582 | 79 | 575 | 89 | 3.59 | 0.51 | 2.75 | 0.86 | .77 | .42 |
| | White | 170 | 122,750 | 608 | 78 | 598 | 86 | 3.69 | 0.47 | 3.12 | 0.78 | .84 | .36 |
| | Two or More Races | 169 | 8,548 | 615 | 80 | 606 | 91 | 3.70 | 0.46 | 3.03 | 0.82 | .85 | .36 |
| | Not Stated | 171 | 3,195 | 570 | 88 | 562 | 98 | 3.52 | 0.52 | 2.84 | 0.89 | .78 | .42 |
| Best Language | English Only | 171 | 184,083 | 601 | 81 | 592 | 90 | 3.67 | 0.47 | 3.05 | 0.81 | .84 | .37 |
| | English and Another | 171 | 35,237 | 575 | 85 | 580 | 101 | 3.66 | 0.47 | 2.91 | 0.81 | .82 | .38 |
| | Another Language | 153 | 3,458 | 562 | 91 | 653 | 116 | 3.70 | 0.46 | 3.04 | 0.78 | .77 | .42 |
| | Not Stated | 163 | 1,080 | 568 | 92 | 568 | 107 | 3.52 | 0.54 | 2.87 | 0.89 | .78 | .41 |
| Highest Parental Education Level | No High School Diploma | 169 | 12,850 | 539 | 77 | 541 | 88 | 3.60 | 0.49 | 2.74 | 0.83 | .79 | .41 |
| | High School Diploma | 171 | 48,127 | 566 | 78 | 559 | 87 | 3.60 | 0.49 | 2.82 | 0.87 | .78 | .41 |
| | Associate Degree | 171 | 15,659 | 572 | 77 | 561 | 83 | 3.62 | 0.49 | 2.89 | 0.86 | .79 | .41 |
| | Bachelor's Degree | 171 | 80,465 | 603 | 76 | 599 | 88 | 3.68 | 0.46 | 3.10 | 0.76 | .85 | .36 |
| | Graduate Degree | 171 | 63,539 | 631 | 78 | 625 | 91 | 3.74 | 0.45 | 3.20 | 0.72 | .87 | .33 |
| | Not Stated | 168 | 3,218 | 539 | 85 | 539 | 96 | 3.46 | 0.53 | 2.70 | 0.92 | .76 | .43 |
| Overall | | 171 | 223,858 | 596 | 83 | 591 | 93 | 3.67 | 0.47 | 3.03 | 0.81 | .83 | .37 |

Note: $n$ = subgroup sample size, $k$ = number of institutions. Retention data are available for $n$ = 204,504 students.

## Differential Validity

Table 4 presents the corrected correlations for SAT ERW section, SAT Math section, SAT ERW and SAT Math sections together, HSGPA, and both SAT sections with HSGPA all with FYGPA for student subgroups. Raw correlations for student subgroups can be found in Table A 1. The findings are summarized below.

### *Gender*

Results show that the SAT is slightly more predictive for females than for males. The correlations between FYGPA and the SAT sections, including the multiple correlation of the two SAT sections with FYGPA, range from .53 to .56 for females and from .47 to .51 for males. As for HSGPA and FYGPA, the results are similar with a larger correlation for females ($r$=.53) than for males ($r$=.51). Combining both SAT sections and HSGPA results in a multiple correlation of .63 for females and .59 for males. These findings are similar to those reported in previous research (Beard & Marini, 2018; Mattern et al., 2008).

### *Race/Ethnicity*

The results for race/ethnicity show that for the individual SAT sections, the SAT has the strongest correlation with FYGPA among Asian students ($r$=.51). Correlations tend to be lower for underrepresented minority groups ranging from .39 to .45. The multiple correlation of the two SAT sections together is the highest for American Indian or Alaska Native students. However, this result should be interpreted with caution due to the smaller sample size of 220 students across 8 institutions.[4] As for HSGPA, three groups share the highest correlation—American Indian or Alaska Native, Asian, and Two or More races. Finally, the multiple correlation between both SAT sections and HSGPA with FYGPA is the strongest for American Indian or Alaska Native students, which once again should be interpreted with caution. The second strongest relationship is found among white students and students reporting Two or More Races ($r$=.62). General trends across race/ethnicity groups are similar to previous research with Asian and white students tending to have among the strongest correlations and underrepresented minority students tending to have lower correlations with FYGPA. (Beard & Marini, 2018; Mattern et al., 2008).

### *Best Language*

For best language analyses, the correlations between each section score and FYGPA range from .40 to .49, with the strongest relationship found among students who identify their best language as English Only. Similar to previous research, section score correlations for students reporting English and Another Language as their best language are in the middle of the three language groups with both section correlations with FYGPA equal to .46, and the lowest correlations were found among those in the Another Language group, ranging from .40 to .41. For all correlations between predictors and FYGPA, including multiple correlations, the relationship is the strongest for those students identifying their best language as English Only (excluding those in the Not Stated category who have an SAT and HSGPA

---

[4] As noted in the table, institutions are included in the correlation calculation if there are at least 15 students within that subgroup at that institution. Therefore, correlations exclude any institutions with 14 or fewer students in the group. This value of 15 was chosen based on previous research (Mattern et al., 2008).

correlation with FYGPA of .66, but a small sample size and therefore should be interpreted with caution). General trends across reported best language groups are consistent with previous research (Beard & Marini, 2013; Mattern et al., 2008).

*Highest Parental Education Level*
All correlations tend to increase as parental education level increases. Correlations are the highest for students who have a parent with a Graduate Degree with correlations ranging from .48 for the SAT Math section with FYGPA to .63 for the multiple correlation between both SAT sections and HSGPA with FYGPA. This is compared to students with parents without a high school diploma, with their correlations ranging from .42 to .54. General trends as parental education level increases are consistent with previous research (Beard & Marini, 2018).

## Table 4: Corrected Correlations of SAT Scores and HSGPA with FYGPA by Student Subgroups

| | Student Subgroup | *k* | *n* | SAT ERW | SAT Math | SAT | HSGPA | SAT, HSGPA |
|---|---|---|---|---|---|---|---|---|
| Gender | Male | 169 | 97,080 | .47 | .48 | .51 | .51 | .59 |
| | Female | 171 | 126,778 | .53 | .53 | .56 | .53 | .63 |
| Race/Ethnicity | American Indian or Alaska Native | 8 | 220 | .41 | .39 | .57 | .50 | .68 |
| | Asian | 125 | 24,858 | .51 | .51 | .55 | .50 | .61 |
| | Black or African American | 142 | 15,741 | .40 | .39 | .44 | .44 | .52 |
| | Hispanic or Latino | 156 | 47,036 | .45 | .42 | .47 | .47 | .54 |
| | Native Hawaiian or Other Pacific Islander | - | - | - | - | - | - | - |
| | Two or More Races | 116 | 8,077 | .50 | .48 | .53 | .50 | .62 |
| | White | 169 | 122,749 | .47 | .46 | .50 | .57 | .62 |
| | Not Stated | 78 | 2,635 | .39 | .36 | .45 | .40 | .55 |
| Best Language | English Only | 171 | 184,083 | .49 | .48 | .51 | .55 | .62 |
| | English and Another | 154 | 35,057 | .46 | .46 | .49 | .47 | .56 |
| | Another Language | 47 | 2,854 | .41 | .40 | .46 | .39 | .53 |
| | Not Stated | 17 | 365 | .45 | .45 | .55 | .51 | .66 |
| Highest Parental Education Level | No High School Diploma | 86 | 12,284 | .44 | .42 | .47 | .44 | .54 |
| | High School Diploma | 166 | 48,081 | .44 | .43 | .46 | .49 | .56 |
| | Associate Degree | 142 | 15,399 | .44 | .42 | .47 | .51 | .58 |
| | Bachelor's Degree | 171 | 80,465 | .48 | .47 | .50 | .55 | .61 |
| | Graduate Degree | 170 | 63,526 | .50 | .48 | .53 | .56 | .63 |
| | Not Stated | 72 | 2,630 | .37 | .38 | .46 | .44 | .56 |
| Overall | | 171 | 223,858 | .49 | .47 | .51 | .53 | .61 |

Note: *n* = subgroup sample size, *k* = number of institutions. Correlations are only calculated at institutions with at least 15 students in a subgroup. SAT indicates the multiple correlation of both SAT sections together.

## Differential Prediction

The results for the differential prediction analyses are presented in Table 5. The sections below describe the findings by each student subgroup. Results tend to be consistent with previous SAT and HSGPA differential prediction research (Mattern et al., 2008).

### Gender

The results by gender indicate that SAT scores tend to underpredict FYGPA for female students with mean residuals ranging from 0.10 to 0.13. The same pattern of underprediction of FYGPA for female students is also evident for HSGPA (0.06) and for the combination of both SAT section scores and HSGPA. This pattern is consistent with previous research (Mattern et al., 2008). See Figure A 1 in the Appendix for a visualization of this relationship.

### Race/Ethnicity

The differential prediction results by race/ethnicity are consistent with findings from previous research showing that underrepresented minority students tend to earn lower grades in college than those predicted by admission tests and high school grades (Zwick, 2019). Specifically, the FYGPAs of American Indian or Alaska Native, black or African American, Hispanic or Latino, and Native Hawaiian or Other Pacific Islander students tended to be overpredicted by all measures and combinations of measures. The amount of FYGPA overprediction tended to be larger for HSGPA than for SAT scores. For the two largest underrepresented minority groups, black or African American and Hispanic or Latino students, the mean residuals when using HSGPA alone were -0.18 and -0.10, respectively. The FYGPAs of students identifying as "Two or More Races" also tended to be overpredicted, although less so than for most of the other race/ethnicity categories on all measures and combinations of measures. The FYGPAs of white students were slightly underpredicted by all measures and combinations of measures with mean residuals ranging from 0.03 to 0.05. Students identifying as Asian also had their FYGPAs slightly underpredicted by most measures and combinations of measures (mean residuals ranging from 0.02 to 0.08), except by the SAT Math section score where FYGPA was very slightly overpredicted (mean residual = -0.01). See Figure A 2 in the Appendix for a visualization of this relationship.

### Best Language

Results for the best language subgroups show that the FYGPAs of students whose best language is English are most accurately predicted by all measures with mean residuals ranging from 0.00 to 0.01. It is meaningful to point out that this is also the largest group and by size alone their mean residual should be the closest to zero. The FYGPAs of students whose best language is English and Another Language are overpredicted by all measures and combinations of measures with mean residuals ranging from -0.07 to -0.03. The FYGPAs of students whose best language was not English were accurately predicted by HSGPA, overpredicted by SAT Math scores (mean residual of -0.11), and underpredicted (mean residuals ranging from 0.03 to 0.13) by the other measures. See Figure A 3 in the Appendix for a visualization of this relationship.

### Highest Parental Education Level

There appears to be a clear delineation of overprediction versus underprediction of FYGPA depending on parental education level. The FYGPAs of students whose parents have no high school diploma, a high

school diploma, or an associate degree are overpredicted by all measures and combinations of measures, with mean residuals tending to be closer to zero for those with parents with an associate degree. The overprediction of FYGPA was greatest when HSGPA was used alone. However, the FYGPAs of students whose parents have a bachelor's degree or graduate degree are underpredicted by all measures and combinations of measures, with mean residuals tending to be smaller (more accurate) for those with parents with a bachelor's degree. See Figure A 4 in the Appendix for a visualization of this relationship.

Though differences in overprediction and underprediction exist, it is important to keep the results in perspective. Consistent with previous research (Fischer, Schult, & Hell, 2013), we calculated standardized mean differences ($d$; Cohen, 1988) between observed FYGPA and predicted FYGPA. The absolute value of $d$ must be .20 or higher to qualify as a "small" effect size, .50 or higher to qualify as a "medium" effect size, and .80 or higher to qualify as a "large" effect size (Cohen, 1988).[5] Aside from the Native Hawaiian or Other Pacific Islander subgroup (n = 319, $d$ = -0.25, overprediction), no subgroup had an effect size with an absolute value of 0.20 or greater when using the SAT and HSGPA to predict FYGPA.[6] While we ideally want the amount of overprediction and underprediction to be close to zero, the fact that the overwhelming majority of effect sizes do not even qualify as small effect sizes suggests that the amount of overprediction and underprediction is minor.

---

[5] Note that Cohen's (1988) guidelines for standardized mean differences (.2, .5. and .8) are not the same as his guidelines for correlations (.1, .3, and .5) discussed in other sections of this report.
[6] The use of HSGPA alone did lead to overpredictions and effect sizes characterized as small for five subgroups: American Indian or Alaska Native, black or African American, Native Hawaiian or Other Pacific Islander, students whose parents did not have a high school diploma, and students who did not report their parents' highest level of education.

Table 5: Average Overprediction (-) and Underprediction (+) of FYGPA by SAT Scores and HSGPA by Student Subgroups

| | Student Subgroup | k | n | SAT ERW | SAT Math | SAT | HSGPA | SAT, HSGPA |
|---|---|---|---|---|---|---|---|---|
| Gender | Male | 169 | 97,080 | -0.13 | -0.17 | -0.16 | -0.08 | -0.11 |
| | Female | 171 | 126,778 | 0.10 | 0.13 | 0.12 | 0.06 | 0.09 |
| Racial / Ethnic Identity | American Indian or Alaska Native | 129 | 658 | -0.13 | -0.14 | -0.12 | -0.17 | -0.12 |
| | Asian | 167 | 25,209 | 0.07 | -0.01 | 0.02 | 0.08 | 0.03 |
| | Black or African American | 171 | 16,004 | -0.16 | -0.15 | -0.13 | -0.18 | -0.10 |
| | Hispanic or Latino | 171 | 47,175 | -0.06 | -0.06 | -0.04 | -0.10 | -0.05 |
| | Native Hawaiian or Other Pacific Islander | 90 | 319 | -0.21 | -0.21 | -0.20 | -0.21 | -0.17 |
| | White | 170 | 122,750 | 0.04 | 0.05 | 0.04 | 0.05 | 0.03 |
| | Two or More Races | 169 | 8,548 | -0.07 | -0.05 | -0.06 | -0.04 | -0.06 |
| | Not Stated | 171 | 3,195 | -0.09 | -0.10 | -0.08 | -0.09 | -0.06 |
| Best Language | English Only | 171 | 184,083 | 0.00 | 0.01 | 0.00 | 0.01 | 0.00 |
| | English and Another | 171 | 35,237 | -0.03 | -0.05 | -0.03 | -0.07 | -0.03 |
| | Another Language | 153 | 3,458 | 0.13 | -0.11 | 0.03 | 0.00 | 0.03 |
| | Not Stated | 163 | 1,080 | -0.07 | -0.09 | -0.06 | -0.07 | -0.03 |
| Highest Parental Education Level | No High School Diploma | 169 | 12,850 | -0.05 | -0.09 | -0.04 | -0.15 | -0.04 |
| | High School Diploma | 171 | 48,127 | -0.09 | -0.09 | -0.08 | -0.12 | -0.08 |
| | Associate Degree | 171 | 15,659 | -0.04 | -0.04 | -0.03 | -0.08 | -0.05 |
| | Bachelor's Degree | 171 | 80,465 | 0.04 | 0.04 | 0.03 | 0.05 | 0.03 |
| | Graduate Degree | 171 | 63,539 | 0.04 | 0.05 | 0.03 | 0.09 | 0.04 |
| | Not Stated | 168 | 3,218 | -0.11 | -0.15 | -0.10 | -0.16 | -0.09 |
| Overall | | 171 | 223,858 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |

Note: *n* = subgroup sample size, *k* = number of institutions. SAT indicates the multiple correlation of both SAT sections together.

## Retention Analyses

The next section of this study examines the relationship between SAT scores, HSGPA, and retention to the second year at the same institution. For these analyses, the sample was further limited to include only those students with retention information from the institutions, resulting in 204,504 students from 156 institutions in the retention sample. See Table A 2 in Appendix A for descriptive information on the study variables by student subgroup for this sample.

Table 6 shows the mean second-year retention rate for student subgroups and overall by SAT Total score bands. First examining the overall sample, as SAT scores increase, the retention rate also increases. In general, this same pattern is seen across all student subgroups, which is consistent with previous research (Mattern & Patterson, 2014).  There are a few deviations from this pattern, but they are likely the result of the smaller sample sizes in those categories and therefore less stable results. See Table A 3 for sample sizes in each of the SAT Total score bands by student subgroups.

## Table 6: Mean Second-Year Retention Rate by SAT Score Band for Student Subgroups

| | Student Subgroup | k | n | SAT Score Band 400-590 | 600-790 | 800-990 | 1000-1190 | 1200-1390 | 1400-1600 |
|---|---|---|---|---|---|---|---|---|---|
| Gender | Male | 154 | 88,428 | -- | .62 | .69 | .77 | .86 | .91 |
| | Female | 156 | 116,076 | -- | .66 | .73 | .82 | .90 | .93 |
| Race/Ethnicity | American Indian or Alaska Native | 119 | 629 | -- | -- | .68 | .74 | .86 | .88 |
| | Asian | 152 | 21,193 | -- | .65 | .78 | .83 | .89 | .90 |
| | Black or African American | 156 | 14,851 | -- | .66 | .71 | .81 | .87 | .92 |
| | Hispanic or Latino | 156 | 42,815 | -- | .67 | .73 | .80 | .87 | .92 |
| | Native Hawaiian or Other Pacific Islander | 83 | 286 | -- | -- | .53 | .74 | .89 | .90 |
| | Two or More Races | 154 | 7,733 | -- | .81 | .68 | .79 | .89 | .94 |
| | White | 155 | 114,051 | -- | .57 | .71 | .80 | .88 | .93 |
| | Not Stated | 156 | 2,946 | -- | .72 | .69 | .77 | .84 | .89 |
| Best Language | English Only | 156 | 169,319 | -- | .63 | .71 | .80 | .88 | .93 |
| | English and Another | 156 | 31,142 | -- | .68 | .75 | .81 | .87 | .89 |
| | Another Language | 138 | 3,085 | -- | .73 | .70 | .74 | .80 | .79 |
| | Not Stated | 148 | 958 | -- | .55 | .70 | .77 | .85 | .87 |
| Highest Parental Education Level | No High School Diploma | 154 | 11,291 | -- | .64 | .73 | .79 | .88 | .86 |
| | High School Diploma | 156 | 43,479 | -- | .62 | .69 | .77 | .86 | .91 |
| | Associate Degree | 156 | 14,361 | -- | .68 | .71 | .77 | .85 | .90 |
| | Bachelor's Degree | 156 | 74,218 | -- | .68 | .73 | .82 | .89 | .92 |
| | Graduate Degree | 156 | 58,220 | -- | .70 | .75 | .83 | .89 | .93 |
| | Not Stated | 153 | 2,935 | -- | .60 | .70 | .76 | .84 | .82 |
| Overall | | 156 | 204,504 | -- | .65 | .72 | .80 | .88 | .92 |

Note: *n* = subgroup sample size, *k* = number of institutions. Results are presented for 15 students or more.

The next set of analyses examines the relationship between retention to the second year and academic overperformance and underperformance for student subgroups. 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.

*Gender*

Figure 1 shows retention rates for students who underperform and for those who perform as well as expected or above by gender. This figure shows higher retention rates for both males and females when they perform as expected or above and this retention rate of 87% is the same as in the overall sample. Males have a slightly higher retention rate (42%) when underperforming compared to females (38%). However, results for both genders are very similar to the rates in the overall sample.

Figure 1: Retention Rates of Students Underperforming and Performing as Expected or Above Total Sample and by Gender
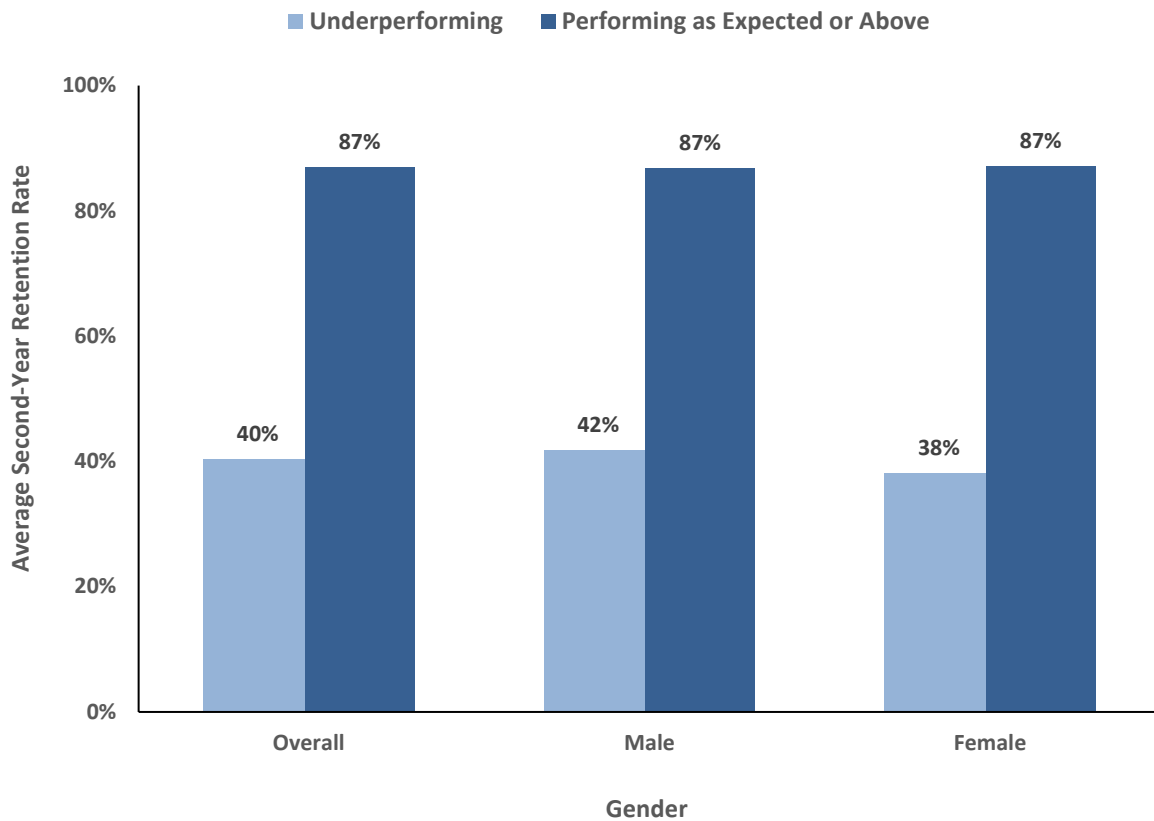
Figure 2 shows retention rates for students who underperform and for those who perform as well as expected or above by race/ethnicity. Retention rates for those performing as expected or above range from 83% to 89% across all race and ethnicities. The largest difference in retention rates between students who underperform and those who perform as expected or above is seen in American Indian or Alaska Native students (27% vs. 83%). The smallest differences are seen in Asian students (56% vs. 89%) and those identifying as Two or More Races (44% vs. 89%). Furthermore, when compared to the overall rates in the sample (40% vs. 87%), underrepresented minorities have lower rates in both categories.

Figure 2: Retention Rates of Students Underperforming and Performing as Expected or Above, Total Sample and by Race/Ethnicity

## *Best Language*

Figure 3 shows retention rates of students underperforming and performing as well as expected or above by best language. Retention rates for both the English Only category and English and Another Language are similar to that of the overall sample. Those whose best language is Another Language have the highest retention rate for students who underperform and the smallest retention difference between those who underperform and those who perform as well as expected or above.

Figure 3: Retention Rates of Students Underperforming and Performing as Expected or Above, Total Sample and by Best Language
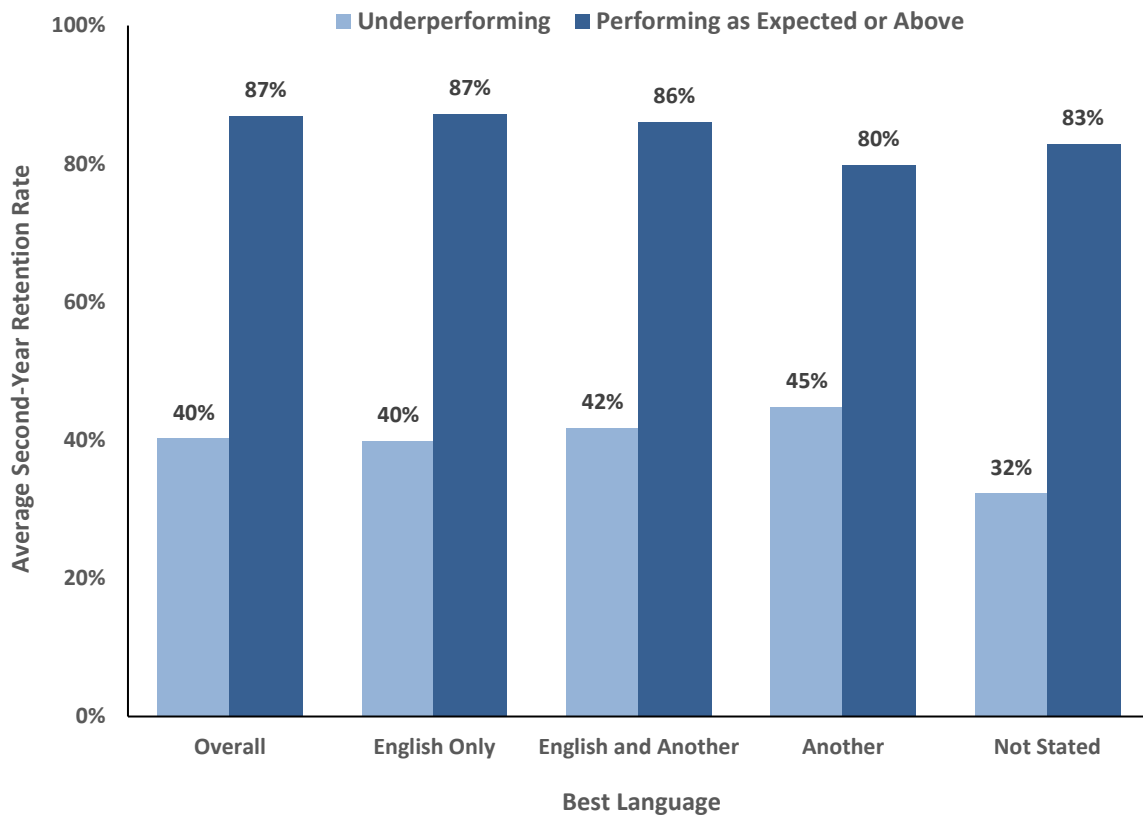


## *Highest Parental Education Level*

Figure 4 shows the retention rates for students who underperform and those who perform as well as expected or above by highest parental education level. There is a general trend that as parental education levels increase, so does that mean second-year retention rate for both groups of students. Students who have parents with bachelor's degrees (44% vs. 88%) and those with graduate degrees (50% vs. 90%) have higher mean retention rates than the full sample (40% vs. 87%) in both categories of student performance. The lowest rates are seen when a student's parents have no high school diploma (30% vs. 84%).
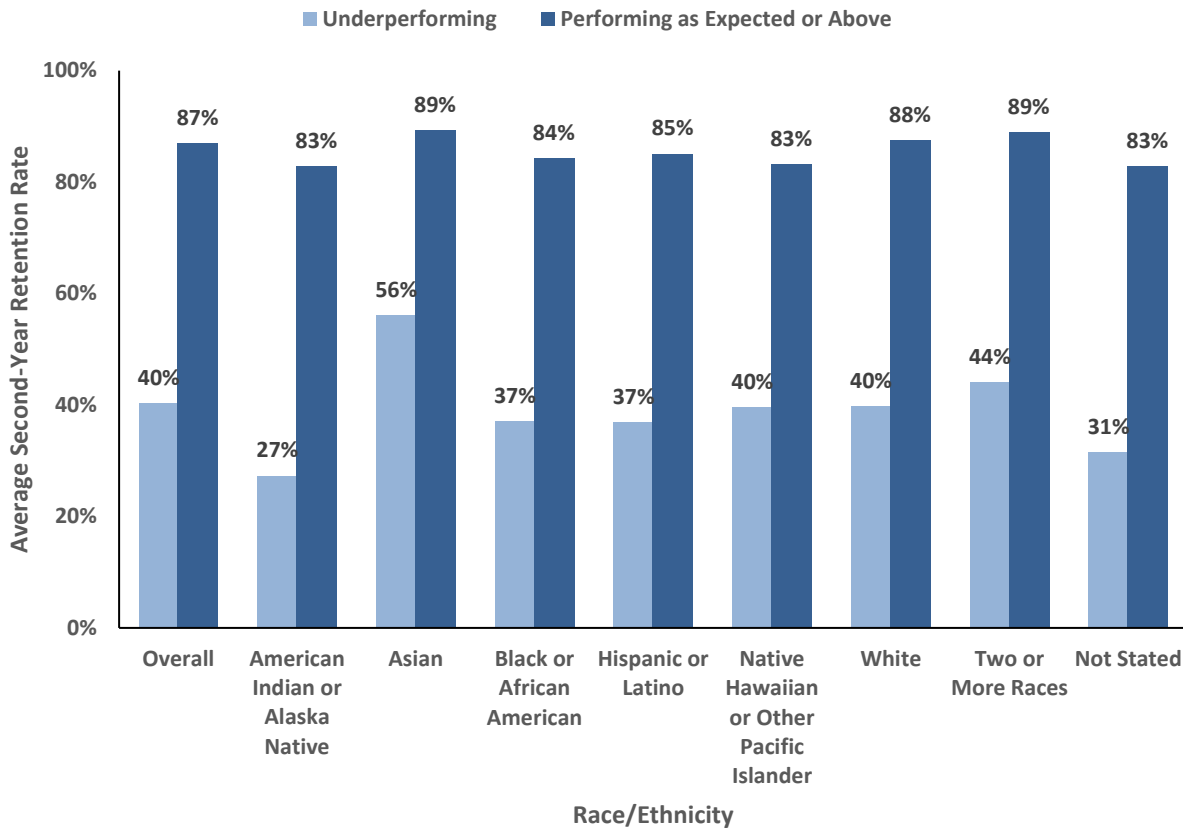
Figure 4: Retention Rates of Students Underperforming and Performing as Expected or Above, Total Sample and by Highest Parental Education Level
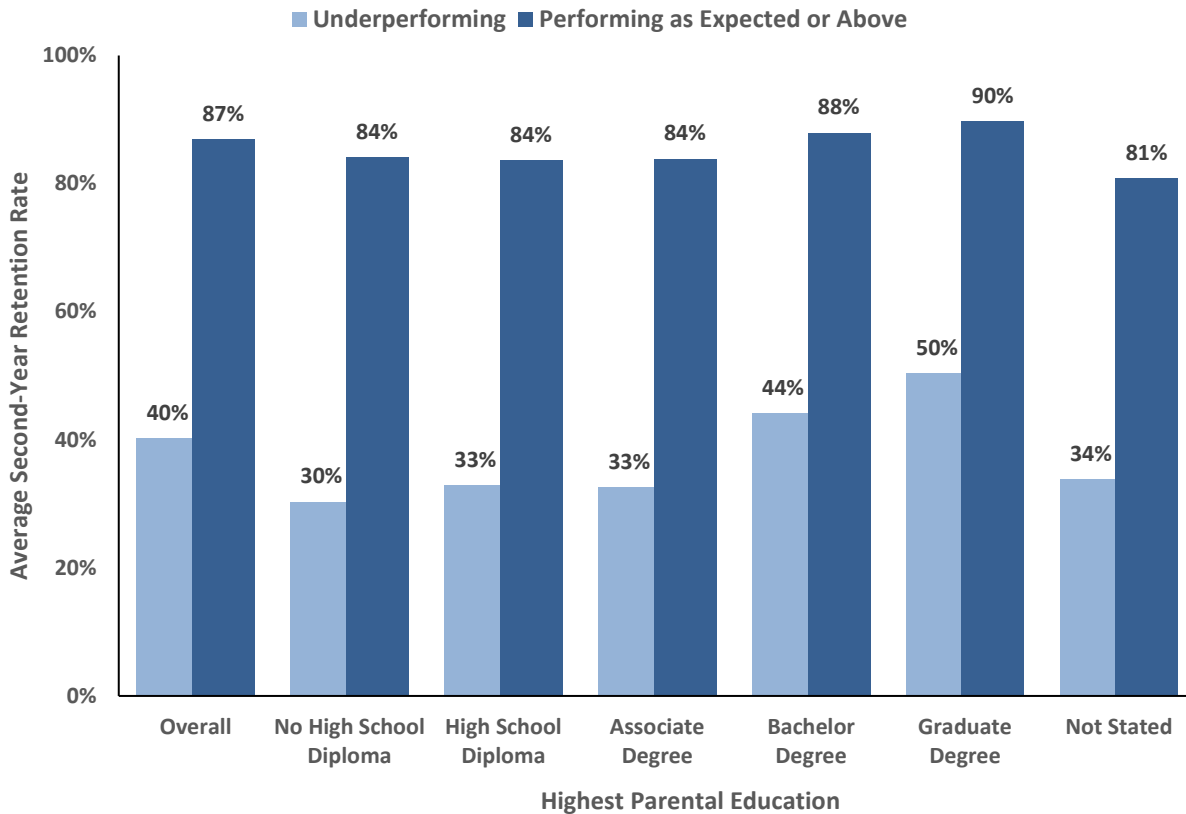


## Institutional Subgroups

After focusing on student level analyses for the first part of the study, the analyses that follow focus on differences in predictive value by institutional subgroups, including control (public/private), selectivity, size, and the intersection of control and selectivity (e.g., a highly selective public institution).

## Descriptive Statistics

Descriptive statistics of the study variables by institutional subgroups are shown in Table 7. Private institutions have higher scores on both SAT sections, HSGPAs, FYGPAs, and rates of retention to the second year compared to public institutions. As the admittance rate of an institution increases (becomes less selective), trends show that average SAT ERW section score, SAT Math section score, HSGPA, FYGPA, and rate of retention to the second year decrease, indicating that higher SAT scores, GPAs, and retention rates are seen at the more selective institutions. When factoring in control with admittance rate, you see that the private counterparts of each admittance rate category often have higher average values on all study variables than their public counterparts. The trend is not as clear in terms of institution size. Very Large institutions tend to have the highest averages on all study variables, followed by Large institutions. However, Medium-sized institutions in this sample have the lowest values across all study variables.

Table 7: Descriptive Statistics of Study Variables by Institutional Subgroups

| Institutional Subgroup | | k | n | SAT ERW | | SAT Math | | HSGPA | | FYGPA | | Retention | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | M | SD | M | SD | M | SD | M | SD | M | SD |
| Control | Private | 89 | 46,981 | 617 | 86 | 611 | 97 | 3.71 | 0.47 | 3.19 | 0.69 | .85 | .35 |
| | Public | 82 | 176,877 | 591 | 81 | 586 | 91 | 3.66 | 0.48 | 2.98 | 0.83 | .83 | .38 |
| Admittance Rate | Under 25% | 20 | 16,075 | 680 | 68 | 687 | 82 | 3.99 | 0.30 | 3.36 | 0.53 | .93 | .25 |
| | 25% to 50% | 30 | 47,268 | 623 | 76 | 629 | 88 | 3.83 | 0.38 | 3.13 | 0.67 | .88 | .32 |
| | 51% to 75% | 73 | 107,840 | 586 | 78 | 578 | 85 | 3.63 | 0.47 | 2.99 | 0.83 | .83 | .38 |
| | Over 75% | 48 | 52,675 | 566 | 78 | 555 | 83 | 3.51 | 0.51 | 2.90 | 0.89 | .78 | .42 |
| Undergraduate Enrollment Size | Small | 67 | 22,899 | 581 | 87 | 567 | 91 | 3.59 | 0.51 | 3.02 | 0.80 | .80 | .40 |
| | Medium | 30 | 23,952 | 563 | 81 | 553 | 88 | 3.48 | 0.53 | 2.92 | 0.87 | .79 | .41 |
| | Large | 30 | 43,540 | 596 | 86 | 589 | 96 | 3.62 | 0.48 | 3.02 | 0.81 | .82 | .38 |
| | Very Large | 44 | 133,467 | 604 | 79 | 603 | 90 | 3.73 | 0.44 | 3.05 | 0.79 | .85 | .36 |
| Admittance Rate x Control | Private, Under 25% | 18 | 9,556 | 695 | 60 | 704 | 71 | 3.98 | 0.31 | 3.43 | 0.49 | .93 | .26 |
| | Private, 25% to 50% | 18 | 13,697 | 629 | 74 | 626 | 86 | 3.75 | 0.41 | 3.22 | 0.65 | .87 | .34 |
| | Private, 51% to 75% | 36 | 15,740 | 583 | 77 | 569 | 83 | 3.59 | 0.48 | 3.08 | 0.73 | .83 | .38 |
| | Private, Over 75% | 17 | 7,988 | 569 | 74 | 556 | 76 | 3.54 | 0.51 | 3.07 | 0.78 | .79 | .41 |
| | Public, Under 25% | 2 | 6,519 | 658 | 74 | 663 | 91 | 4.01 | 0.28 | 3.27 | 0.56 | .94 | .23 |
| | Public, 25% to 50% | 12 | 33,571 | 620 | 76 | 630 | 89 | 3.86 | 0.35 | 3.09 | 0.67 | .89 | .32 |
| | Public, 51 to 75% | 37 | 92,100 | 587 | 78 | 579 | 85 | 3.64 | 0.47 | 2.98 | 0.85 | .83 | .38 |
| | Public, Over 75% | 31 | 44,687 | 566 | 79 | 555 | 84 | 3.50 | 0.51 | 2.87 | 0.91 | .77 | .42 |
| Overall | | 171 | 223,858 | 596 | 83 | 591 | 93 | 3.67 | 0.47 | 3.03 | 0.81 | .83 | .37 |

Note: $n$ = subgroup sample size, $k$ = number of institutions. Undergraduate enrollment size 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. Retention data are available for $n$ = 204,504 students.

## Differential Validity

Table 8 presents the corrected FYGPA correlations for the SAT ERW section, SAT Math section, SAT ERW and SAT Math sections together, HSGPA, and both SAT sections with HSGPA by institutional subgroups. Raw correlations for the institutional subgroups can be found in Table B 1. The headings below summarize the findings in each subgroup.

### *Control*

Results by institutional control show that the SAT is slightly more predictive for Private institutions than for Public institutions. For Private institutions, correlations for the two sections range from .52 to .54, with a multiple correlation of both sections together of .56. HSGPA follows the same trend, being more predictive of FYGPA at Private institutions (*r*=.56) than at Public institutions (*r*=.53). The multiple correlation of both sections and HSGPA together are also strongest for Private institutions (*r*=.65) compared to Public (*r*=.59).

### *Admittance Rate*

The SAT is the most predictive of FYGPA at the most selective institutions (Admittance Rate of 0% to 25%) with section score correlations ranging from .55 to .58 compared to the least selective institutions (Admittance Rate of Over 75%) with correlations ranging from .44 to .46. The multiple correlation of the two SAT sections scores with FYGPA is the highest for institutions with an admittance rate of 0% to 25% (*r*=.61), compared to institutions with an admittance rate of Over 75% (*r*=.48). However, HSGPA alone has the reverse relationship as it is the most predictive at institutions with an admittance rate of Over 75% (*r*=.56), compared to those with an admittance rate of 0% to 25% (*r*=.49). When the two SAT section scores and HSGPA are used in tandem, the strongest predictive relationship is seen at the most selective institutions with a correlation of .65.

### *Undergraduate Enrollment Size*

In terms of size, the SAT is the most predictive of FYGPA at small institutions compared to institutions of any other size. For small institutions the correlations for the two section scores as well as the multiple correlation of both SAT sections scores range from .51 to .56. In contrast, medium size institutions show the smallest correlations among the sizes with SAT section score correlations as well as the multiple correlation of SAT ranging from .46 to .50. HSGPA is the most predictive of FYGPA at small institutions (*r*=.57) and is slightly less predictive as size increases. Finally, the multiple correlation of the SAT section scores and HSGPA together to predict FYGPA is the largest at small institutions (*r*=.65) and almost equal at all other sized institutions.

### *Admittance Rate x Control*

For the interaction between admittance rate and control, results show that the SAT is the most predictive at Private selective institutions with correlations ranging from .57 to .59 for both section scores and .62 for the multiple correlation of the two section scores. HSGPA is the most predictive of FYGPA at Private institutions with admittance rates of 51% to 75% and Over 75%, both equal to .58. When using both the SAT sections and HSGPA together, the predictive relationship with FYGPA is the strongest again at Private institutions with an admittance rate of 0% to 25%.

Table 8: Corrected Correlations of SAT Scores and HSGPA with FYGPA by Institutional Subgroups

| Institution Subgroup | | *k* | *n* | SAT ERW | SAT Math | SAT | HSGPA | SAT, HSGPA |
|---|---|---|---|---|---|---|---|---|
| Control | Private | 89 | 46,981 | .54 | .52 | .56 | .56 | .65 |
| | Public | 82 | 176,877 | .47 | .46 | .50 | .53 | .59 |
| Admittance Rate | Under 25% | 20 | 16,075 | .58 | .55 | .61 | .49 | .65 |
| | 25% to 50% | 30 | 47,268 | .53 | .51 | .56 | .53 | .63 |
| | 51% to 75% | 73 | 107,840 | .47 | .46 | .49 | .53 | .59 |
| | Over 75% | 48 | 52,675 | .46 | .44 | .48 | .56 | .60 |
| Undergraduate Enrollment Size | Small | 67 | 22,899 | .53 | .51 | .56 | .57 | .65 |
| | Medium | 30 | 23,952 | .48 | .46 | .50 | .54 | .60 |
| | Large | 30 | 43,540 | .49 | .47 | .51 | .54 | .61 |
| | Very Large | 44 | 133,467 | .48 | .47 | .51 | .52 | .60 |
| Admittance Rate x Control | Private Under 25% | 18 | 9,556 | .59 | .57 | .62 | .52 | .67 |
| | Private 25% to 50% | 18 | 13,697 | .54 | .52 | .56 | .57 | .65 |
| | Private 51% to 75% | 36 | 15,740 | .53 | .51 | .55 | .58 | .65 |
| | Private Over 75% | 17 | 7,988 | .49 | .48 | .52 | .58 | .64 |
| | Public Under 25% | 2 | 6,519 | .58 | .53 | .59 | .46 | .62 |
| | Public 25% to 50% | 12 | 33,571 | .53 | .51 | .56 | .51 | .62 |
| | Public 51% to 75% | 37 | 92,100 | .46 | .45 | .48 | .52 | .58 |
| | Public Over 75% | 31 | 44,687 | .45 | .44 | .47 | .55 | .59 |
| Overall | | 171 | 223,858 | .49 | .47 | .51 | .53 | .61 |

Note: *n* = subgroup sample size, *k* = number of institutions. SAT indicates the multiple correlation of both SAT sections together. Undergraduate enrollment size 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.

## Retention Analyses

This section once again examines the relationship between retention to the second year at the same institution and the SAT for the 204,504 students studied in the student subgroup section. See Table B 2 in Appendix B for descriptive information on the study variables by institutional subgroup for this sample. Table 9 shows the mean retention rate to the second year by SAT score bands across institutional subgroups. In general, across all institutional subgroups, as SAT scores increase, the retention rate also increases. The most noticeable exception to this trend is seen in the Private, Under 25% subgroup in the Admittance Rate x Control category. This is most likely because the 800–990 SAT score band has fewer students in it (n=38) than the remaining score bands, so that percentage should be interpreted with caution. See Table B 3 in Appendix B for sample sizes in each of the SAT Total score bands by institutional subgroups.

## Table 9: Mean Second-Year Retention Rate by SAT Score Band for Institutional Subgroups

| Institutional Subgroup | | k | n | 400-590 | 600-790 | 800-990 | 1000-1190 | 1200-1390 | 1400-1600 |
|---|---|---|---|---|---|---|---|---|---|
| Control | Private | 81 | 43,061 | -- | .65 | .74 | .81 | .89 | .92 |
| | Public | 75 | 161,443 | -- | .64 | .71 | .80 | .88 | .92 |
| Admittance Rate | Under 25% | 16 | 11,611 | -- | -- | .92 | .92 | .94 | .93 |
| | 25% to 50% | 29 | 42,944 | -- | .71 | .76 | .84 | .90 | .92 |
| | 51% to 75% | 65 | 100,399 | -- | .67 | .73 | .80 | .88 | .92 |
| | Over 75% | 46 | 49,550 | -- | .60 | .69 | .77 | .84 | .89 |
| Undergraduate Enrollment Size | Small | 61 | 20,970 | -- | .57 | .69 | .79 | .87 | .91 |
| | Medium | 28 | 22,592 | -- | .61 | .70 | .78 | .86 | .93 |
| | Large | 27 | 37,460 | -- | .70 | .71 | .78 | .87 | .91 |
| | Very Large | 40 | 123,482 | -- | .67 | .73 | .81 | .89 | .93 |
| Admittance Rate x Control | Private, Under 25% | 15 | 8,074 | -- | -- | .97 | .92 | .93 | .93 |
| | Private, 25% to 50% | 18 | 13,697 | -- | .74 | .76 | .83 | .89 | .90 |
| | Private, 51% to 75% | 31 | 13,302 | -- | .71 | .74 | .81 | .87 | .89 |
| | Private, Over 75% | 17 | 7,988 | -- | .50 | .71 | .78 | .86 | .93 |
| | Public, Under 25% | 1 | 3,537 | -- | -- | -- | -- | -- | -- |
| | Public, 25% to 50% | 11 | 29,247 | -- | .69 | .76 | .85 | .90 | .93 |
| | Public, 51% to 75% | 34 | 87,097 | -- | .66 | .73 | .80 | .88 | .92 |
| | Public, Over 75% | 29 | 41,562 | -- | .62 | .68 | .77 | .83 | .89 |
| Overall | | 156 | 204,504 | -- | .65 | .72 | .80 | .88 | .92 |

Note: $n$ = subgroup sample size, $k$ = number of institutions. Results are presented for 15 students or more. Undergraduate enrollment size 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.

The next set of analyses examines the relationship between retention to the second year and academic overperformance and underperformance for institutional subgroups.

*Control*

Figure 5 shows the mean second-year retention rate for students who underperform and those who perform as well as expected or above by institutional control. The retention rates for students who perform as well as expected or above are very similar for both public (87%) and private (88%) institutions. However, those students who underperform have a higher mean retention rate (51%) at private institutions than at public institutions (38%).

Figure 5: Retention Rates of Students Underperforming and Performing as Expected or Above, Total Sample and by Control
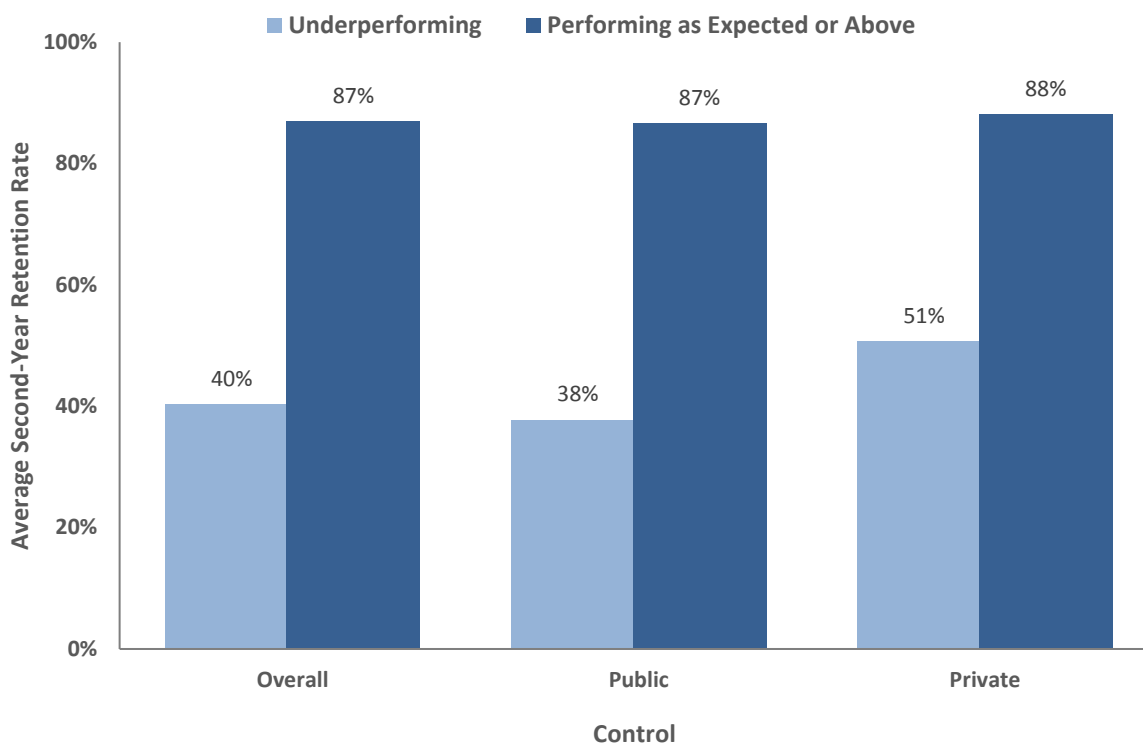


*Admittance Rate*

Figure 6 displays the retention rates of students who underperform and those who perform as well as expected or above by institutional admittance rate. As an institution admits a higher percentage of applicants, the retention rates in both student categories decrease. For the most selective institutions, those admitting under 25% of students, the retention rate for students who underperform (83%) is almost equal to that for students who perform as expected or above in the overall sample (87%). Furthermore, for the admittance rate of under 25%, a very high percentage of students (94%) who perform as well as expected or above return for the second year. For the admittance rate category of

25% to 50%, the retention rates for both student categories (Underperforming: 55%; Performing as Expected or Above: 91%) exceed what is seen in the overall sample (40% and 87%, respectively). The least selective admission rate category (over 75%) has the lowest rates (28% and 82%) of all groups and these rates are also lower than what is seen in the full sample. This is also the largest difference seen between the student performance groups in any of the admittance rate categories.

Figure 6: Retention Rates of Students Underperforming and Performing as Expected or Above, Total Sample and by Admittance Rate
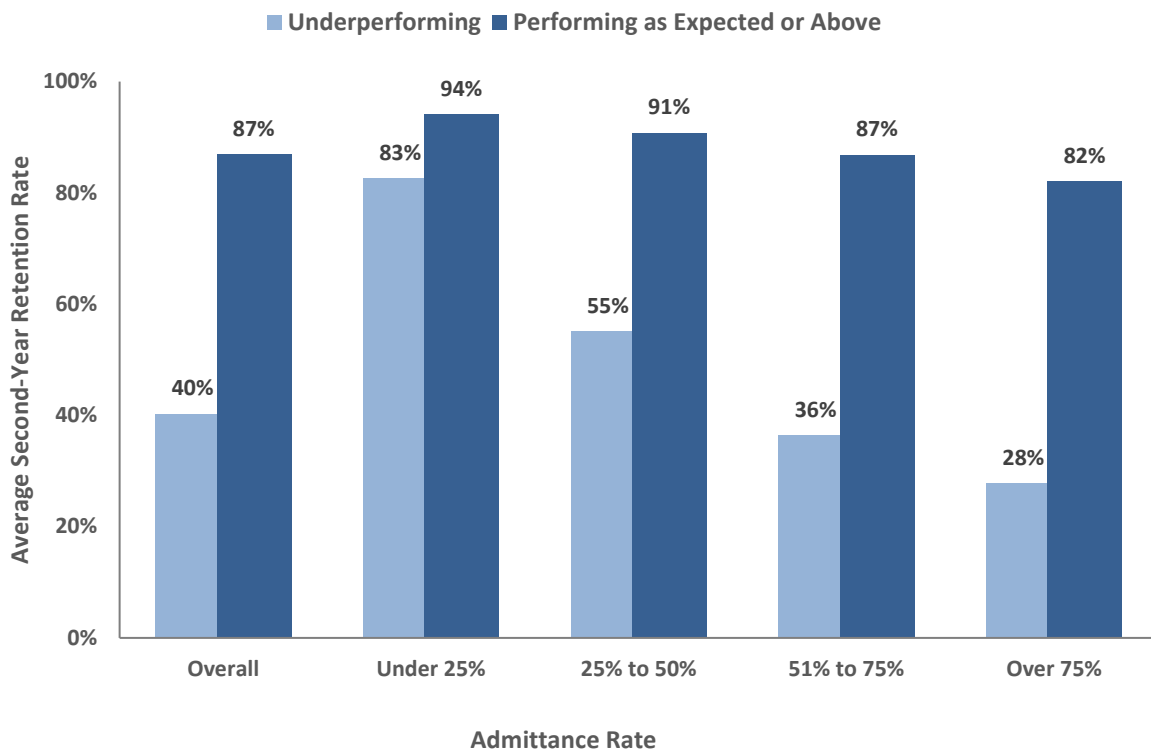


*Undergraduate Enrollment Size*
Figure 7 shows retention rates of students who underperform and those who perform as expected or above by institution size. There is no clear trend for these two groups as institution size changes. Medium size intuitions have the lowest rate of retention for students who underperform (30%) than other institutional size groups and the overall sample. Medium size institutions also have the lowest rate of retention to the second year for students who perform as well as expected or above (83%) than all other size institutions and the overall sample.

Figure 7: Retention Rates of Students Underperforming and Performing as Expected or Above, Total Sample and by Undergraduate Enrollment Size
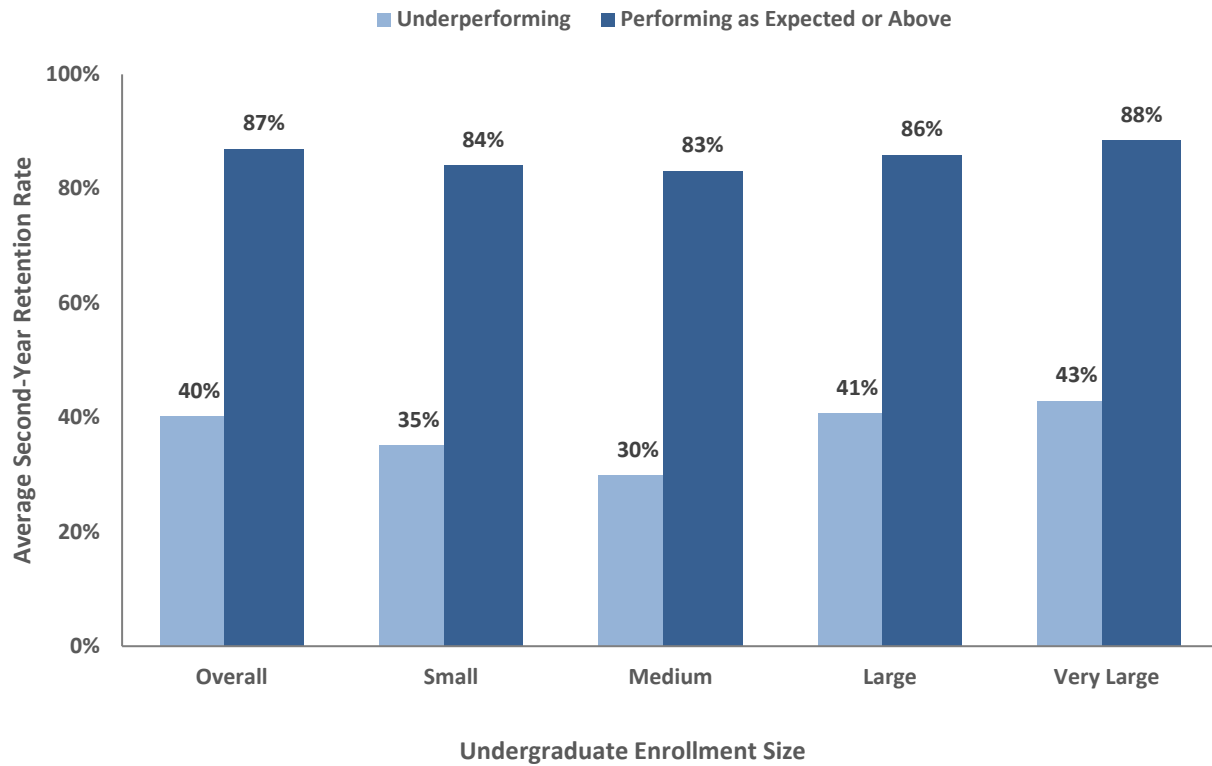


*Admittance Rate x Control*

Figure 8 shows the mean retention rate when the institutions' control and admittance rate are examined in tandem. The figure somewhat reflects the trends that were seen by control alone (Figure 5) and admittance rate alone (Figure 6)—retention rates decrease as the admittance rate increases and, in general, private institutions have higher retention rates of students who underperform compared to the comparable public admittance rate subgroup. These differences are very small, varying from 1%–4% by control, when holding selectivity constant. It is interesting that in three of the four admittance rate categories (all but over 75% admittance rate), public institutions have a slightly higher retention rate for students who perform as well as expected or above than private institutions.

# Discussion

This study explored the relationships between SAT scores, HSGPA, FYGPA, and retention by relevant student and institutional subgroups to understand whether SAT scores and HSGPA function differently for different subgroups of students. The findings from the current study are consistent with previous differential validity and prediction research on SAT and HSGPA relationships with FYGPA and retention (Mattern, et al., 2008; Mattern & Patterson, 2014; Ramist, Lewis, & McCamley-Jenkins, 1994; Young, 2001; Zwick, 2006, 2019). In general, while we see differences in predictive relationships by subgroups, they are relatively small. Moreover, SAT scores remain predictive of FYGPA and retention to the second year across student and institutional subgroups.

## Differential Validity, FYGPA

A consistent finding was that the relationship between the SAT and FYGPA is positive for all subgroups, with correlations ranging from .44 to .57. These are considered "medium" to "large" correlations (Cohen, 1988). Moreover, the multiple correlations for the SAT and HSGPA with FYGPA always exceed the correlations between HSGPA and FYGPA. The increases in the correlations indicate that the SAT adds

incremental validity beyond HSGPA for all student subgroups when predicting FYGPA. More information is better than less.

As noted above, the differences between groups are generally small, but the relationships are somewhat stronger for some groups than they are for others. In terms of gender, the SAT and HSGPA are more predictive for females than for males. In the analyses by race/ethnicity, the results indicate that the SAT is slightly more predictive for students who identify as American Indian or Alaska Native, Asian, and Two or More Races, and the correlations for the SAT are equal to those for HSGPA for black or African American and Hispanic or Latino students. Only for white students is HSGPA the strongest predictor of FYGPA. With regard to best language, the SAT and HSGPA are the most predictive for students who identify English as their best language versus the other language groups, though the SAT has a stronger relationship with FYGPA than does HSGPA for students who best language is either English and Another Language or Another Language. In terms of parental education level, both the SAT and HSGPA are more predictive as parental education level increases.

When examining institutional subgroups, similar to previous research, the SAT is somewhat more predictive at private institutions, more selective institutions, and small institutions (Beard & Marini, 2015, 2018; Mattern & Patterson, 2014). In all the analyses, the SAT added incremental validity beyond that of HSGPA. One notable result is that the validity of the SAT is the strongest at the most selective institutions, exceeding the validity of HSGPA. It is also at these institutions that the SAT adds the greatest incremental validity beyond HSGPA.

## Differential Prediction

Results from the differential prediction analyses are consistent with previous research. In terms of gender, male students' FYGPA tends to be slightly overpredicted by both the SAT and HSGPA and female students' FYGPA tends to be slightly underpredicted by both the SAT and HSGPA. The SAT and HSGPA both tend to overpredict the FYGPA of students who identify as American Indian or Alaska Native, black or African American, Hispanic or Latino, Native Hawaiian or Other Pacific Islander, and Two or More Races. Though the results vary across subgroupings, in most cases the amount of overprediction or underprediction is the smallest when the SAT and HSGPA are used together. The amount of overprediction or underprediction is the smallest when using HSGPA alone in the gender analyses, but there is usually less overprediction and underprediction when using the SAT alone than there is when using HSGPA alone in racial/ethnic, best language, and highest parental education level analyses.

Though there is some differential prediction for all subgroups, be it overprediction or underprediction, these differences are minor. When the SAT and HSGPA are used to predict FYGPA, the absolute values of the effect sizes—Cohen's $d$—are less than 0.20 for all but one small subgroup. The take home message is that institutions should use multiple measures to predict FYGPA, and that the amount of overprediction and underprediction will be minor.

## Differential Validity, Retention

The general trend across all student subgroups is that retention rates increase in tandem with SAT score bands.[7] Female students return for the second year at slightly higher rates than males across the SAT score bands. In the racial/ethnic group analyses, it is worth noting that the overall retention rates for black or African American students and Hispanic or Latino students are very similar to the retention rates of the total sample when looking within SAT score bands, with differences of only -.01 to +.02. From the analyses of the three language groups, the students in the Another Language group had retention rates lower than those in the total sample across SAT score bands in all but one score band (600–790). This suggests that these students may need additional support from their institutions. When looking at highest parental education level, second-year retention rates also increase by SAT score band for all the groups, though retention rates are generally higher for those with higher parental education levels.

Additional analyses focused on the retention rates for students who underperformed and for students who performed as expected or above. Students were categorized according to how much their actual FYGPAs differed from their predicted FYGPA based on HSGPA and SAT scores. Across all subgroups, students who earn FYGPAs much lower than predicted have lower retention rates than do students who perform as expected or above. Male and female students who perform as expected or above return at the same rate, but female students who underperform in the first year of college return for the second year at a slightly lower rate than underperforming male students. Students from underrepresented minority groups tend to have lower rates of returning when they underperform versus students in the other racial/ethnic groups. Interestingly, students in the Another Language group who underperform in their first year tend to return at slightly higher rates than the underperformers in the full sample and in the other two language groups. However, the students performing as expected or above in that Another Language group tended to return at lower rates than do students in the total sample and the other two language groups. Finally, retention rates increase as highest parental education level increases for both underperforming students and for students performing as expected or above. In particular, students in the No High School Diploma, High School Diploma, and Associate Degree groups have retention rates below those found in the total sample. Conversely, students in the Bachelor's Degree and Graduate Degree groups have retention rates higher than those for the total sample.

When examining institutional subgroups for the retention analyses, we see that as SAT score bands increase, retention rates also increase across virtually all institutional subgroups. Higher retention rates, however, are generally found at the private and selective institutions, with no clear pattern by size. Students at private institutions who underperform have a higher rate of returning to the second year than do underperforming students at public institutions. Retention rates are highest for both

---

[7] Atypical results in the current study are associated with small sample sizes within score bands for smaller subgroups. Examples of these cases are the students who identified as being two or more races (SAT score band 600–790, n = 26), students whose best language was a language other than English (SAT score band 600–790, n = 52), and students whose parents had no high school diploma (SAT scores band 1400–1600, n = 226). Whether these results are truly atypical, or whether they represent a consistent pattern, is an area for future research in the coming years.

underperforming students and those performing as expected or above at the most selective institutions, and retention rates decrease as institutions become less selective. Finally, medium-sized institutions have the lowest rates of retention for both student performance categories. Perhaps this is because medium-sized institutions are just big enough for students to get lost in the crowd and not large enough to have substantial programs and resources to promote student success.

## Future Research

When interpreting findings from differential validity and prediction research, it is helpful to keep in mind previous research that has identified a variety of factors that may be influencing the results (Zwick, 2002, 2019). Much of this research has focused on gender and race/ethnicity. To contextualize the findings regarding the underprediction of female students' FYGPA, researchers have explored differential course taking (Ramist et al., 1994), choice of academic major (Pennock-Román, 1994), and gender differences in studiousness (Mattern, et al., 2008; Young, 2004; Zwick, 2002). Past research has also found that the amount of overprediction and underprediction of male and female students varies by race/ethnicity (Bridgeman, McCamley-Jenkins, & Ervin, 2000). Furthermore, research also suggests that differences between male and female students, and for students in different racial/ethnic groups, are smaller at more selective institutions (Bridgeman et al., 2000; Ramist et al, 1994).

Any of these reasons can contribute to our understanding of the findings we see in this report. The importance of fairness and equity for all students compels us to continue our research efforts, so we intend to replicate and extend the research discussed above to further understand the relationships between SAT scores and academic outcomes across student subgroups.

## Conclusion

Taken together, these study findings show that, in general, the utility of the SAT and its added informational value above HSGPA to predict FYGPA holds across the student and institutional subgroups examined. Institutions can feel confident using SAT scores and HSGPA for admission, scholarship, and advising/retention decisions across various student and institutional subgroups. The small differences found in the prediction of FYGPA by subgroup echo what has been found in previous research and in most cases are not even categorized as small effect sizes, but they are still important to study, document, and further examine.

Although it's critical to conduct differential validity and prediction research nationally, institutions are encouraged to conduct their own campus-specific research. There are likely unique institutional circumstances that can impact and contribute to the relationship between SAT scores, HSGPA, and FYGPA across subgroups that can best be understood in local contexts. The College Board's Admitted Class Evaluation Service (ACES) system remains available to institutions at no cost to conduct this research.

# References

Aud, S., Fox, M., & KewalRamani, A. (2010). *Status and Trends in the Education of Racial and Ethnic Groups* (NCES 2010-015). U.S. Department of Education, National Center for Education Statistics. Washington DC: U.S. Government Printing Office.

Beard, J., & Marini, J. P. (2015). *Validity of the SAT for predicting first-year grades: 2012 SAT validity sample* (College Board Statistical Report 2015-2). New York: College Board.

Beard, J., & Marini, J. (2018). *Validity of the SAT for predicting first-year grades: 2013 SAT validity sample*. New York: College Board

Breland, H. M. (1979). *Population validity and college entrance measures.* (Research Monograph No. 8). New York: College Board.

Bridgeman, B., McCamley-Jenkins, L., & Ervin, N. (2000). *Prediction of freshman grade-point average from the revised and recentered SAT I: Reasoning Test* (College Board Report 2000-1). New York: College Board.

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

College Board (2017). *SAT Suite of Assessments Annual Report.* New York: College Board.

Fischer, F. T., Schult, J., & Hell, B. (2013). Sex-specific differential prediction of college admission tests: A meta-analysis. *Journal of Educational Psychology, 105* (2), 478–488.

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.

Linn, R. L. (1973). Fair test use in selection. *Review of Educational Research, 43*, 139–161.

Linn, R. L. (1978). Single-group validity, differential validity, and differential prediction. *Journal of Applied Psychology, 63*(4), 507–512.

Linn, R. L. (1982). Ability testing: Individual differences, prediction, and differential prediction. In R. L. Linn (Ed.), *Ability testing: Uses, consequences, and controversies*. Washington, DC: National Academy Press.

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: 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 2008-4). New York: College Board.

Maxey, J., & Sawyer, R. (1981, July). *Predictive validity of the ACT Assessment for Afro-American/Black, Mexican-American/Chicano, and Caucasian-American/White students* (ACT Research Bulletin 81-1). Iowa City, IA: American College Testing.

Pascarella, E. T. & Terenzini, P. T. (2005). *How college affects students: A third decade of research* (Vol. 2). San Francisco: Jossey-Bass.

Pennock-Román, M. (1994). *College major and gender differences in the prediction of college grades* (College Board Report 94-2). New York: College Board.

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

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., & 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.

Westrick, P. A., Le, H., Robbins, S. B., Radunzel, J. M. R., & Schmidt, F. L. (2015) College Performance and Retention: A Meta-Analysis of the Predictive Validities of ACT® Scores, High School Grades, and SES, *Educational Assessment, 20*(1), 23–45.

Westrick, P. A., Marini, J. P, Young, L., Ng, H., Shmueli, D., & Shaw, E. J. (2019). *Validity of the SAT for predicting first-year grades and retention to the second year.* (College Board Research Report). New York: College Board.

Young, J. W. (2001). *Differential validity, differential prediction, and college admissions testing: A comprehensive review and analysis.* (Research Report No. 2001-6). New York: College Board.

Young, J. W. (2004). Differential validity and prediction: Race and sex differences in college admissions testing. In R. Zwick (Ed.), *Rethinking the SAT: The future of standardized testing in university admissions*, pp.289–301. New York: RoutledgeFalmer.

Zwick, R. (2002). *Fair game: The use of standardized admissions tests in higher education*. New York: RoutledgeFalmer.

Zwick, R. (2006). Higher education admissions testing. In R. Brennan (Ed.), *Educational Measurement* (4th ed., pp. 647-679). Westport, CT: American Council on Education, Praeger.

Zwick, R. (2019). Assessment in American higher education: The role of admission tests. *The ANNALS of the American Academy of Political and Social Science, 683*(1), 130–148. https://doi.org/10.1177/0002716219843469

Zwick, R., & Schlemer, L. (2004). SAT validity for linguistic minorities at the University of California, Santa Barbara. *Educational Measurement: Issues and Practice*, 25, 6–16.

Zwick, R., & Sklar, J. C. (2005). Predicting college grades and degree completion using high school grades and SAT scores: The role of student ethnicity and first language. *American Educational Research Journal,* 42, 439–464.

# Appendix A: Student Subgroups

Table A 1: Raw Correlations of SAT Scores and HSGPA with FYGPA by Student Subgroups

|  | Student Subgroup | k | n | SAT ERW | SAT Math | SAT | HSGPA | SAT, HSGPA |
|---|---|---|---|---|---|---|---|---|
| Gender | Male | 169 | 97,080 | .27 | .29 | .32 | .32 | .41 |
|  | Female | 171 | 126,778 | .33 | .33 | .37 | .32 | .44 |
| Race/Ethnicity | American Indian or Alaska Native | 8 | 220 | .23 | .21 | .39 | .38 | .53 |
|  | Asian | 125 | 24858 | .27 | .27 | .33 | .23 | .39 |
|  | Black or African American | 142 | 15,741 | .21 | .18 | .25 | .27 | .36 |
|  | Hispanic or Latino | 156 | 47,036 | .25 | .22 | .27 | .27 | .36 |
|  | Native Hawaiian or Other Pacific Islander | - | - | - | - | - | - | - |
|  | Two or More Races | 116 | 8,077 | .30 | .27 | .34 | .30 | .43 |
|  | White | 169 | 122,749 | .27 | .25 | .30 | .37 | .42 |
|  | Not Stated | 78 | 2,635 | .24 | .19 | .32 | .27 | .42 |
| Best Language | English Only | 171 | 184,083 | .30 | .28 | .32 | .35 | .43 |
|  | English and Another | 154 | 35,057 | .26 | .26 | .31 | .25 | .38 |
|  | Another Language | 47 | 2,854 | .23 | .21 | .29 | .18 | .36 |
|  | Not Stated | 17 | 365 | .30 | .29 | .43 | .37 | .53 |
| Highest Parental Education Level | No High School Diploma | 86 | 12,284 | .23 | .21 | .27 | .23 | .35 |
|  | High School Diploma | 166 | 48,081 | .25 | .23 | .28 | .31 | .39 |
|  | Associate Degree | 142 | 15,399 | .25 | .23 | .29 | .34 | .41 |
|  | Bachelor's Degree | 171 | 80,465 | .27 | .25 | .30 | .35 | .42 |
|  | Graduate Degree | 170 | 63,526 | .29 | .27 | .32 | .34 | .42 |
|  | Not Stated | 72 | 2,630 | .21 | .22 | .31 | .30 | .43 |
| Overall |  | 171 | 223,858 | .29 | .27 | .32 | .33 | .42 |

Note: Correlations are only calculated at institutions with at least 15 students in a subgroup. SAT indicates the multiple correlation of both SAT sections together.

Figure A 1: Average Overprediction (-) and Underprediction (+) of FYGPA by SAT Scores and HSGPA —Gender



**Male**

**Female**

UNDERPREDICTION

OVERPREDICTION

0.12

0.09

0.06

-0.08

-0.11

-0.16

■ HSGPA  ■ SAT  ■ SAT,HSGPA
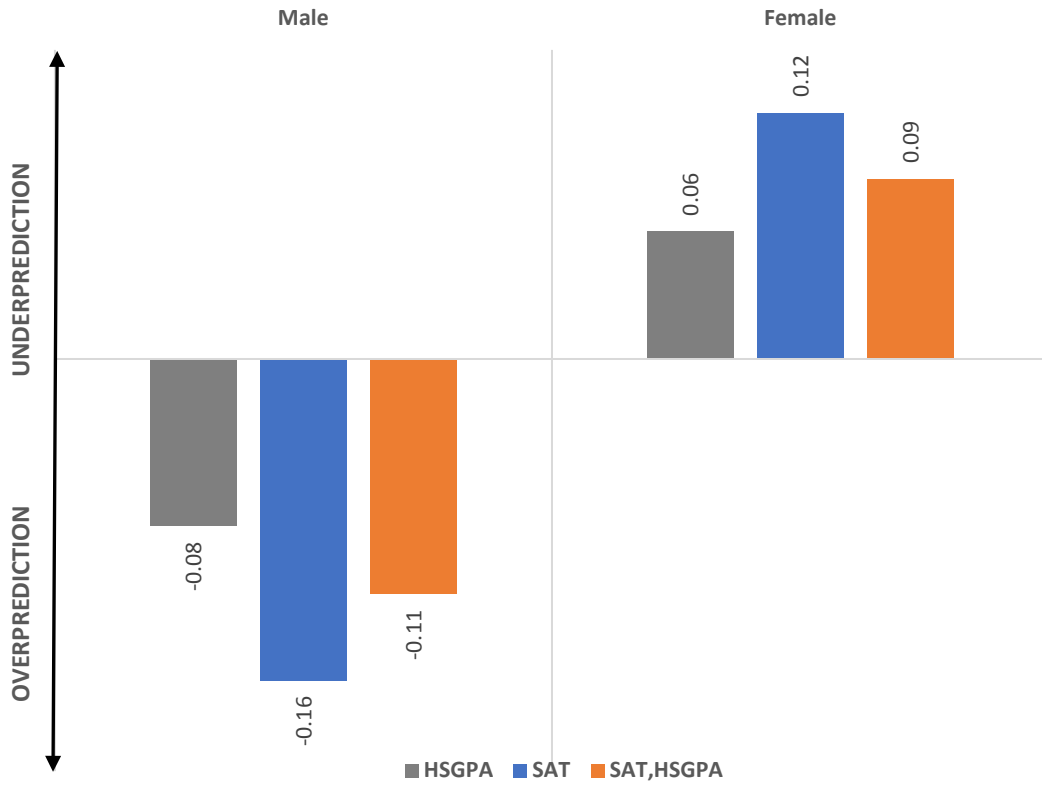
Figure A 2: Average Overprediction (-) and Underprediction (+) of FYGPA by SAT Scores and HSGPA —Race/Ethnicity



UNDERPREDICTION

OVERPREDICTION

| | American Indian or Alaska Native | Asian | Black or African American | Hispanic or Latino | Native Hawaiian or Other Pacific Islander | White | Two or more races |
|---|---|---|---|---|---|---|---|
| HSGPA | -0.17 | 0.08 | -0.18 | -0.10 | -0.21 | 0.05 | -0.04 |
| SAT | -0.12 | 0.02 | -0.13 | -0.04 | -0.20 | 0.04 | -0.06 |
| SAT,HSGPA | -0.12 | 0.03 | -0.10 | -0.05 | -0.17 | 0.03 | -0.06 |

■ HSGPA ■ SAT ■ SAT,HSGPA

Figure A 3: Average Overprediction (-) and Underprediction (+) of FYGPA by SAT Scores and HSGPA —Best Language



**HSGPA** **SAT** **SAT,HSGPA**
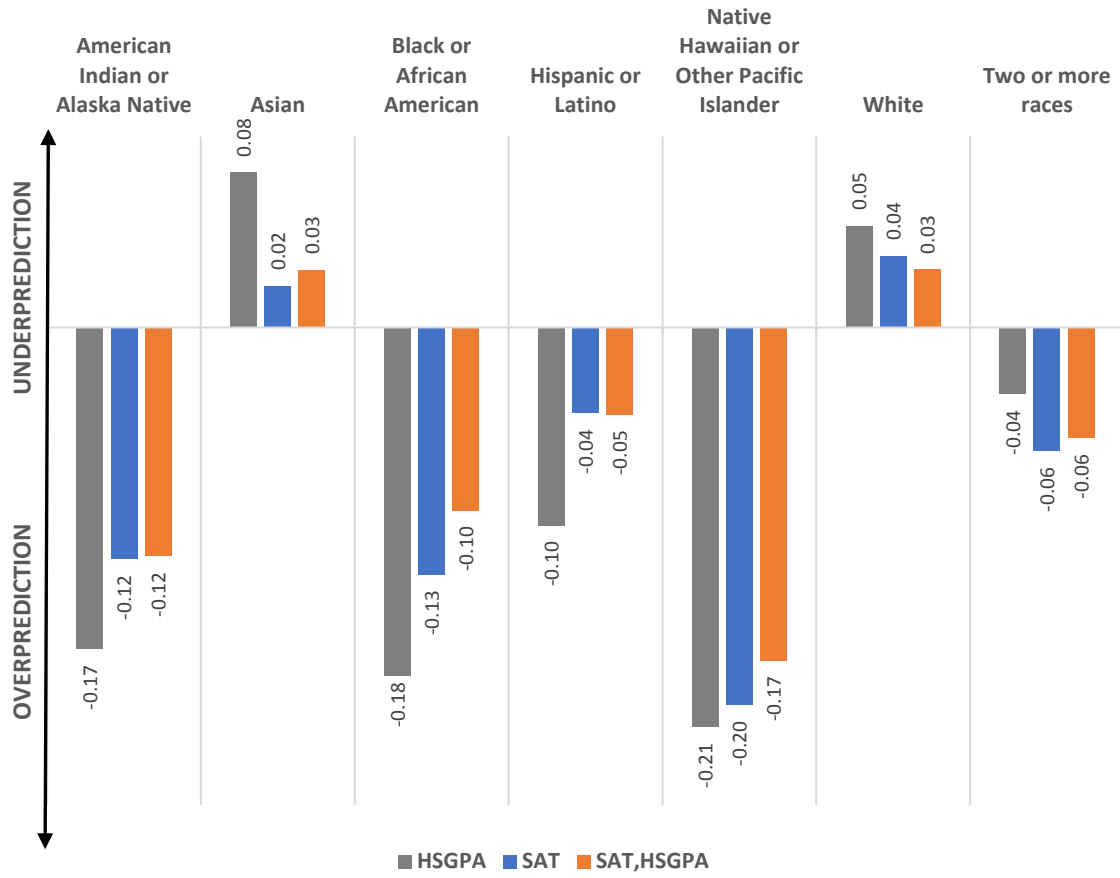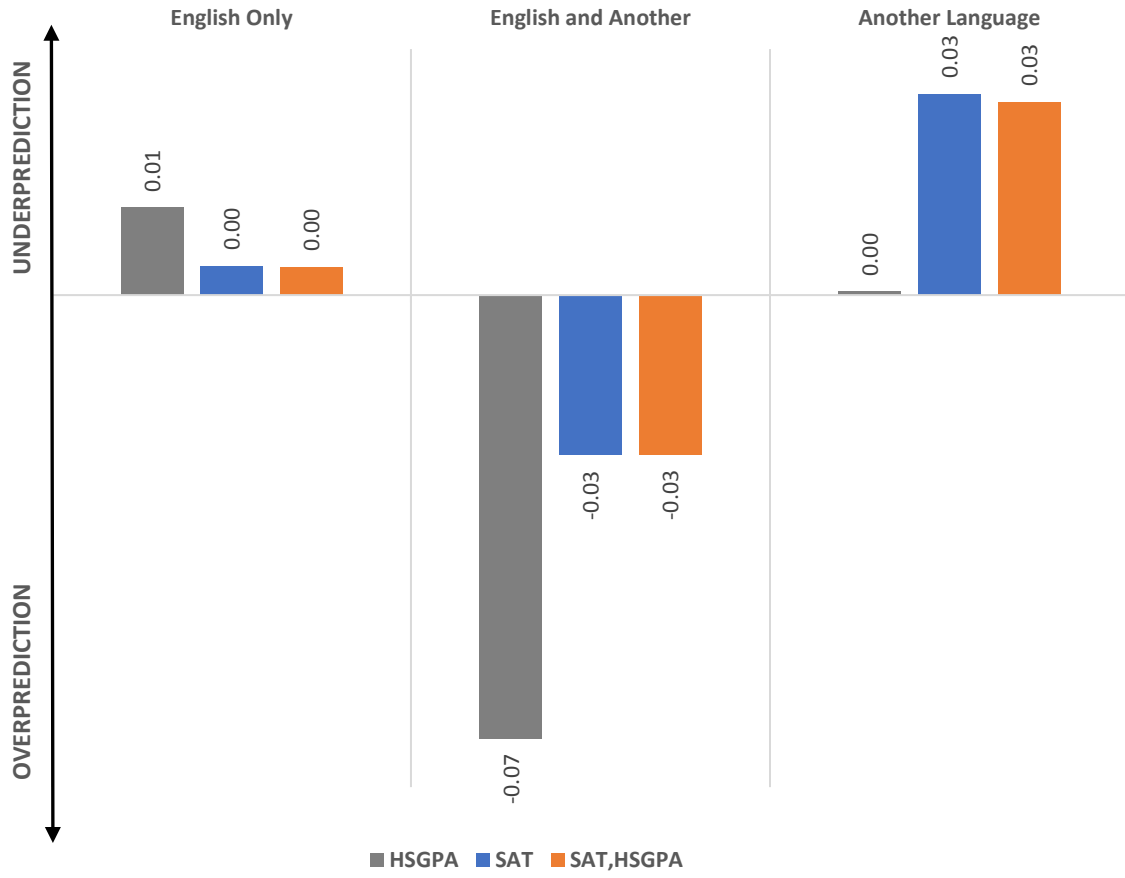
Figure A 4: Average Overprediction (-) and Underprediction (+) of FYGPA by SAT Scores and HSGPA —Highest Parental Education Level

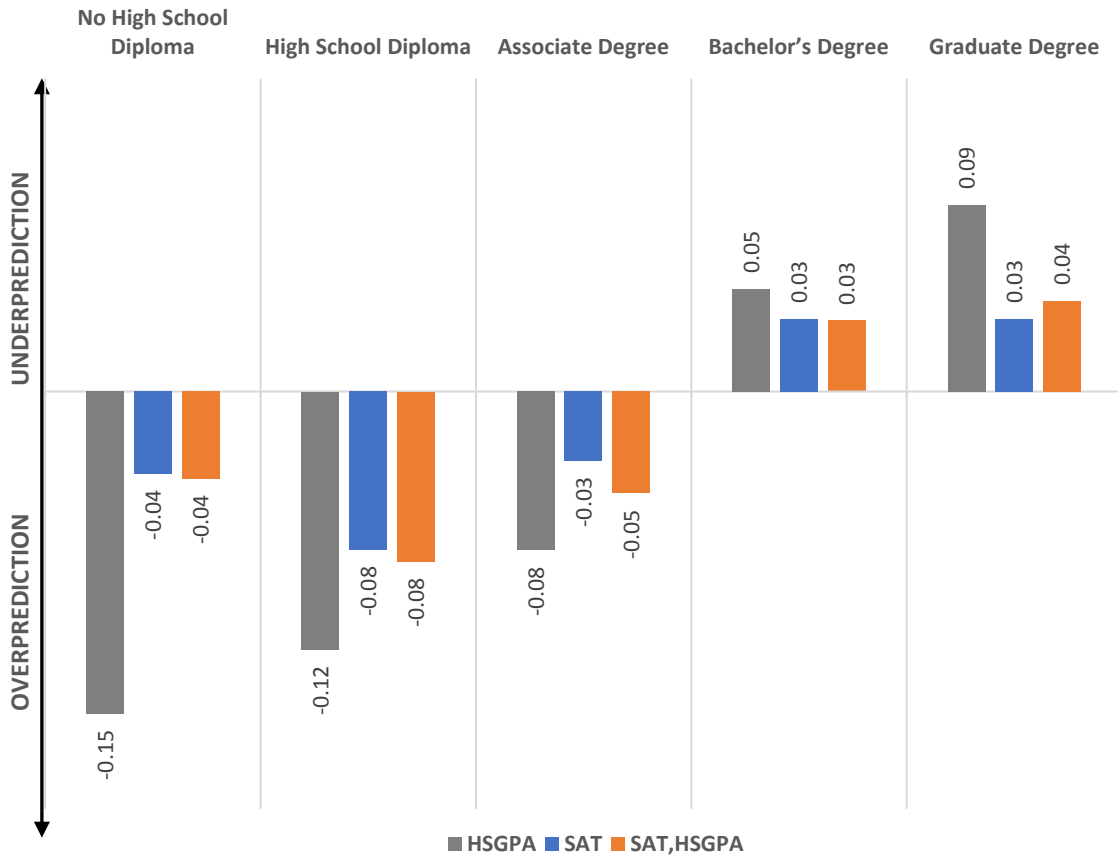Table A 2: Descriptive Statistics for Study Variables by Student Subgroups for the Retention Sample

| | | | | SAT ERW | | SAT Math | | HSGPA | | FYGPA | | Retention | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Student Subgroup | k | n | M | SD | M | SD | M | SD | M | SD | M | SD |
| Gender | Male | 154 | 88,428 | 602 | 83 | 612 | 92 | 3.62 | 0.50 | 2.91 | 0.86 | .82 | .38 |
| | Female | 156 | 116,076 | 591 | 82 | 573 | 89 | 3.71 | 0.45 | 3.12 | 0.76 | .84 | .36 |
| Race/Ethnicity | American Indian or Alaska Native | 119 | 629 | 558 | 77 | 548 | 85 | 3.57 | 0.50 | 2.71 | 0.92 | .76 | .43 |
| | Asian | 152 | 21,193 | 618 | 83 | 656 | 94 | 3.76 | 0.41 | 3.18 | 0.68 | .87 | .34 |
| | Black or African American | 156 | 14,851 | 547 | 81 | 528 | 84 | 3.47 | 0.53 | 2.68 | 0.88 | .79 | .41 |
| | Hispanic or Latino | 156 | 42,815 | 567 | 81 | 557 | 86 | 3.63 | 0.48 | 2.84 | 0.84 | .81 | .40 |
| | Native Hawaiian or Other Pacific Islander | 83 | 286 | 580 | 80 | 573 | 91 | 3.58 | 0.51 | 2.75 | 0.88 | .77 | .42 |
| | White | 155 | 114,051 | 608 | 77 | 599 | 85 | 3.69 | 0.46 | 3.12 | 0.78 | .84 | .36 |
| | Two or More Races | 154 | 7,733 | 615 | 79 | 605 | 90 | 3.70 | 0.46 | 3.02 | 0.83 | .85 | .36 |
| | Not Stated | 156 | 2,946 | 568 | 87 | 561 | 97 | 3.52 | 0.52 | 2.84 | 0.88 | .78 | .42 |
| Best Language | English Only | 156 | 169,319 | 601 | 81 | 592 | 90 | 3.67 | 0.47 | 3.05 | 0.81 | .84 | .37 |
| | English and Another | 156 | 31,142 | 573 | 85 | 577 | 101 | 3.65 | 0.47 | 2.91 | 0.81 | .82 | .38 |
| | Another Language | 138 | 3,085 | 559 | 91 | 649 | 116 | 3.69 | 0.46 | 3.04 | 0.78 | .77 | .42 |
| | Not Stated | 148 | 958 | 567 | 92 | 565 | 105 | 3.51 | 0.53 | 2.87 | 0.88 | .78 | .41 |
| Highest Parental Education Level | No High School Diploma | 154 | 11,291 | 536 | 77 | 538 | 88 | 3.58 | 0.49 | 2.74 | 0.84 | .79 | .41 |
| | High School Diploma | 156 | 43,479 | 565 | 78 | 558 | 87 | 3.59 | 0.50 | 2.81 | 0.88 | .78 | .41 |
| | Associate Degree | 156 | 14,361 | 571 | 77 | 560 | 83 | 3.62 | 0.49 | 2.89 | 0.86 | .79 | .41 |
| | Bachelor's Degree | 156 | 74,218 | 603 | 76 | 598 | 87 | 3.69 | 0.46 | 3.10 | 0.76 | .85 | .36 |
| | Graduate Degree | 156 | 58,220 | 630 | 78 | 624 | 90 | 3.74 | 0.45 | 3.20 | 0.72 | .87 | .33 |
| | Not Stated | 153 | 2,935 | 539 | 85 | 539 | 95 | 3.46 | 0.53 | 2.70 | 0.92 | .76 | .43 |
| Overall | | 156 | 204,504 | 596 | 82 | 590 | 92 | 3.67 | 0.47 | 3.03 | 0.81 | .83 | .37 |

Table A 3: Sample Size by Student Subgroup and SAT Score Band

| Student Subgroup | | SAT Score Band | | | | | |
|---|---|---|---|---|---|---|---|
| | | 400-590 | 600-790 | 800-990 | 1000-1190 | 1200-1390 | 1400-1600 |
| Gender | Male | 4 | 502 | 7,873 | 31,025 | 36,492 | 12,532 |
| | Female | 6 | 856 | 16,472 | 49,286 | 40,604 | 8,852 |
| Racial / Ethnic Identity | American Indian or Alaska Native | 1 | 9 | 119 | 329 | 155 | 16 |
| | Asian | 0 | 63 | 1,135 | 5,070 | 9,646 | 5,279 |
| | Black or African American | 2 | 416 | 4,218 | 6,990 | 2,912 | 313 |
| | Hispanic or Latino | 2 | 514 | 8,529 | 19,963 | 11,915 | 1,892 |
| | Native Hawaiian or Other Pacific Islander | 0 | 3 | 40 | 134 | 89 | 20 |
| | Two or More Races | 0 | 26 | 639 | 2,698 | 3,300 | 1,070 |
| | White | 4 | 284 | 9,043 | 43,881 | 48,245 | 12,594 |
| | Not Stated | 1 | 43 | 622 | 1,246 | 834 | 200 |
| Best Language | English Only | 7 | 896 | 17,963 | 66,295 | 66,208 | 17,950 |
| | English and Another | 2 | 381 | 5,750 | 12,775 | 9,427 | 2,807 |
| | Another Language | 1 | 52 | 432 | 864 | 1,195 | 541 |
| | Not Stated | 0 | 29 | 200 | 377 | 266 | 86 |
| Highest Parental Education Level | No High School Diploma | 1 | 279 | 3,266 | 5,336 | 2,183 | 226 |
| | High School Diploma | 2 | 473 | 8,392 | 20,798 | 12,163 | 1,651 |
| | Associate Degree | 2 | 114 | 2,464 | 6,983 | 4,238 | 560 |
| | Bachelor's Degree | 3 | 272 | 6,278 | 29,018 | 31,120 | 7,527 |
| | Graduate Degree | 1 | 121 | 3,073 | 16,942 | 26,775 | 11,308 |
| | Not Stated | 1 | 99 | 872 | 1,234 | 617 | 112 |
| Overall | | 10 | 1,358 | 24,345 | 80,311 | 77,096 | 21,384 |

# Appendix B: Institutional Subgroups

Table B 1: Raw Correlations of SAT Scores and HSGPA with FYGPA by Institution Subgroups

| Institution Subgroup | | k | n | SAT ERW | SAT Math | SAT | HSGPA | SAT, HSGPA |
|---|---|---|---|---|---|---|---|---|
| Control | Private | 89 | 46,981 | .33 | .30 | .36 | .37 | .46 |
| | Public | 82 | 176,877 | .28 | .27 | .31 | .32 | .41 |
| Admittance Rate | Under 25% | 20 | 16,075 | .38 | .34 | .40 | .22 | .43 |
| | 25% to 50% | 30 | 47,268 | .33 | .30 | .36 | .28 | .42 |
| | 51% to 75% | 73 | 107,840 | .26 | .25 | .29 | .33 | .40 |
| | Over 75% | 48 | 52,675 | .30 | .28 | .32 | .42 | .46 |
| Undergraduate Enrollment Size | Small | 67 | 22,899 | .35 | .33 | .38 | .41 | .50 |
| | Medium | 30 | 23,952 | .29 | .26 | .31 | .37 | .44 |
| | Large | 30 | 43,540 | .29 | .25 | .31 | .34 | .42 |
| | Very Large | 44 | 133,467 | .29 | .27 | .32 | .31 | .41 |
| Admittance Rate x Control | Private Under 25% | 18 | 9,556 | .34 | .31 | .37 | .25 | .41 |
| | Private 25% to 50% | 18 | 13,697 | .32 | .27 | .35 | .35 | .43 |
| | Private 51% to 75% | 36 | 15,740 | .35 | .32 | .38 | .41 | .49 |
| | Private Over 75% | 17 | 7,988 | .32 | .30 | .35 | .44 | .49 |
| | Public Under 25% | 2 | 6,519 | .43 | .39 | .45 | .18 | .46 |
| | Public 25% to 50% | 12 | 33,571 | .33 | .31 | .36 | .25 | .41 |
| | Public 51% to 75% | 37 | 92,100 | .25 | .24 | .28 | .31 | .39 |
| | Public Over 75% | 31 | 44,687 | .29 | .27 | .32 | .42 | .46 |
| Overall | | 171 | 223,858 | .29 | .27 | .32 | .33 | .42 |

Note: Correlations are only calculated at institutions with at least 15 students in a subgroup. SAT indicates the multiple correlation of both SAT sections together.

Table B 2: Descriptive Statistics for Study Variables by Institution Subgroups for the Retention Sample

| Institutional Subgroup | | k | n | SAT ERW M | SAT ERW SD | SAT Math M | SAT Math SD | HSGPA M | HSGPA SD | FYGPA M | FYGPA SD | Retention M | Retention SD |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Control | Private | 81 | 43,061 | 616 | 85 | 610 | 96 | 3.71 | 0.46 | 3.19 | 0.69 | 0.85 | 0.35 |
| | Public | 75 | 161,443 | 590 | 81 | 585 | 91 | 3.66 | 0.48 | 2.99 | 0.83 | 0.83 | 0.38 |
| Admittance Rate | Under 25% | 16 | 11,611 | 682 | 67 | 689 | 81 | 3.99 | 0.30 | 3.37 | 0.53 | 0.93 | 0.25 |
| | 25% to 50% | 29 | 42,944 | 625 | 77 | 630 | 89 | 3.83 | 0.38 | 3.14 | 0.67 | 0.88 | 0.32 |
| | 51% to 75% | 65 | 100,399 | 587 | 78 | 578 | 85 | 3.64 | 0.47 | 3.00 | 0.84 | 0.83 | 0.38 |
| | Over 75% | 46 | 49,550 | 568 | 78 | 557 | 84 | 3.51 | 0.51 | 2.92 | 0.88 | 0.78 | 0.42 |
| Undergraduate Enrollment Size | Small | 61 | 20,970 | 581 | 88 | 567 | 92 | 3.58 | 0.51 | 3.02 | 0.81 | 0.80 | 0.40 |
| | Medium | 28 | 22,592 | 562 | 81 | 552 | 88 | 3.49 | 0.52 | 2.90 | 0.87 | 0.79 | 0.41 |
| | Large | 27 | 37,460 | 601 | 86 | 594 | 96 | 3.64 | 0.48 | 3.05 | 0.79 | 0.82 | 0.38 |
| | Very Large | 40 | 123,482 | 603 | 79 | 600 | 90 | 3.72 | 0.45 | 3.05 | 0.80 | 0.85 | 0.36 |
| Admittance Rate x Control | Private, Under 25% | 15 | 8,074 | 694 | 61 | 703 | 72 | 3.99 | 0.31 | 3.42 | 0.50 | 0.93 | 0.26 |
| | Private, 25% to 50% | 18 | 13,697 | 629 | 74 | 626 | 86 | 3.75 | 0.41 | 3.22 | 0.65 | 0.87 | 0.34 |
| | Private, 51% to 75% | 31 | 13,302 | 582 | 76 | 569 | 82 | 3.60 | 0.47 | 3.08 | 0.74 | 0.83 | 0.38 |
| | Private, Over 75% | 17 | 7,988 | 569 | 74 | 556 | 76 | 3.54 | 0.51 | 3.07 | 0.78 | 0.79 | 0.41 |
| | Public, Under 25% | 1 | 3,537 | 653 | 73 | 655 | 91 | 4.00 | 0.29 | 3.26 | 0.58 | 0.94 | 0.23 |
| | Public, 25% to 50% | 11 | 29,247 | 623 | 78 | 632 | 91 | 3.87 | 0.36 | 3.11 | 0.67 | 0.89 | 0.32 |
| | Public, 51% to 75% | 34 | 87,097 | 588 | 78 | 580 | 86 | 3.64 | 0.47 | 2.98 | 0.85 | 0.83 | 0.38 |
| | Public, Over 75% | 29 | 41,562 | 568 | 79 | 557 | 85 | 3.50 | 0.51 | 2.89 | 0.90 | 0.77 | 0.42 |
| Overall | | 156 | 204,504 | 596 | 82 | 590 | 92 | 3.67 | 0.47 | 3.03 | 0.81 | 0.83 | 0.37 |

Table B 3: Sample Size by Institution Subgroup and SAT Score Band

| Institutional Characteristic | | SAT Score Band | | | | | |
|---|---|---|---|---|---|---|---|
| | | 400–590 | 600–790 | 800–990 | 1000–1190 | 1200–1390 | 1400–1600 |
| Control | Private | 1 | 195 | 3,833 | 14,511 | 16,872 | 7,649 |
| | Public | 9 | 1,163 | 20,512 | 65,800 | 60,224 | 13,735 |
| Admittance Rate | Under 25% | 0 | 4 | 126 | 1,296 | 4,332 | 5,853 |
| | 25% to 50% | 1 | 129 | 2,412 | 11,281 | 21,435 | 7,686 |
| | 51% to 75% | 6 | 670 | 12,772 | 43,448 | 37,412 | 6,091 |
| | Over 75% | 3 | 555 | 9,035 | 24,286 | 13,917 | 1,754 |
| Undergraduate Enrollment Size | Small | 0 | 263 | 3,667 | 9,051 | 6,304 | 1,685 |
| | Medium | 1 | 303 | 4,980 | 10,517 | 5,903 | 888 |
| | Large | 2 | 269 | 4,481 | 14,042 | 13,633 | 5,033 |
| | Very Large | 7 | 523 | 11,217 | 46,701 | 51,256 | 13,778 |
| Admittance Rate x Control | Private, Under 25% | 0 | 2 | 38 | 493 | 2,892 | 4,649 |
| | Private, 25% to 50% | 1 | 38 | 660 | 3,641 | 7,049 | 2,308 |
| | Private, 51% to 75% | 0 | 87 | 1,835 | 6,198 | 4,654 | 528 |
| | Private, Over 75% | 0 | 68 | 1,300 | 4,179 | 2,277 | 164 |
| | Public, Under 25% | 0 | 2 | 88 | 803 | 1,440 | 1,204 |
| | Public, 25% to 50% | 0 | 91 | 1,752 | 7,640 | 14,386 | 5,378 |
| | Public, 51% to 75% | 6 | 583 | 10,937 | 37,250 | 32,758 | 5,563 |
| | Public, Over 75% | 3 | 487 | 7,735 | 20,107 | 11,640 | 1,590 |
| Overall | | 10 | 1,358 | 24,345 | 80,311 | 77,096 | 21,384 |