Executive Summary

The purpose of this document is to identify and dispel rumors that are frequently cited about the SAT. The following is a compilation of nine popular rumors organized into three areas: Student Demographics, Test Preparation/Test Prediction, and Test Utilization.

Student Demographics

Student demographics claims are those that primarily center around a specific demographic group that is said to over-/underperform on the SAT. Presumably, it is reasoned, if the SAT were a fair test, no student demographic characteristics would matter, as average scores would be similar across groups. Therefore, some people assume that any difference in SAT performance by demographics must indicate test bias for/against a demographic group.

Rumor 1: The SAT Is a Wealth Test.

According to Peter Sacks, author of Standardized Minds, “one can make a good guess about a child’s standardized test scores simply by looking at how many degrees her parents have and what kind of car they drive.”

Fact: Rigorous Course Taking in High School Better Explains Why Students Do Well (or Not) on the SAT.

While SAT scores may be correlated with socioeconomic status (parental income and education), correlation does not mean two phenomena are causally related (e.g., parental income causes students to do well on the SAT). Generally, students do well on the SAT because they are exposed to and apply knowledge gained in rigorous course material in high school and take rigorous core courses. Specifically, research has shown that taking rigorous courses in math is positively associated with SAT scores.

Rumor 2: The SAT Is Biased Against Minorities.

According to Joseph Soares, author of SAT Wars, “The SAT and ACT are fundamentally discriminatory” because “… they’ve inadvertently ended up with a statistical algorithm that is systematically biased against racial minorities and women.”

Fact: Disparities in SAT Scores Are the Product of Different Social Conditions, Policy Choices, and Educational Opportunities That Advantage Some Groups Over Others.

If the SAT were biased against some minority groups (e.g., African Americans and Hispanics) we would expect: (1) minority students to miss a certain question or series of questions more so than white students, because presumably minorities would not have been previously exposed to that content or be able to comprehend that content, or (2) the SAT would “hold back” minorities because it underpredicts their future performance in college and in the labor force.

In reality, every question that contributes to a section or total SAT score undergoes a rigorous pretest phase so it is highly unlikely that a biased question, let alone series of questions, would ever reach the point where it would be scored. Psychometricians, or professionals who study the theory and technique of educational measurement, perform differential item functioning (DIF) analysis on test items to make sure that items do not unfairly advantage one group over another. Using DIF analysis, psychometricians ensure that people of the same ability level have the same probability of answering an item correctly, regardless of the race or gender group to which they belong. Items that exhibit even moderate DIF during the pretest phase are not included in the operational test. In addition, psychometricians examine if there is any difference in predictive ability among groups. Moreover, all items undergo extensive qualitative review by expert fairness reviewers.

As for some minority students being “held back” because their scores underpredict future performance in college and the labor force, the opposite is actually true. Minority (Hispanics and African Americans) performance on the SAT slightly overpredicts future college and employment performance. Psychometricians utilize differential validity and differential prediction procedures to examine the relationship between test scores and criterion variables (such as first-year grade point average, or FYGPA, in college). For example,
a test would exhibit differential validity if the correlation between SAT score and FYGPA differed for white students compared to black students. In addition, the issue of differential prediction occurs when a test under- or overpredicts a specific criterion (such as FYGPA) for a certain demographic group (see Rumor 3, below).5

Despite these measures to ensure good test development, differences in group average can still occur across different race and class groups.6 For example, the average difference in SAT score between white and African American students is roughly 100–110 points. Similar differences in average score are also found on other standardized tests, such as the NAEP, ACT, and GRE.7 These differences in scores on standardized tests do not occur as a result of bias of the test itself, but rather, external factors that play a role in quality education and ultimately student achievement, such as access to rigorous course work, books and stable peer groups, family support and resources, and academic preparation. Social disparities are seen across most metrics of attainment from standardized test scores to retention and graduation rates. Thus, academic achievement, including performance on the SAT, is a product of a set of social conditions, policy choices, and educational opportunities. Taken together, these activities play a role in student performance outcomes on the SAT and in school in general, and help to explain differences in average SAT score across social groupings.8

Attempts to address racial group differences through changes to the test alone have largely proven unsuccessful. For example, in 1976, the Golden Rule Insurance Company sued the Illinois Department of Insurance and Educational Testing Service (ETS), arguing that Illinois insurance licensure exams were biased against minority candidates. In an out-of-court settlement, the Illinois DOI and ETS agreed to attempt to reduce racial group differences in average scores by giving priority to questions that similar percentages of black and white candidates answered correctly.9 The approach was widely criticized by the psychometric community.10 A University of Illinois study, in applying the Golden Rule procedures to the SAT, found that including only items with small differences between black and white students resulted in a less reliable and valid test and did not necessarily reduce group differences in performance. Furthermore, the authors noted that the Golden Rule procedures were ineffective in detecting biased items.11

Rumor 3: The SAT Is Biased Against Women.
Opponents of the SAT often argue the test is biased against women. This rumor consists of two parts: (1) that average scores for females are lower than for males, and (2) the difference in average score means the SAT is biased against women. For example, according to the Center for Women Policy Studies:

"Next time your daughter comes home with a College Board score that's about 40 points below what you thought she'd get, don't blames yourself for not making her the right breakfast the day of the exam. Instead, you can blame the Scholastic Assessment Test [sic] itself."12

Fact: Disparities in SAT Scores Across Gender Occur as a Result of a Variety of Structural Factors Such as Course Preparation and Different Interests and Aspirations.
Similar to the claims that the SAT is biased against minorities, we would expect (1) women to miss a certain question or series of questions more than male students, because presumably women would not have been previously exposed to that content or be able to comprehend that content, or (2) the SAT would “hold back” women because it underpredicts their future performance in college. As noted in the response to Rumor 2, above, three procedures are used to ensure fairness on tests across demographic groups. First, psychometricians use differential item functioning (DIF) on test items to make sure that items that favor one group over another do not appear on the test. Pretest items that exhibit moderate DIF are not included in the operational test and any items thought to exhibit DIF on an operational test are excluded. Second, psychometricians utilize differential validity and differential prediction procedures to examine the relationship between test scores and criterion variables (see Rumor 2 for a detailed discussion of SAT bias). The issue of differential prediction occurs when a test under- or overpredicts a specific criterion (such as FYGPA) for females compared to males. Third, all items undergo extensive qualitative review by fairness expert reviewers. Overall, research shows that the SAT underpredicts female FYGPA, that is, that women perform better in college than expected given test scores.13 However, issues of differential validity or differential prediction are not measures of test bias.

Studies have examined the issue of underprediction for females and found that it may occur because of differences in grading stringency in the courses females typically take in college as compared with courses taken by men; further, including writing test scores in the prediction equation reduces underprediction of female first-year GPA.14

Tests are also reviewed to ensure that items do not include racist, sexist, or offensive content or content that could be gender biased. For example, the SAT does not include information about sports or military issues because these items may put males at an unfair advantage.15 Moreover, item review committees often review gender differences on item performance by comparing a matched group of males and females on certain items. Researchers at ETS, for example, found that males tend to score higher than females on scientific topics on the SAT. In addition, while the average math score of males tends to be higher, females typically score better on "pure mathematics" problems while males outperform women on word problems.16 In addition, tests are also designed to be gender and race balanced through including items that acknowledge the contributions of people of different genders, races, and nationalities.

Test Preparation/Test Prediction.
Test Preparation and Test Prediction claims are those that center around a specific strategy or predicted outcome related to the SAT. Test preparation rumors generally center on specific instances or small groups that follow a specific strategy or skill and significantly
over-/underperform as a direct result of this strategy. Test prediction rumors center around claims that the SAT is a very poor predictor of student performance. Test prediction claims can be offered with or without qualification (controlling for ...), and with or without reference ("it is a poor predictor relative to high school GPA").

**Rumor 4: Coaching Schools Can Raise SAT Scores Significantly.**
The entire rationale behind coaching schools is that students can significantly improve their SAT scores through participating in expensive testing preparation programs. For example, for some preparation courses, The Princeton Review promises to "improve your score by at least 150 points or [you] get your money back."17

**Fact: SAT Coaching Schools’ Programs Yield Modest Score Gains.**
Coaching schools offer courses where SAT preparation is structured by an official instructor who educates students in "test-taking strategies or ‘tricks’ [that] enable test takers to ‘beat the test’ … to take advantage of flaws in the test or in the testing system."18 Although test preparation companies like The Princeton Review and Kaplan offer money-back guarantees and tout increases of 100 points or more, the College Board has long maintained that test preparation is largely ineffective and the SAT is not “coachable."19 In addition to coaching schools, students may also hire a private tutor designed to help them prepare for the SAT.20

To be sure, students should prepare for the SAT by reviewing practice exams that are available through preparation guides as well as the College Board website. Students can also take advantage of free online SAT preparation content developed jointly between Khan Academy and the College Board to refresh key academic skills. However, research suggests that, on average, coaching results in a small, positive effect on SAT scores. While studies have provided slightly different estimates, on average, the effect of coaching on the math section is between 10–20 points, and between 5–10 points for the verbal section (15 to 30 points in total).21 These findings correspond with a meta-analysis conducted by Becker22 who found that, from 1953 to 1988, the expected gain from coaching was 9 points on verbal and 16 points on math. Grodsky offers one reason such importance is placed on such modest gains:

> "Expensive test preparation activities such as private classes and tutors allow affluent parents to assert their power over a process largely beyond their control, that of elite college admissions. These efforts appear to do little to increase their children’s chances of gaining admission to an elite college or university, but serve to at least partially assuage the needs of the affluent to feel they have done everything they can to pave the way for their children."23

**Rumor 5: The SAT Is Not a Good Predictor of Student Achievement in College.**
Critics of the SAT often argue that the test is not a good predictor of student achievement in college. For example, according to a recent article in *Time* by Leon Botstein,24 president of Bard College:

> "The blunt fact is that the SAT has never been a good predictor of academic achievement in college …. knowing something or how to do something in real life is never defined by being able to choose a “right” answer from a set of possible answers (some of them intentionally misleading) put forward by faceless test designers who are rarely eminent experts.”

Moreover, Elizabeth Kolbert,25 a staff writer for *The New Yorker*, writes: "the SAT measures those skills — and really only those skills — necessary for the SATs."

**Fact: There Is Strong Evidence That the SAT Is a Good Predictor of Student Achievement in College.**
The SAT is a good predictor of first-year performance in college, student retention, as well as cumulative college GPA.26 While high school grades are typically a slightly better predictor of college grades, SAT scores add significantly to the prediction.27 Thus, the best predictor of college performance is a combination of high school GPA and SAT critical reading, mathematics, and writing scores, which produce a correlation of .46 (uncorrected for range restriction) with first-year college GPA.28 From a college enrollment management perspective, evaluating both a student’s prior achievement along with a student’s SAT score makes a lot of sense. A student’s grade point average and course-taking history should matter, but how can a college decide if geometry at High School A is the same as geometry at High School B? A standardized test, such as the SAT, can help address that question. However, only looking at a single test on a single day probably should not count as much as a student’s high school experience. Thus, considering both together gets you a fairer and better picture of student achievement for college admission.

Research by Mattern and Patterson29 has shown that students with higher SAT scores have higher first-year college retention rates. For example, in their 2009 study, students with the highest SAT scores (2100–2400) had a more than 95% chance of returning for their second year of college compared to students with the lowest scores (600–890), who had less than a 64% chance of returning the second year. In addition, research has found that the SAT is also a good predictor of first-, second-, and third-year college performance as well as college completion.30

**Rumor 6: The SAT Does Not Add Much Beyond High School GPA in Predicting College Performance.**
Critics of the SAT often argue that high school GPA is a better predictor of college performance than the SAT. In addition, they argue that the SAT does not substantially add to the prediction of college success when high school grades are also considered.31 For example, as Joseph Soares, author of *SAT Wars* argues, "... the statistical case against the SAT is that it does not significantly enhance the ability of admissions staff to predict the academic potential of applicants."32
Fact: There Is Strong Evidence That the SAT Adds Substantially to High School GPA in Predicting College Performance.
Although the SAT may better predict college performance than high school grades for some students and for some institutions, overall, a large body of research has shown that the best predictor of first-year college performance is a combination of high school GPA, SAT verbal scores, and SAT math scores. For example, Rebecca Zwick found that the correlation between SAT verbal and math scores and first-year GPA was about .36 and between high school grades and first-year GPA was slightly higher (about .39). However, the best predictor of college performance was a combination of high school GPA and SAT verbal and math scores, which produced a correlation of .48 with first-year GPA, resulting in an increment of .09 when SAT was added (uncorrected for range restriction).

In addition, while first-year GPA tends to be overpredicted for some minority groups, this overprediction tends to be greater when high school GPA is used as the sole predictor of college performance, compared to using a combination of SAT and high school GPA. A variety of factors are thought to contribute to this greater overprediction, such as different grading standards, rigorous course work, and access to resources across high schools. For example, while high schools tend to have similar grade distributions, they vary in achievement levels of students. Thus, including SAT scores in the prediction of college performance helps to mitigate this overprediction by offering a standard for comparing students across different schools.

Test Utilization
Test preparation and test prediction claims often support the notion that excluding the SAT from admission criteria and utilizing high school GPA and class rank, as well as other factors, may help to increase diversity across college campuses.

Rumor 7: SAT-Optional Institutions Have Significantly More Diverse Entering Classes Than Institutions That Use the SAT.
Organizations such as the National Center for Fair and Open Testing (FairTest) argue that standardized admission tests such as the SAT create a significant barrier to entry for academically qualified minority, first-generation, and low-income applicants. In an attempt to address these concerns, some colleges and universities, such as Mount Holyoke College and Bates College, and more recently Temple University, have adopted test-optional policies. However, out of more than 1,800 four-year nonprofit colleges and universities across the United States, 78% still require or recommend a college entrance exam as part of their admission policy.

Fact: Removing the SAT as a Criterion for College Admission Does Not Significantly Enhance Diversity on College Campuses.
Generally, the most common test-optional approach that colleges have used to attempt to increase diversity on their campuses is class rank (or percentage plans). Recent research has shown that as a whole, institutions that have shifted to test-optional policies have "done little to meet their manifest goals of expanding educational opportunity for low-income and minority students." Some critics of the SAT have argued that scores should be replaced with class rank as an admission criterion because it is more equitable than SAT scores and grades; and some university systems (California, Texas, Florida) have adopted these policies. Through this approach, a fixed percent of students (usually the top 4–20%) are admitted from each high school, regardless of test scores, school quality, or rigor of course work. However, percent plans do not ultimately increase diversity on campus. For example, a recent study of 193 selective colleges and universities by Carnevale, Rose, and Strohl found that using a 10% plan would have resulted in a qualified applicant pool that was, on average, 6% black and 11% Hispanic, which did not significantly differ from the existing racial/ethnic composition of schools, which was 4% black and 7% Hispanic. In addition, the percent plan approach may also create other problems, such as admitting top-ranked students from lower quality high schools, which can lead to higher dropout rates and lower retention rates.

Rumor 8: Students Who Do Not Submit SAT Scores Perform Just as Well as Those That Do Submit Scores at SAT-Optional Institutions.
In recent years, an increasing number of colleges and universities across the United States have instituted test-optional admission policies. These policies are thought to increase the diversity of applicant pools. Supporters of this admission strategy argue that students who do not submit SAT scores do just as well in college as students that do submit their scores, and thus, the SAT is not a necessary component of the admission process.

Fact: The Best Predictor of College Success Is a Combination of SAT Scores and High School GPA.
The SAT is a useful predictor of college success. For example, research has found that, overall, students that had high school GPAs of A- (3.70) or better had an 85% chance of success in their first year of college (as measured by receiving at least a 2.5 GPA), but success rate varied by SAT score. Specifically, students that scored between 1410 and 1600 on the SAT had a 95% chance of success, while students with scores of 1000 or lower only had a 62% chance of success.

One issue related to test-optional policies is the difference in SAT score of nonsubmitter students versus students who did report their scores. For example, with 98% of enrolled students’ scores collected at Providence College, SAT submitters had an average score of roughly 1200, while nonsubmitters had an average score of 1100. The average SAT scores for nonsubmitter students versus submitter students is also lower at Bowdoin College, which adopted a test-optional policy in 1969. Research on student performance at Bowdoin has shown that, on average, the students that did not submit their scores (who also typically have lower SAT scores) perform worse in their first year of college compared to students who do submit their scores (who typically have higher scores). In addition, research on admission decisions by Mount Holyoke College, a test-optional institution, found that, overall, students who matriculated to the college without submitting their SAT scores had a slightly lower first-year GPA (3.24) compared to SAT submitters (3.35).
Ironically, then, knowledge of whether or not a candidate submits SAT scores to a test-optional school can itself predict performance in college, a vivid illustration of the incremental information admission test scores offer.

**Rumor 9: The SAT Tests “Aptitude” Rather Than Skills People Actually Learn in School.**

The SAT, originally named the Scholastic Aptitude Test, was developed from the Army Alpha test, which was used to support military assignments in World War I. These tests were direct descendants of Intelligence Quotient (IQ) tests that originated in the 1900s in the United States. In the early 20th century, the eugenics movement became popular and scientists in this field argued that intelligence was genetic, not part of learned behavior. As researchers learned more about educational measurement and testing, they realized that knowledge (i.e., reading, writing, math) is much more a product of a student’s environment than a reflection of one’s capacity at birth. According to sociologist Christopher Jenks:

> “Many people … think that both intelligence and aptitude are innate traits … yet almost all psychologists now agree that while an individual’s score on an intelligence or aptitude test depends partly on his or her genetic makeup, it also reflects a multitude of environmental influences.”

Despite this research, there are still misconceptions that standardized tests are a measure of aptitude rather than knowledge learned in school.

**Fact: The SAT Tests Student Knowledge of Skills (Writing, Reading, and Math) That Are Taught Every Day in School.**

The SAT is a standardized assessment of critical reading, mathematical reasoning, and writing skills that students learn over time in school and need to be successful in college. Generally, students do well on the SAT because they have taken and are exposed to rigorous course work in high school. For example, taking rigorous courses in math is positively associated with SAT scores. In addition, the SAT is a useful predictor of student success in college including first-year GPA, student retention, and cumulative college GPA.

The redesigned SAT will be grounded in the real world, and cover questions directly related to the work performed in college and career. For example, the math section of the exam will focus on essential areas that students need to be quantitatively literate such as using ratios, percentages, and proportional reasoning to solve problems in science and social science and career contexts. In addition, the evidence-based reading and writing section of the exam will focus on words that students will use throughout their lives — in high school, college, and beyond — and will include texts in the humanities, science, history, and social studies and career-related sources.

Notes


13. Mattern, Patterson, Shaw, Kobrin, and Barbuti, *Differential Validity and Prediction of the SAT*.


15. Rebecca Zwick, *Fair Game*?

16. Ibid.


29. Mattern and Patterson, *Is Performance on the SAT Related to College Retention?*


40. Rebecca Zwick, *Fair Game?*


