

SAT Benchmarks

Development of a College Readiness Benchmark and its
Relationship to Secondary and Postsecondary School Performance

By Jeffrey Wyatt, Jennifer Kobrin, Andrew Wiley, Wayne J. Camara, and Nina Proestler



BENCHMARKS

Jeffrey Wyatt is an assistant research scientist at the College Board.

Jennifer Kobrin is a research scientist at the College Board.

Andrew Wiley is an executive director at the College Board.

Wayne J. Camara is vice president of research and development at the College Board.

Nina Proestler is a graduate intern at the College Board.

Mission Statement

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

About the College Board

The College Board is a mission-driven not-for-profit organization that connects students to college success and opportunity. Founded in 1900, the College Board was created to expand access to higher education. Today, the membership association is made up of more than 5,900 of the world's leading educational institutions and is dedicated to promoting excellence and equity in education. Each year, the College Board helps more than seven million students prepare for a successful transition to college through programs and services in college readiness and college success — including the SAT® and the Advanced Placement Program®. The organization also serves the education community through research and advocacy on behalf of students, educators and schools.

For further information, visit www.collegeboard.org.

© 2011 The College Board. College Board, Advanced Placement, Advanced Placement Program, AP, SAT and the acorn logo are registered trademarks of the College Board. Inspiring Minds and ReadStep are trademarks owned by the College Board. PSAT/NMSQT is a registered trademark of the College Board and National Merit Scholarship Corporation. All other products and services may be trademarks of their respective owners. Visit the College Board on the Web: www.collegeboard.org.

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



Contents

Executive Summary	5
Introduction	6
The Need for College Readiness Benchmarks.....	7
The Current Study.....	8
Method	8
Samples	8
Measures	9
Analyses	11
Results	12
Determination of College Readiness Benchmarks	13
SAT® College Readiness Benchmarks and College Enrollment and Performance.....	14
SAT College Readiness Benchmarks and High School Performance	15
SAT Benchmarks and Student Demographics	21
Discussion	22
Limitations	24
References.....	25

Appendix A: Comparison of Sample 1 to Population of Four-Year Institutions by Key Variables.....	28
Appendix B: Percent of the 2010 Cohort (Sample 3) Meeting the SAT Benchmark by Academic Rigor Index (ARI).....	29
Appendix C: Student Performance on English Advanced Placement® (AP®) Exams For Those Meeting and Not Meeting the SAT-CR Benchmark.....	29
Appendix D: Student Performance on English Advanced Placement (AP) Exams for Those Meeting and Not Meeting the SAT-W Benchmark	30
Appendix E: Student Performance on Mathematics Advanced Placement (AP) Exams for Those Meeting and Not Meeting the SAT-M Benchmark	30

Tables

Table 1: Demographic Characteristics of the 2007 Sample of Students (Sample 1), the 2007 U.S. Cohort (Sample 2), and the 2010 U.S. Cohort (Sample 3)	13
Table 2: SAT Benchmarks and Attainment	13
Table 3: The Percentage of Students Enrolled in Postsecondary Education Who Met and Did Not Meet the SAT Benchmark (Sample 2)	14
Table 4: Mean FYGPA for Students Who Met and Did Not Meet the SAT Benchmark	15
Table 5: Percentage of the 2010 Cohort (Sample 3) that Met the Benchmark by Academic Variables	16
Table 6: Percentage of the 2010 Cohort (Sample 3) that Obtained the Benchmark: By Demographic Variables.....	22

Figures

Figure 1: The percentage of students retained by benchmark status	15
Figure 2: Percentage meeting the SAT benchmark by Academic Rigor Index (ARI) score	17
Figure 3: Student performance on English Advanced Placement (AP) exams by college readiness status on the critical reading section	18
Figure 4: Student performance on English Advanced Placement (AP) Exams by college readiness status on the writing section	19
Figure 5: Student performance on Math Advanced Placement (AP) Exams by college readiness status on the math section.....	20

Executive Summary

The current study was part of an ongoing effort at the College Board to establish college readiness benchmarks on the SAT®, PSAT/NMSQT®, and ReadStep™ as well as to provide schools, districts, and states with a view of their students' college readiness. College readiness benchmarks were established based on SAT performance, using a sample of approximately 68,000 students across 110 four-year institutions. The college readiness benchmark was calculated as the SAT score associated with a 65 percent probability of earning a first-year GPA of 2.67 (B-) or higher. The SAT benchmark determined in this study was 1550 for the composite¹. Individual benchmark scores were also calculated for the critical reading, mathematics, and writing sections to provide indicators of student proficiency in each of these subjects, resulting in a benchmark score of 500 on each section. Once the benchmark scores were obtained, a series of analyses were conducted to establish the validity of the benchmarks for indicating college readiness. These analyses examined the relationship between college readiness benchmark attainment and high school academic performance measures (curriculum, HSGPA, and AP performance), along with college indicators including enrollment, FYGPA, and retention. The results showed that students meeting the benchmark are more likely to enroll in college; return for their second and third years of college; earn higher grades in both high school and college; and are more likely to have taken a core curriculum as well as more rigorous courses in high school than those not meeting the benchmark.

¹ The college readiness benchmark refers to the composite benchmark score of 1550, unless the individual section score benchmarks are mentioned specifically.

Introduction

Over the past 10 to 20 years, educational attainment of students in the United States has stagnated, while college degree attainment rates in most other industrialized nations have continued to grow. The National Center for Higher Education Management Systems (NCHEMS) estimates that only 39 percent of adults aged 25 to 34 in the United States hold an associate degree or higher (NCHEMS, 2009). In 1995, the United States was ranked second in the percentage of students who received a postsecondary degree, but fell to 15th among 25 countries in 2005 (Organization for Economic Cooperation and Development [OECD, 2010]) as increases in graduation rates failed to keep pace with those of other developed nations (OECD, 2008). In 2007, 31 percent of 25- to 34-year-olds in the United States had attained a bachelor's degree, which represents only a 2 percent increase since the year 2000 (U.S. Department of Education, 2008). These types of trends formed the backdrop when President Obama set a goal for the United States to have the highest proportion of college graduates in the world by 2020 (<http://www.whitehouse.gov/issues/education>).

Previous research has demonstrated myriad economic and societal benefits associated with obtaining a college degree. A college education is related to significantly higher wages and having less of a financial burden on the state (e.g., incarceration, public health, welfare) (Baum & Payea, 2004). In 2005, the typical full-time year-round worker in the United States with a four-year degree earned \$50,900, 62 percent more than the \$31,500 earned by the typical full-time year-round worker with only a high school diploma (Baum & Ma, 2007). A college degree is also associated with better citizenship, political activism, volunteering, and lawful behavior. The presence of such externalities and positive spillover effects for the nation and state are an essential reason for public support of transfer systems in higher education (Goldberg & Smith, 2008).

Although the benefits of obtaining a degree have been widely documented, colleges and universities still struggle with a high percentage of enrolled students who do not complete their degree. NCHEMS (2009) reported that 56 percent of students who entered a four-year U.S. college or university in 2001 graduated within six years, which is only a slight increase from 52 percent of students entering in 1991. Degree attainment rates are even lower at two-year schools as approximately 32 percent of entering students earn either an associate or bachelor's degree within six years (Attewell, Heil, & Reisel, 2010).

Research on college completion has examined a wide range of factors. A present theoretical model posits two factors: academic and social integration (Tinto, 1993). The first factor addresses the match between a student's cognitive skills and abilities and the demands of the academic institution and program, while the second factor concerns a student's social engagement to the institution and other students. A variety of factors have been examined in research on college completion: academic preparation in high school; financial support; student demographic characteristics; employment; remediation; and institutional differences.

Attewell et al. (2010) recently reported that no single dominant factor was associated with college completion and that there is substantial variation in the factors related to graduation within six years from a two-year or four-year institution. Remediation and high school academic preparation were not significant factors in degree attainment at community colleges. Instead, financial support, hours worked, and demographic characteristics were the most significant predictors after controlling for all other factors. At four-year institutions, the results were different. High school academic preparation, as measured by admission test scores, high school grades, and academic rigor, was the most important factor, although remedial courses were only significantly related to degree attainment at the least selective

four-year colleges, after controlling for other factors. The number of hours a student worked was also consistently related to graduation within six years across all types of institutions, and nontraditional status (e.g., part-time, delayed entry) and integration (social and academic) into college were significant predictors of graduation in the least selective and moderately selective four-year institutions.

Students who require remedial course work in college have starkly lower graduation rates than those who enter college prepared for college-level course work. Among 12th-graders attending a postsecondary institution in 1992, 17 percent of those enrolled in a remedial reading course eventually earned a bachelor's degree, and 13 percent earned an associate degree or certificate. Among those taking two or fewer remedial mathematics courses, 27 percent earned a bachelor's degree, and 14 percent earned an associate degree or certificate. Among those not taking any remedial course, 58 percent earned a bachelor's degree and 11 percent earned an associate degree or certificate (Wirt, Choy, Rooney, Provasnik, Sen, & Tobin, 2004).

The exact percentage or number of students who require remediation is hard to pinpoint, with some studies estimating that 28 percent of entering college students are remediated (Wirt et al., 2004). Remediation rates differ by institution type as public two-year schools have remediation rates of 42 percent, while public four-year colleges have remediation rates of 20 percent, and private four-year colleges have remediation rates of 12 percent. Remediation rates are closely related to high school academic preparation as measured by tests, grades, and academic course-taking patterns (Attewell et al., 2010). Remediation rates are also closely associated with other key student demographic variables, such as income, race/ethnicity, and parental education. High school graduates from the highest income levels are three times more likely than students in the lowest income level to be academically prepared for college (Presley & Gong, 2005). A recent report by the National Council for Education Statistics (NCES) divided college students into quintiles based on socioeconomic status (SES) and found that 63 percent of students in the bottom quintile (lowest SES) enrolled in a remedial course, compared to 25 percent in the top quintile (highest SES). Remediation rates also differ by race and ethnicity. NCES estimates that 62 percent of African American and 63 percent of Hispanic students take at least one remedial class, compared to 35 percent of white students (Wirt et al., 2004). Further compounding matters, African American, Hispanic, and low-income students are also more likely to be the first in their family to attend college (Chen & Carroll, 2005). First-generation students have generally been exposed to a less rigorous curriculum during high school than their non-first-generation counterparts. Approximately 55 percent of first-generation students require remediation (at least one remedial class), about twice as often as students whose parent(s) had obtained a bachelor's degree (27 percent) (Chen & Carroll, 2005).

The Need for College Readiness Benchmarks

As the education community pays increasing attention to issues surrounding retention and remediation, objective and fair measures of student preparedness for college become increasingly critical. There is a critical need to inform students, teachers, parents, and counselors about whether students have the academic skills to succeed in college and to design interventions to help students correct deficiencies so that they can enter college prepared to succeed. There has been increasing attention on benchmarking both at the state and national level, as there is widespread recognition of the need to gauge the college readiness of students (McNeil, 2008). Benchmarks can also serve the role of strengthening the college culture and expectations for students (Corwin & Tierney, 2007). One of the most comprehensive and far-reaching initiatives is the Common Core Standards and Assessments,

which is a major effort to establish consistent content and performance standards related to college readiness.

Several states have incorporated empirically based benchmarks (e.g., NAEP, ACT, College Board) in setting cut scores on state tests to ensure college readiness (Camara, in preparation).

The Current Study

The current study was part of an ongoing effort at the College Board to establish college readiness benchmarks on the SAT, PSAT/NMSQT, and ReadStep; and to provide schools, districts, and states with a view of their students' college readiness. Kobrin (2007) used a model-based method (i.e., logistic regression) to derive two SAT benchmarks, one corresponding to a 65 percent probability of getting a 2.7 first-year grade point average (FYGPA) and one at a 65 percent probability of getting a 2.0 FYGPA. Kobrin's study was conducted on the pre-2005 SAT, which had only the mathematics and critical reading sections, and was scored on a 400 to 1600 scale. Kobrin determined that scores of 1180 and 800 were associated with a 65 percent probability of obtaining a FYGPA of B- (2.7) or higher and C (2.0) or higher, respectively.

The primary purpose of the current study is to extend the work of Kobrin (2007) by identifying college readiness benchmarks on the SAT; and to collect evidence of the validity of the benchmarks for indicating college readiness by examining the relationship of these benchmarks to other measures of high school performance, college performance, and student demographic characteristics. These benchmarks are intended to provide information on the college readiness of groups of students (e.g., aggregated by school, district, state, or nation). In considering the college readiness of individual students, many factors should be considered in addition to test scores. These may include high school GPA (HSGPA), completed course work, recommendations, and noncognitive factors.

Method

Samples

The current study used data from students taking the SAT who were reported to graduate from high school in 2007 and 2010 (hereafter referred to as the 2007 and 2010 graduating seniors cohorts or the 2007 and 2010 cohorts). The students' most recent SAT score was used for all analyses in this study. Three separate samples were created for the analyses in this study.

Sample 1. The first sample was derived from the 2007 graduating seniors' cohort and was used to investigate the relationship between SAT benchmark scores and college grades. In order to create this sample, SAT records from the College Board were matched to college performance data for students who entered one of 110 higher education institutions that participated in a national validity study for the SAT (Kobrin, Patterson, Shaw, Mattern, & Barbuti, 2008) and supplied data (i.e., course grades, FYGPA, and retention) for their 2007 entering first-year class. Institutions were recruited to be representative of the target population, which included the 726 four-year institutions that received at least 200 SAT score reports in 2005. The sample of institutions was diverse with respect to region of the U.S., control (i.e., public versus private), selectivity, and size.

Data from the final sample of 110 institutions were matched to College Board records that

included SAT scores, self-reported HSGPA, and demographic information. This sample used the same dataset restrictions specified in prior research reports (Wyatt, Wiley, Camara, & Proestler, in press) and was limited to students who had taken the SAT on or after March 2005 (which included writing), reported their HSGPA, and provided responses to the course work questions on the SAT Questionnaire (SAT-Q). By including only students with HSGPA and course work data, the relationship between benchmark attainment and other measures of high school performance could be investigated. There were 67,644 students in Sample 1.

Sample 2. The second sample was also derived from the 2007 SAT graduating seniors cohort. This sample was limited to students from the United States who took the SAT on or after March 2005 and was matched to data from the National Student Clearinghouse (NSC) to obtain initial postsecondary enrollment data for these students. This dataset was used to investigate the relationship between SAT benchmark scores and college enrollment. NSC tracks student enrollment and degree attainment for over 3,100 two- and four-year colleges and universities in the United States (a list of participating institutions is located at www.studentclearinghouse.org), equivalent to 91 percent of the U.S. college-going population. This dataset contains 1,419,714 students.

Sample 3. The third sample was used to examine the relationship between the SAT benchmark scores and overall student performance, demographic characteristics, and other high school performance measures. The dataset included 1,457,489 students from the 2010 cohort who attended a high school within the United States.

Measures

SAT® Scores. SAT scores were obtained for all three samples. The SAT consists of the critical reading, mathematics, and writing sections; each section has a score scale range of 200 to 800 with 10-point increments. The SAT composite score is the sum of all three section scores and therefore has a score scale range of 600 to 2400. Further information on the SAT can be found at <http://professionals.collegeboard.com/testing/sat-reasoning>.

Gender. Students reported their gender (female or male) when they completed the SAT-Q.

Ethnicity. Students indicated their race/ethnicity on the SAT-Q in one of eight categories: (1) American Indian or Alaska Native, (2) Asian, Asian American, or Pacific Islander, (3) black or African American, (4) Mexican or Mexican American, (5) Puerto Rican, (6) Other Hispanic, Latino, or Latin American, (7) white, and (8) other. The categories 4, 5, and 6 were combined into a single category titled “Hispanic.”

Best Language. Students reported their best language on the SAT-Q. Response options included “English Only,” “English and Another Language,” and “Another Language.”

Highest Parental Education. Students’ highest level of parental education was also derived from self-reported data on the SAT-Q. Student responses were provided for both mother’s and father’s highest educational level. The highest degree (i.e., No High School Diploma, High School Diploma, Associate Degree, Bachelor’s Degree, or Graduate Degree) of either parent was used to create this variable.

High School GPA (HSGPA). Cumulative HSGPA was self-reported by students registering to take the SAT. Scores were reported in letter grades ranging from an F (below 65) to an A+ (97–100). High school grades were then converted to a 0–4.33 scale. While HSGPA was self-reported, a number of studies have suggested that the correlation between self-reported HSGPA and actual HSGPA is between 0.74 and 0.85 (Kuncel, Crede, & Thomas, 2005; Maxey

& Ormsby, 1971; Schiel & Noble, 1991; Shaw & Mattern, 2009), indicating that self-reported HSGPA is a reliable measure of high school performance.

Core Curriculum. A core curriculum was defined as having completed four years of English, three years of math, three years of science, and three years of social science/history during high school. Students reported the courses that they took or planned to take during high school on the SAT-Q.

Academic Rigor Index (ARI). Wyatt et al. (in press) developed an academic rigor index (ARI), scaled from 0 to 25, that is designed to measure the degree of challenge associated with students' high school course work. This index is calculated from students' self-reported high school course work within five academic subjects: English, mathematics, science, social science/history, and foreign/classical languages. Within each of the five subjects, between 0 and 5 points are awarded based on the difficulty of courses, with 5 points indicating the most rigorous curriculum. The number of points awarded within each subject is summed to create the ARI on a 0-25 scale, with 25 representing the highest level of rigor and 0 representing the lowest. For more information see Wyatt et al. (in press).

Advanced Placement® (AP®) Performance. AP Exams are traditionally administered at the conclusion of college-level courses taught to high school students within their normal high school setting. These courses must conform to an AP curriculum, which provides guidance on the depth and breadth of content that should be covered during the course. At the completion of the course, students may choose to complete a standardized exam that measures domain-specific college-level knowledge and skills. The exam is scored from 1 to 5, with a 5 representing the equivalent of an A in the corresponding introductory college-level course, a 4 representing a B, a 3 representing a C, a 2 representing a D, and 1 representing an F. Most often, colleges award credit for AP Exam scores of 3 or higher. Data from Sample 3 were matched to students' AP records, and all students from the 2010 graduating seniors cohort who also took an AP Exam in either English or mathematics were included in this analysis. Appendixes C–E provide the number of students who took both the SAT and English or mathematics AP Exams.

Percent Enrolled in College. Data from the 2007 SAT graduating seniors cohort were matched to the NSC database (Sample 2). Using the 2007 sample, the percent of students who actually enrolled in a higher education institution was calculated. As a further refinement, the percent of students enrolled in college was disaggregated by college type (two-year and four-year).

First-Year GPA (FYGPA). For Sample 1, FYGPA was obtained from participating colleges and universities. The values of FYGPA ranged from 0.00 to 4.19 (mean = 2.93, SD = .73), with only 24 students having an FYGPA greater than 4.00.

Retention. For Sample 1, institutions indicated whether students returned for the fall semester of their sophomore year (retention to second year) and whether students returned for their junior year (retention to third year). It should be noted that while this variable does provide a reasonable estimation of student persistence, it does not account for students who have transferred to another institution and persisted at that other institution.

Analyses

The first phase of the study used Sample 1 and was designed to set a benchmark that could be used to assess college readiness. Logistic regression was used to set the SAT benchmarks, using as a criterion a 65 percent probability of obtaining an FYGPA of a B- or higher within each of the 110 institutions. Logistic regression is a statistical method that uses binary outcome information (e.g., success versus failure) to predict the probability of success based on one or more predictor variables (in this case, an SAT score). Separate logistic regression equations were estimated for each institution, using the sum of SAT critical reading, mathematics, and writing section scores to predict the probability that a student would be successful in terms of earning an FYGPA of 2.67 (or B-) or higher. The 65 percent probability of success was chosen because this level has been used in other research, including research focused on the National Assessment of Educational Progress (NAEP) and other educational studies as an appropriate standard for defining success in a domain. A 65 percent probability level has also been recommended by subject-area experts as an appropriate standard for knowledge or success in a domain (Beaton & Allen, 1992; Zwick, Senturk, & Wang, 2001).

In 2008, the College Board assembled an expert panel of educators and policymakers to participate in a judgmental standard-setting process to recommend both probability and criterion for defining college readiness. The panel agreed that a probability in the range of 60 to 75 percent would be the most appropriate. The FYGPA criterion of 2.67 was also recommended because it represents a B- at most colleges and seems appropriate and sufficiently rigorous when considering academic success of freshmen (Kobrin, Patterson, Wiley, & Mattern, under review). While the expert panel provided a probability level and criterion to define college readiness, they strongly recommended using six-year graduation as an indicator of college success. However, as these data were not available, the committee agreed that an FYGPA of a B- or higher was indicative of future success in college and could be used as a reasonable criterion. Research has established a strong correlation between FYGPA and retention, and the likelihood of continuing college for four years increases substantially for students with higher FYGPAs (Allen, 1999; Murtaugh, Burns, & Schuster, 1999).

Benchmarks were established for each SAT section (critical reading, mathematics, and writing), and for the SAT composite (the sum of the scores on the three SAT sections) within each of the 110 institutions. Separate logistic regression equations were estimated for each of 109 institutions, and benchmark scores were calculated. One institution was not used because 100 percent of their students obtained a FYGPA of 2.67 or higher. After the institution-level benchmarks were computed, any out-of-range values (e.g., lower than 600 or higher than 2400) were excluded, and the estimates were averaged, weighted by the institution-level sample sizes.*

As a result, the number of valid institutions differed by section: there were 104 valid institutions for critical reading, 102 for mathematics, 107 for writing, and 107 for the composite. Once the SAT benchmark score was computed, it was rounded down to the nearest legitimate interval (e.g., a critical reading score of 504 would be rounded down to 500 since it is not possible for a student to score between 500 and 510).

*Inverse prediction was used to obtain an estimate of the exact SAT composite score that is associated with a particular probability of success. One potential weakness of this approach is that it is possible to obtain benchmarks that fall outside of the actual SAT score range. This issue was handled by excluding those institutions whose benchmarks were outside the range of scores that a student could actually obtain.

Once benchmark scores were obtained, a series of analyses were conducted to compare the academic performance of students who met or exceeded the benchmarks and those who did not. These comparisons were made both on high school academic measures (curriculum, HSGPA, and AP performance) and college indicators including enrollment, FYGPA, and retention.

Results

Table 1 compares the demographic characteristics for all three samples. Sample 1 is largely representative of the SAT taker population (Sample 2 and Sample 3), with slight differences in composition with respect to gender, ethnic/racial, and best language subgroups. The colleges participating in Sample 1 are largely representative of all four-year institutions with respect to size, selectivity, geography, and institutional control (public/private). More detailed information is available in Appendix A.

Table 1

Demographic Characteristics of the 2007 Sample of Students (Sample 1), the 2007 U.S. Cohort (Sample 2), and the 2010 U.S. Cohort (Sample 3).

		Sample 1		Sample 2		Sample 3	
		Number	Percent	Number	Percent	Number	Percent
All		67,644		1,419,714		1,457,489	
Gender	Female	39,189	57.9	762,803	53.7	782,804	53.7
	Male	28,455	42.1	651,243	45.9	674,685	46.3
Race/ Ethnicity	American Indian	358	0.5	9,659	0.7	8,295	0.6
	African American	5,796	8.6	155,244	10.9	191,566	13.1
	Asian American	6,809	10.1	119,026	8.4	133,351	9.1
	Hispanic	6,951	10.3	162,633	11.5	215,181	14.8
	White	43,130	63.8	812,551	57.2	817,915	56.1
	Other	1,908	2.8	46,552	3.3	39,769	2.7
	No Response	2,692	4.0	114,049	8.0	51,412	3.5
Best Language	English	61,503	90.9	1,215,036	85.6	1,181,661	81.1
	English and Another Language	4,594	6.8	103,479	7.3	214,827	14.7
	Another Language	732	1.1	27,142	1.9	31,337	2.2
	No Response	815	1.2	74,057	5.2	29,664	2.0

Determination of College Readiness Benchmarks

As described earlier, college readiness benchmarks were obtained through logistic regression to determine the SAT score associated with a 65 percent probability of obtaining an FYGPA of a B- or higher. Once the SAT benchmark scores were obtained, each score was rounded down to the nearest valid SAT score. The SAT benchmarks were 1550 for the composite and 500 for each section, critical reading (SAT-CR), mathematics (SAT-M), and writing (SAT-W). Table 2 shows the percent of SAT takers from the 2010 cohort that met the Composite benchmark and each of the section score benchmarks. About half of all students met the college readiness benchmark in the critical reading section, while slightly more (54 percent) met the mathematics section benchmark and slightly fewer (46 percent) met the writing section benchmark. About 43 percent of students met the composite benchmark score and were considered ready for college.

Table 2

SAT Benchmarks and Attainment

	Composite	Critical Reading	Mathematics	Writing
Benchmark	1550	500	500	500
Number Meeting Benchmark	630,704	734,749	792,448	670,256
Percent Meeting Benchmark	43	50	54	46

Note. There were 1,457,489 students in the sample.

SAT® College Readiness Benchmarks and College Enrollment and Performance

Student enrollment data were evaluated using students from Sample 2. Table 3 shows the percentage of students enrolled in a two-year or four-year postsecondary institution, and the percentage of students not enrolled in a postsecondary institution, based on whether or not they met the SAT College Readiness benchmark. For students meeting the benchmark, there was a very high likelihood of initial enrollment (78 percent) in a four-year institution. For students who did not reach the SAT College Readiness benchmark, this percentage was notably lower (46 percent). Students not meeting the benchmark were more than three times as likely to enroll in a two-year institution as those who met the college readiness benchmark. About 25 percent of students who did not meet the college readiness benchmark failed to enroll in any higher education institution compared to 14 percent of students who did meet the benchmark.

Table 3

The Percent of Students Enrolled in Postsecondary Education Who Met and Did Not Meet the SAT Benchmark (Sample 2)

Readiness Status	Not Enrolled	Enrolled in 2-Yr.	Enrolled in 4-Yr.	N
Met Benchmark	14	8	78	629,552
Did not Meet Benchmark	25	29	46	790,162

Note: Initial postsecondary enrollment data were obtained from the National Student Clearinghouse (NSC).

Table 4 compares the mean FYGPA for students in Sample 1 based on whether or not they met the SAT College Readiness benchmark. The mean FYGPA for students achieving the benchmark was 3.12, compared to 2.57 for those who did not meet the benchmark. (The overall mean FYGPA for all students in Sample 1 was 2.93). The difference in means between the two groups was 0.55 and was statistically significant ($t(40,135) = 92.45, p < .001, d = 0.78$). The medium-to-large effect size suggests that student attainment of the benchmark score is substantially related to subsequent college performance as measured by FYGPA.

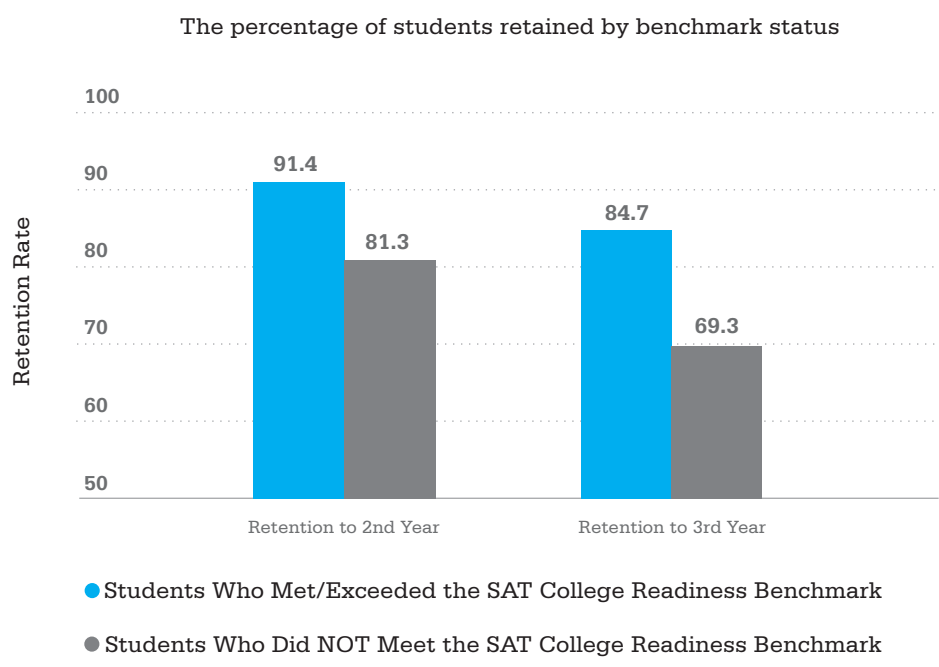
Table 4 also shows that the percent obtaining an FYGPA of 2.67 or higher was considerably higher for students meeting the SAT benchmark score (79.3 percent) than for students not meeting the benchmark (50.4 percent). It is important to note that the students used for this analysis (Sample 1) were all enrolled in a four-year institution and therefore were on average more academically prepared than the general population of SAT takers. As for students who did not meet the benchmark, this sample draws from the 46 percent of those enrolled in a four-year college who have higher average HSGPA and SAT scores than the overall population of students not meeting the benchmark. Thus, fewer than 50.4 percent of the general population of SAT takers not meeting the benchmark would be expected to obtain an FYGPA of 2.67 or higher (Table 4) if enrolled in a four-year college.

Table 4
Mean FYGPA for Students Who Met and Did Not Meet the SAT Benchmark

Readiness Status	Mean	SD	Percent with 2.67 or Higher	Range
Met Benchmark	3.12	.64	79.3	0.00–4.19
Did Not Meet Benchmark	2.57	.76	50.4	0.00–4.03

Retention statistics were also calculated using a subset of Sample 1. Ninety-one of the original 110 institutions participating in the SAT validity study continued their participation through year 2 and up to the beginning of year 3. This subset of 58,287 students was used to obtain retention data. Figure 1 shows the percent of students retained to the second and third years. The retention rate to the second year of college was about 10 percentage points higher for students meeting the benchmark compared to that of students who did not meet the benchmark. For retention to the third year, the gap widened to approximately 15 percentage points.

Figure 1



SAT College Readiness Benchmarks and High School Performance

A retrospective analysis was conducted to determine if students who met the SAT College Readiness benchmark had higher academic achievement during high school. Table 5 reports the percent of students meeting the benchmark across a series of other measures of high school preparation and performance. As would be expected, there is a strong relationship between the SAT College Readiness benchmark and these measures of high school performance. For example, when looking at HSGPA, approximately 9 to 12 percent of students with a HSGPA of C (C+, C, or C-) or lower met the benchmark, compared to over 57 to 84 percent of those with a HSGPA of A (A+, A, or A-).

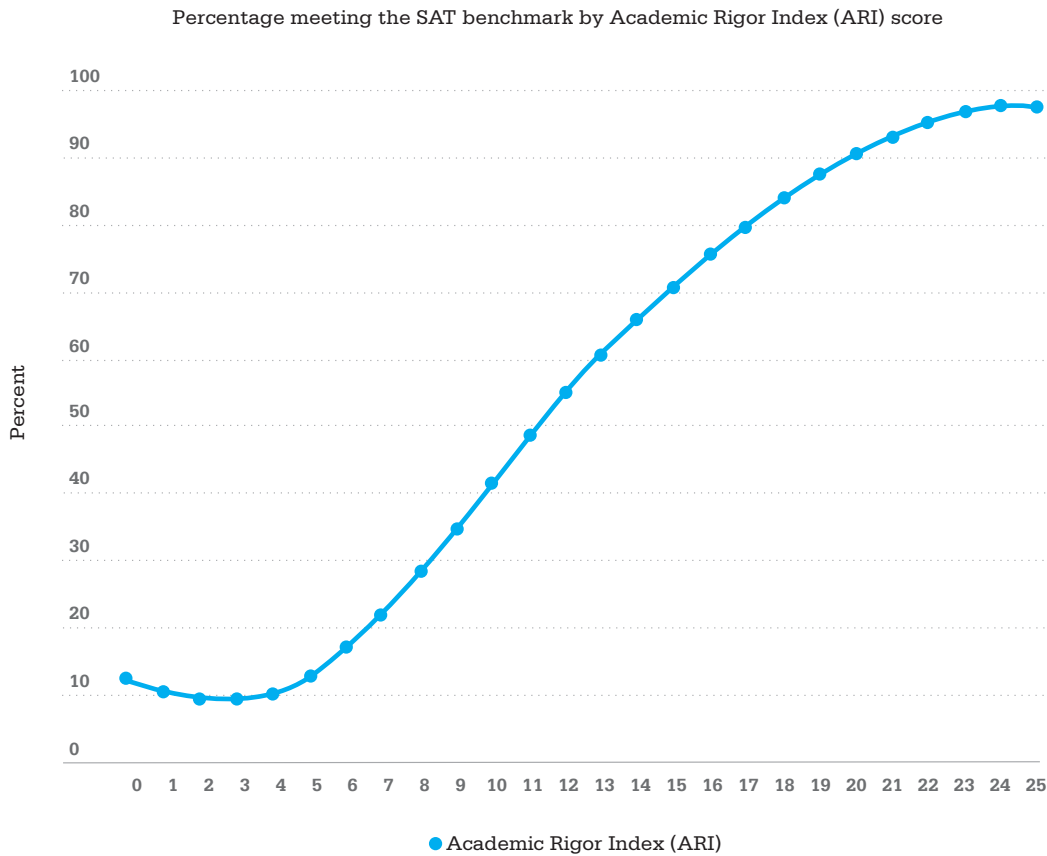
College readiness also varied considerably with academic course-taking behavior. One measure of college preparation is a core curriculum which consists of four years of English and three years each of mathematics, science, and social science/history. Half of the students who took a core curriculum met the benchmark, compared to 29 percent of those who did not. The academic rigor index (ARI), which measures the challenge associated with high school course work, is highly related to the core curriculum because students who have taken a core curriculum have a mean ARI score of 11.5 and a median ARI score of 11.0, compared to a mean ARI score of 7.7 and a median score of 7.0 for students who did not take a core curriculum. In general, students with a more rigorous high school curriculum were much more likely to meet the SAT College Readiness benchmark. Approximately 13 percent of students with an ARI of 5 or lower met the benchmark. Each successive increase on the ARI is associated with an increased percent of students considered to be ready for college. For example, 29 percent of students with an ARI between 6 and 10 met the benchmark, compared to 60 percent with an ARI between 11 and 15, 83 percent with an ARI between 16 and 20, and 95 percent of those with an ARI of 21 or more. Thus, as expected, benchmark attainment appears to be highly related to other measures of academic performance during high school. Figure 2 and Appendix B contain the percentage meeting the benchmark by each ARI score point.

Table 5

Percentage of the 2010 Cohort (Sample 3) that Met the Benchmark by Academic Variables

Academic Variables	Number	Met Benchmark (%)	
HSGPA	A+	80,417	84
	A	259,152	71
	A-	265,289	57
	B+	264,280	38
	B	237,944	27
	B-	125,475	18
	C+	83,253	12
	C	46,127	10
C- or Lower	19,477	9	
Curriculum	Non-Core	271,050	29
	Core	959,396	50
Academic Rigor (ARI)	0–5	223,451	13.2
	6–10	411,250	29.2
	11–15	268,145	60.0
	16–20	184,478	82.9
	21–25	63,359	95.1

Figure 2



While the SAT composite benchmark is used to provide information on general college readiness, the individual section benchmarks provide feedback on student performance on the critical reading, mathematics, and writing sections. Figure 3 and Figure 4 provide the distribution of AP English Language and AP English Literature Exam scores for students who met or exceeded the SAT College Readiness benchmarks in critical reading and writing, and for those who have not. The difficulty of AP courses is on par with that of introductory college-level courses, and so evaluating the relationship between the SAT benchmarks and AP performance is one way to validate the SAT college readiness benchmarks in terms of college outcomes. An AP Exam score of 3 indicates that a student has achieved the equivalent of a “C” (a passing grade) in a college-level course, the minimum score for which most colleges award credit.

Figure 3 shows that 9 percent of students who did not meet the SAT-CR benchmark obtained a 3 or higher in AP English Language, compared to 78 percent of students who met the benchmark. The results are similar for the AP Exam in English Literature, as 6 percent of those not meeting the College Readiness benchmark on SAT-CR obtained a 3 or higher compared to 74 percent of those who met the benchmark.

Figure 3

Student performance on English Advanced Placement® (AP®) Exams by college readiness status on the critical reading section

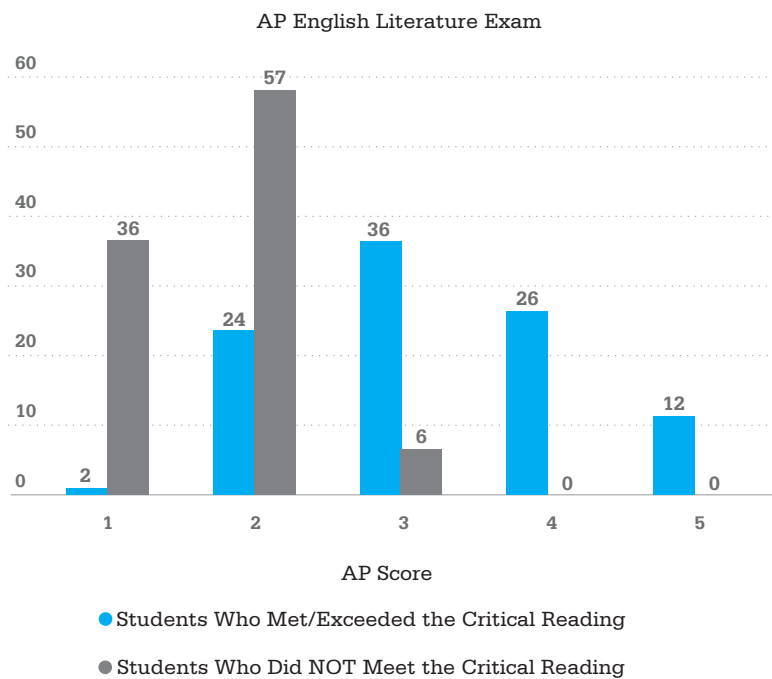
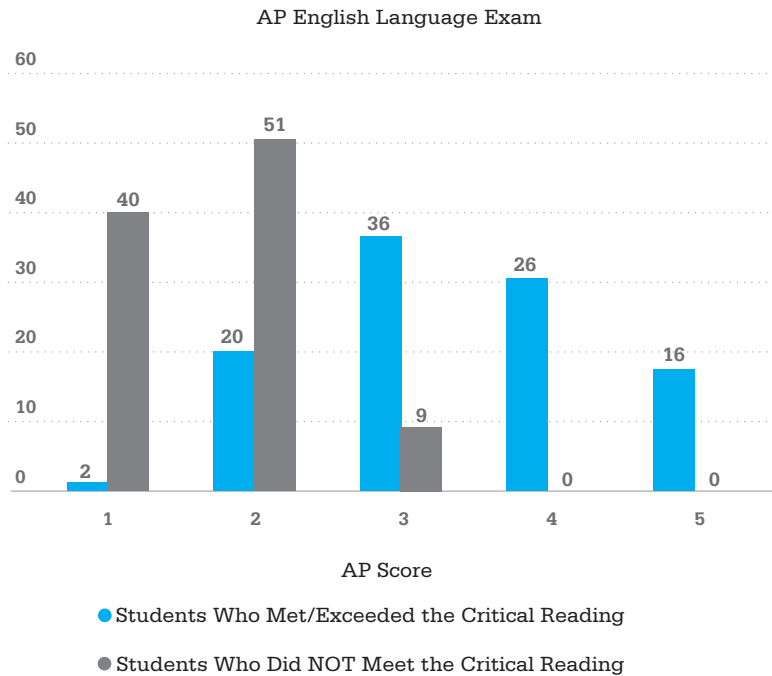


Figure 4 shows the relationship between the College Readiness benchmark on the SAT writing section and performance on the AP English Language and English Literature Exams. Only 14 percent and 12 percent of students who do not meet the SAT-W benchmark scored a 3 or higher on the AP English Language and English Literature Examinations, respectively, compared to 79 percent and 75 percent of those students who met the benchmark.

Figure 4

Student performance on English Advanced Placement (AP) Exams by college readiness status on the writing section

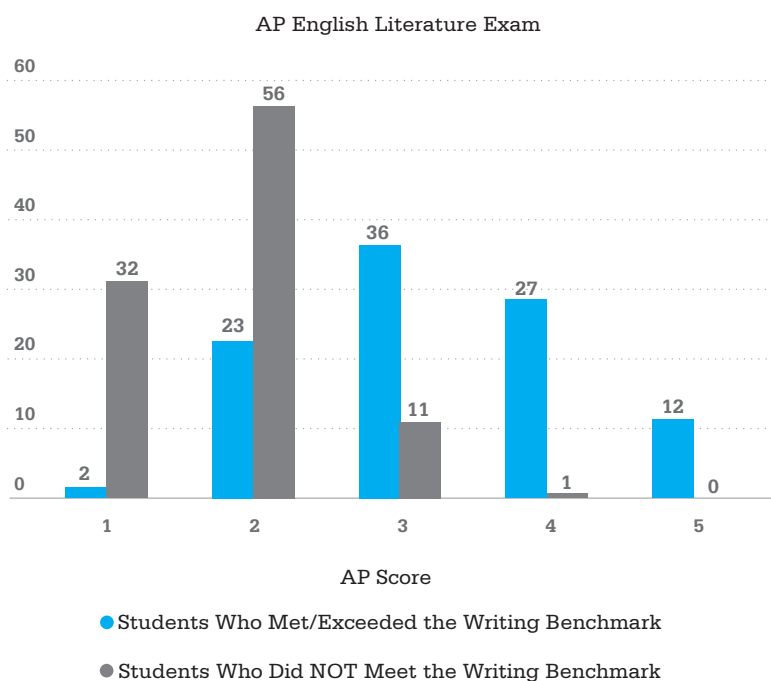
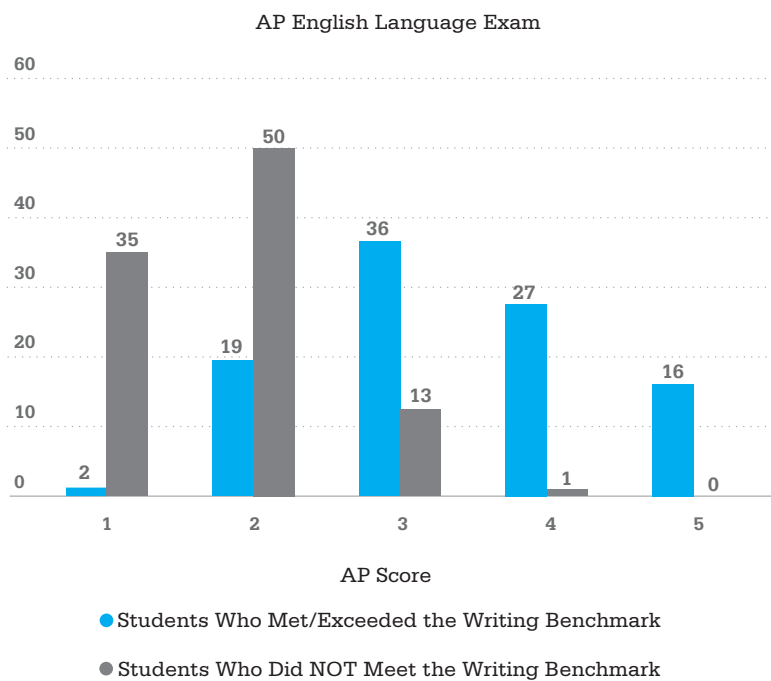
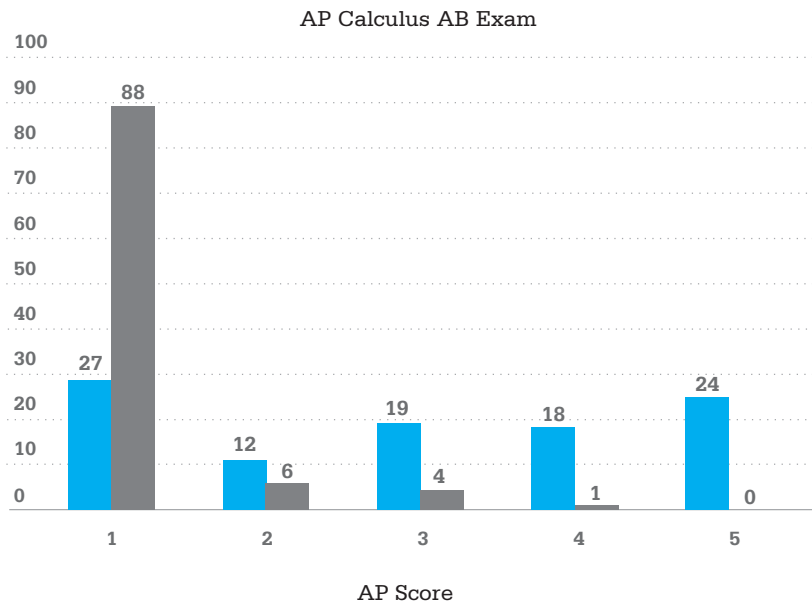


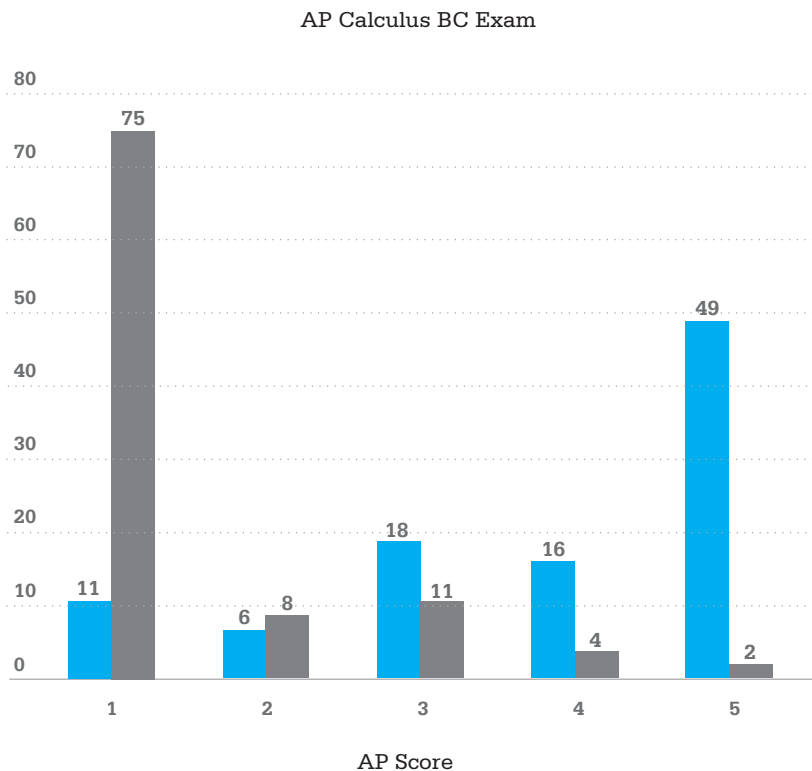
Figure 5 shows the relationship between achieving the college readiness benchmark on the SAT mathematics section and performance on relevant AP Exams. Between 5 percent and 17 percent of those not meeting the SAT-M benchmark scored a 3 or higher on an AP Calculus or AP Statistics Exam, compared to between 61 percent and 83 percent for those who met the SAT-M benchmark. Thus, it appears that students meeting the section score benchmarks are much more likely to successfully complete college-level course work than students not meeting the benchmark.

Figure 5

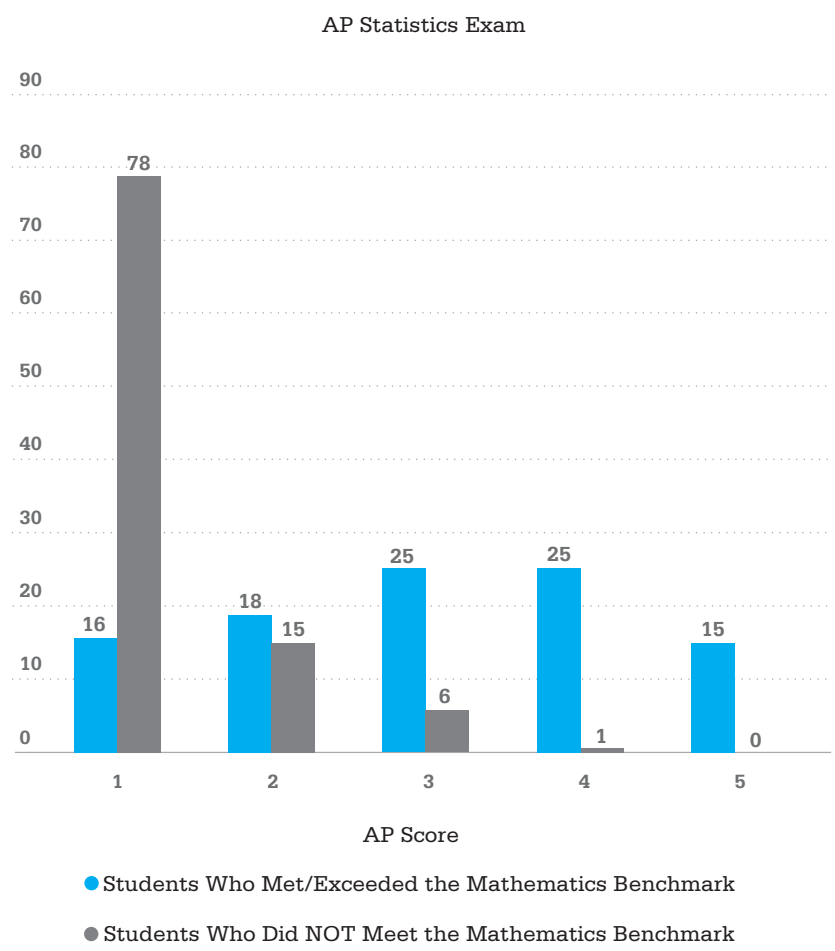
Student performance on Math Advanced Placement (AP) Exams by college readiness status on the math section



- Students Who Met/Exceeded the Mathematics Benchmark
- Students Who Did NOT Meet the Mathematics Benchmark



- Students Who Met/Exceeded the Mathematics Benchmark
- Students Who Did NOT Meet the Mathematics Benchmark



SAT Benchmarks and Student Demographics

Table 6 reports the percent of students who met the benchmark by demographic characteristics. Overall, 43 percent of 2010 graduates who took the SAT met the SAT College Readiness benchmark. The results were similar across gender, with a slightly larger percentage of male students meeting the SAT benchmark than female students.

The differences in benchmark achievement among race and ethnicity subgroups are consistent with prior research showing similar gaps on precollege measures, including HSGPA, NAEP, SAT, and ACT (Kobrin, Sathy, & Shaw, 2006; Camara & Schmidt, 1999), college outcomes including remediation rates (Wirt et al., 2004) and degree completion (see, for example, Attewell et al., 2010). The percent of Asian and white students who met the SAT benchmark was more than twice that of Hispanic students and more than three times that of African American students. Students reporting English as their best language were more likely to meet the SAT benchmark than students whose best language was either another language and English or solely another language. College readiness also differed by parental education levels with fewer than one in six students of parents without a high school diploma meeting the SAT benchmark, compared to half of those who have one or more parents with a bachelor’s degree and over two-thirds of students with one or more parents with a graduate degree or higher.

Table 6

Percentage of the 2010 Cohort (Sample 3) that Obtained the Benchmark: By Demographic Variables

Demographic Variables		Number	Obtained Benchmark (%)
Overall		1,457,489	43
Gender	Female	782,804	41
	Male	674,685	46
Race/Ethnicity	American Indian	8,295	35
	African American	191,566	15
	Asian American	133,531	56
	Hispanic	215,181	24
	White	817,915	53
	Other	39,769	42
	No Response	51,412	44
Best Language	English Only	1,181,661	46
	English & Another	214,827	31
	Another	31,337	26
	No Response	29,664	34
Parental Education	Less than High School	71,653	15
	High School	416,206	27
	Associate Degree	117,701	33
	Bachelor's Degree	393,374	52
	Graduate Degree	327,231	68

Discussion

In order for the United States to continue to prosper, it is imperative that all students have the access and opportunity to attend college and earn a college degree. Educators, families, communities, and policymakers all have the responsibility to ensure that all students, including those from low-income backgrounds, graduate from high school ready for college success (College Board, 2007). One barrier to achieving this goal is inadequate high school preparation that leaves many students unprepared for college-level work. This may result in students either failing to attend college or remediation for those who do attend. This has become particularly important in a globally competitive setting in which the United States finds its 15 and 16-year-old students failing to keep pace with the gains of international students (OECD Programme for International Student Assessment, 2009). The challenge for the United States will likely increase in the coming years as a larger percent of traditionally underserved students enter the school system.

Given these trends, the College Board has embarked on a series of initiatives centered on increasing the number of students who are ready and able to attend college. One important part of this challenge is to develop empirical measures to identify if students are ready for college. The goal is to provide states, districts, administrators, teachers, parents, and students with information regarding students' preparedness for and ability to succeed in college. The SAT benchmarks described in this report were created to establish a threshold for students that, if met, would ensure a reasonable probability of college success and eventual completion.

The SAT benchmarks are designed to measure the college readiness of high school students, using the SAT, a college entrance examination taken by nearly 1.45 million students in all 50 United States and the District of Columbia. The SAT benchmark determined in this study was 1550 for the composite. Individual benchmark scores were also calculated for the critical reading, mathematics, and writing sections to provide indicators of student proficiency in each of these subjects, resulting in a benchmark score of 500 on each section. In the 2010 cohort of college-bound students, 43 percent met the SAT college readiness benchmark. This report provided evidence to validate the use of the benchmark for assessing college readiness by showing the association of the benchmarks with other measures of student performance in high school and college. Specifically, students meeting the benchmark of 1550 on the SAT:

- are more likely to enroll in college and are more likely to enroll in a four-year as opposed to a two-year college;
- are more likely to be retained for their second and third year of college;
- earn a higher FYGPA, on average, compared to those not meeting the benchmark; and
- are more likely to have earned higher grades in high school, are more likely to have taken a core curriculum, and are more likely to have taken rigorous courses in high school.

In addition, students meeting the benchmark on the SAT critical reading and writing sections are more likely to score a 3 or higher on an AP Exam in English; and students meeting the benchmark on the SAT mathematics section are more likely to score a 3 or higher on an AP exam in mathematics.

The SAT benchmarks have several advantages, including the ability to easily measure the college readiness of students; the potential for enhanced aggregate reporting to assist schools and districts; and the ability to provide early indicators of college readiness. In addition, the SAT benchmark could prove useful in assessing changes in student preparedness over time. Additionally, academic behaviors associated with benchmark attainment could be identified and encouraged on a wide scale. For example, a discovery that a particular course sequence in mathematics is positively associated with achieving the SAT-M benchmark could suggest benefits in introducing the sequence to more students.

PSAT/NMSQT benchmarks in the 10th and 11th grade have also been created by linking PSAT/NMSQT scores to SAT benchmark scores. The goal is to provide information as to whether younger high school students are on track to be ready for college (see Proctor, Wyatt, & Wiley, 2010). By extending indications of college readiness to 10th and 11th grade, more time would be available to assist students in academic need. Because 1.5 million students take the PSAT/NMSQT as juniors and 1.5 million take the exam as sophomores, this tool has the potential to assist a large number of students in becoming ready for college. A 10th-grade PSAT/NMSQT benchmark score of 145 (60–240 scale) and an 11th-grade PSAT/NMSQT benchmark score of 152 indicate that a student is on track to meeting the SAT benchmark. For students testing in the 2009-10 academic year, only 27 percent of sophomores and 38 percent of juniors were on track to meet the SAT benchmark (Proctor et al., 2010). These results provide the opportunity to give increased academic support to students not currently on the path to. Additionally, the College Board has developed ReditStep as an assessment for students to be administered in the eighth grade. A link is currently being developed between ReditStep and the PSAT/NMSQT benchmark scores, allowing early indicators to be calculated for eighth-grade students. The development of an eighth-grade indicator could provide additional benefits by enabling early intervention for students who may require extra support.

Limitations

One limitation of the proposed SAT benchmark is that students intending to attend college are more likely to take the SAT and generally have stronger academic credentials than those not taking the exam. This effect is likely to be magnified in states where a low percentage of the student population take the exam, since SAT takers in those states are likely to be high achievers and are less representative of the total student population. However, in schools, districts, and states where a high percentage of students take the SAT, the college readiness benchmark should be an accurate indicator of group preparedness.

While the SAT college readiness benchmark can be an accurate indicator of the academic preparedness of students, it does not consider other noncognitive factors such as motivation and persistence, which are also linked to success in college. Due to the omission of noncognitive factors and other measures of high school achievement, the benchmark is designed to evaluate the aggregate readiness of a group of students rather than the individual student. When evaluating the individual student, the SAT should not be the only piece of information considered in making decisions on readiness for college. Other measures of academic performance and other factors should be considered in conjunction with the SAT. These factors may traditionally include HSGPA, the difficulty of high school course work, letters of recommendation, a personal statement, and extracurricular activities.

Future research should examine the stability of the benchmarks over time, across different student subgroups, different types of postsecondary institutions, and across college majors. Finally, other outcome measures, in particular college graduation as well as certain nonacademic measures of college success, should be examined in relationship to the benchmarks.

References

- ACT. (2004). *Crisis at the core: Preparing all students for college and work*. Iowa: ACT, Inc.
- Allen, D. (1999). Desire to finish college: An empirical link between motivation and persistence. *Research in Higher Education*, 40, 461–485.
- Attewell, P., Heil, S., & Reisel, L. (2010). Competing explanations of undergraduate completion. *American Educational Research Journal*, 47, 1–22.
- Baum, S., & Ma, J. (2007). *Education pays: The benefits of higher education for individuals and society*. The College Board: New York.
- Baum, S., & Payea, K. (2004). *Education pays*. New York, NY: The College Board. Retrieved, April 9, 2009, from http://www.collegeboard.com/prod_downloads/press/cost04/EducationPays2004.pdf.
- Beaton, A.E., & Allen, N. L. (1992). Interpreting scales through scale anchoring. *Journal of Educational Statistics*, 17 (2), 191–204.
- Camara, W. J. (2011). *Empirical benchmarks in a judgmental standard setting process*. Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans, LA.
- Camara, W.J., & Schmidt, A.E. (1999). Group differences in standardized testing and social stratification, (College Board Research Report No. 1999-5). New York: The College Board.
- Chen, X., & Carroll, C.D. (2005). *First Generation Students in Postsecondary Education* (NCES 2005-171). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office. U.S.
- College Board (2007). *The CollegeKeys compact: Getting ready, getting in, and getting through college: Expanding options for low-income students*. New York: The College Board.
- Corwin, Z. B., & Tierney, W. G. (2007). *Getting there — and beyond: Building a culture of college-going in high schools*. Los Angeles, CA: University of Southern California Center for Higher Education Policy Analysis.
- Goldberg, J., & Smith, J. (2008). The effects of education on labor market outcomes. In Helen F. Ladd and Edward B. Fiske (eds.) *Handbook of research in education finance and policy* (688-708). New York, NY: Routledge.

- Kobrin, J. (2007). Determining SAT benchmarks for college readiness. College Board Research Note RN-30. New York: The College Board.
- Kobrin, J. L., Patterson, B.F., Shaw, E. J., Mattern, K.D., & Barbuti, S. M. (2008). Validity of the SAT for Predicting First-Year College Grade Point Average. (College Board Research Report No. 2008-5). New York: The College Board.
- Kobrin, J.L., Patterson, B.F., Wiley, A., & Mattern, K.D. (under review). A Standard Setting Study to Set Benchmarks on the SAT. New York: The College Board.
- Kobrin, J.L., Sathy, V. & Shaw, E.J. (2006). A historical view of subgroup differences on the SAT Reasoning Test. (College Board Research Report No. 2006-5). New York: The College Board.
- Kuncel, N.R., Credé, M., & Thomas, L.L. (2005). The validity of self-reported grade point average, class ranks, and test scores: A meta-analysis and review of the literature. *Review of Educational Research*, 75, 63–82.
- Maxey, E. J., & Ormsby, V. J. (1971). The accuracy of self-report information collected on the ACT test battery: High school grades and items of nonacademic achievement. (ACT Research Report No. 45). Iowa City, IA: ACT.
- McNeil, M. (2008). Benchmarks momentum on increase: Governors' group, state chiefs eyeing international yardsticks. *Education Week* 27(27), 12–13
- Morgan, D.L., & Michaelides, M.P. (2005). Setting cut scores for college placement. (College Board Research Report No. 2005-9). New York: The College Board.
- Murtaugh, P. A., Burns, L. D., & Schuster, J. (1999). Predicting the retention of university students. *Research in Higher Education*, 40, 355–371.
- National Center for Higher Education Management Systems (2009). NCHEMS Information Center. Retrieved from: <http://www.higheredinfo.org/dbrowser/?level=nation&mode=map&state=0&submeasure=239>
- OECD Programme for International Student Assessment (2009). PISA 2009 Results. Retrieved from: http://www.oecd.org/document/61/0,3746,en_32252351_46584327_46567613_1_1_1_1,00.html
- Organization for Economic Co-Operation and Development (2010). Education at a Glance. Retrieved from <http://www.oecd.org/dataoecd/45/39/45926093.pdf>.

- Organization for Economic Co-Operation and Development (2008). Education at a Glance. Retrieved from: <http://www.oecd.org/dataoecd/23/46/41284038.pdf>
- Presley, J. B., & Gong, Y. (2005). The Demographics and Academics of College Readiness in Illinois, IERC 2005-3, (Edwardsville, IL: Illinois Education Research Council, 2005). Retrieved from www.schooldata4all.org.
- Proctor, T., Wyatt, J., & Wiley, A. (2010). The development of early indicators of college readiness using for the PSAT/NMSQT and ReadStep examination programs (College Board Research Report 2010-4). New York: The College Board.
- Schiel, J., & Noble, J. (1991). Accuracy of self-reported course work and grade information of high school sophomores. (ACT Research Report No. 91-6). Iowa City, IA: ACT.
- Shaw, E. J., & Mattern, K. D. (2009). Examining the accuracy of self-reported high school grade point average. (College Board Research Report No. 2009-5). New York: The College Board.
- Tinto, V. (1993). Leaving college: Rethinking the cause of student attrition (2nd ed.). Chicago, IL: University of Chicago Press.
- U.S. Department of Education. (2008). Leading education indicators: Trends in education. Retrieved from <http://www2.ed.gov/nclb/accountability/results/trends/index.html>.
- Wirt, J., Choy, S., Rooney, P., Provasnik, S., Sen, A., & Tobin, R. (2004). The Condition of Education 2004 (NCES 2004-077). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office. U.S.
- Wyatt, J., Wiley, A., Camara, W. J., & Proestler, N. (in press). The development of an index of academic rigor for college readiness.
- Zwick, R., Senturk, D., & Wang, J. (2001). An investigation of alternative methods for item mappings in the National Assessment of Educational Progress. *Educational Measurement: Issues and Practice*, 20(2): 15–25.

Appendix A

Comparison of Sample 1 to Population of Four Year Institutions by Key Variables: Variable, Class, Population, Sample, Sample N				
Variable	Class	Population	Sample	Sample N
Region of the U.S.	Midwest	16%	16%	18
	Mid-Atlantic	18%	21%	23
	New England	13%	18%	20
	South	25%	14%	15
	Southwest	10%	13%	14
	West	18%	18%	20
Selectivity	Admits under 50%	20%	19%	21
	Admits 50 to 75%	44%	57%	63
	Admits over 75%	36%	24%	26
Size	Small	18%	22%	24
	Medium to large	43%	37%	41
	Large	20%	17%	19
	Very large	19%	24%	26
Control	Public	57%	46%	51
	Private	43%	54%	59

Appendix B

Percent of the 2010 Cohort (Sample 3) Meeting the SAT Benchmark by Academic Rigor Index (ARI)

ARI	N	Percent Meeting Benchmark
0	3,439	12
1	8,102	10
2	19,382	9
3	39,604	10
4	65,463	11
5	87,461	13
6	93,649	17
7	92,579	22
8	84,054	29
9	74,231	35
10	66,737	42
11	61,469	49
12	56,719	55
13	53,475	61
14	49,644	66
15	46,838	71
16	43,780	76
17	41,010	80
18	37,310	84
19	33,459	88
20	28,919	91
21	23,505	93
22	17,773	95
23	11,924	97
24	7,205	98
25	3,152	98

Appendix C

Student Performance on English Advanced Placement (AP) Exams for Those Meeting and Not Meeting the SAT-CR Benchmark

AP Exam		Percent Distribution by AP Score					Number
		1	2	3	4	5	
English Language	Did Not Meet Benchmark	40	51	9	0	0	49,108
	Met Benchmark	2	20	36	26	16	198,713
English Literature	Did Not Meet Benchmark	36	57	6	0	0	47,310
	Met Benchmark	2	24	36	26	12	207,563

Appendix D

Student Performance on English Advanced Placement (AP) Exams for Those Meeting and Not Meeting the SAT-W Benchmark

AP Exam		Percent Distribution by AP Score					Number
		1	2	3	4	5	
English Language	Did Not Meet Benchmark	35	50	13	1	0	57,003
	Met Benchmark	2	19	36	27	16	190,818
English Literature	Did Not Meet Benchmark	32	56	11	1	0	55,545
	Met Benchmark	2	23	36	27	12	199,328

Appendix E

Student Performance on Mathematics Advanced Placement (AP) Exams for Those Meeting and Not Meeting the SAT-M Benchmark

AP Exam		Percent Distribution by AP Score					Number
		1	2	3	4	5	
Calculus AB	Did Not Meet Benchmark	88	6	4	1	0	10,745
	Met Benchmark	27	12	19	18	24	167,219
Calculus BC	Did Not Meet Benchmark	75	8	11	4	2	513
	Met Benchmark	11	6	18	16	49	61,196
Statistics	Did Not Meet Benchmark	78	15	6	1	0	9,635
	Met Benchmark	16	18	25	25	15	88,555

