SAT

This article is about the college admission test in the United States. For the exams in England colloquially known as SATs, see National Curriculum assessment.

For other uses, see SAT (disambiguation).

SAT Test

Type	Paper-based standardized test
Developer / administrator	College Board, Educational Testing Service.
Knowledge / skills tested	Writing, critical reading, mathematics.
Purpose	Admission to undergraduate programs of universities or colleges.
Year started	1926
Duration	3 to 4 hours
Score / grade range	200–800 (in 10-point increments) on each of two sections (total 400–1600). Essay scored on scale of 0–24, in 1-point increments.

Offered	Seven times annually
Countries / regions	Worldwide
Languages	English
Annual number of test takers	Over 1.69 million high school graduates in the class of 2015
Prerequisites / eligibility criteria	No official prerequisite. Intended for high school students. Fluency in English assumed.
Fee	US$52.50 to US$101.50, depending on country.
Scores / grades used by	Most universities and colleges offering undergraduate programs in the U.S.
Website	sat.collegeboard.org

The **SAT** is a standardized test widely used for college admissions in the United States. Introduced in 1926, its name and scoring have changed several times; originally called the **Scholastic Aptitude Test**, it was later called the **Scholastic Assessment Test**, then the **SAT I: Reasoning Test**, then the **SAT Reasoning Test**, and now, simply the **SAT**.

The SAT is owned and published by the College Board, a private, not-for-profit corporation in the United States. It is developed and administered on behalf of the College Board by the Educational Testing Service. The test is intended to assess students' readiness for college. The SAT was originally designed not to be aligned with high school curricula, but several adjustments have been made for the version of the SAT introduced in 2016, and College Board president, David Coleman, has said that he also wanted to make the test reflect more closely what students learned in high school.

On March 5, 2014, the College Board announced that a redesigned version of the SAT would be administered for the first time in 2016. The current SAT, introduced in 2016, takes three hours to finish, plus 50 minutes for the SAT with essay, and as of 2017 costs US$45 (US$57 with the optional essay), excluding late fees, with additional processing fees if the SAT is taken outside the United States. Scores on the SAT range from 400 to 1600, combining test results from two 800-point sections: mathematics, and critical reading and writing. Taking the SAT, or its competitor, the ACT, is required for freshman entry to many, but not all, universities in the United States. Starting with the 2015-16 school year, the College Board also announced it would team up with Khan Academy, the free, online education site to provide SAT prep, free of charge.

Function

Education in the United States

- By state + insular areas
- By subject area
- History of
- Issues: *Finance* – *Law* – *Literacy* – *Reform*
- Levels: *Primary* – *Secondary* – *Higher*
- Organizations

Education portal

United States portal

The SAT is typically taken by high school juniors and seniors. The College Board states that the SAT measures literacy and writing skills that are needed for academic success in college. They state that the SAT assesses how well the test takers analyze and solve problems—skills they learned in school that they will need in college. However, the test is administered under a tight time limit (speeded) to help produce a range of scores.

The College Board also states that use of the SAT in combination with high school grade point average (GPA) provides a better indicator of success in college than high school grades alone, as measured by college freshman GPA. Various studies conducted over the lifetime of the SAT show a statistically significant increase in correlation of high school grades and college freshman grades when the SAT is factored in. A large independent validity study on the SAT's ability to predict college freshman GPA was performed by the University of California. The results of this study found how well various predictor variables could explain the variance in college freshman GPA. It found that independently high school GPA could explain 15.4% of the variance in college freshman GPA, SAT I (the SAT Math and Verbal sections) could explain 13.3% of the variance in college freshman GPA, and SAT II (also known as the SAT subject tests; in the UC's case specifically Writing, Mathematics IC or IIC, plus a third subject test of the student's choice) could explain 16% of the variance in college freshman GPA. When high school GPA and the SAT I were combined, they explained 20.8% of the variance in college freshman GPA. When high school GPA and the SAT II were combined, they explained 22.2% of the variance in college freshman GPA. When SAT I was added to the combination of high school GPA and SAT II, it added a .1 percentage point increase in explaining the variance in college freshman GPA for a total of 22.3%.

There are substantial differences in funding, curricula, grading, and difficulty among U.S. secondary schools due to U.S. federalism, local control, and the prevalence of private, distance, and home schooled students. SAT (and ACT) scores are intended to supplement the secondary school record and help admission officers put local data—such as course work, grades, and class rank—in a national perspective. However, independent research has shown that high school GPA is better than the SAT at predicting college grades regardless of high school type or quality.

This map of the United States shows the states in which (blue color) more seniors in the class of 2016 took the SAT than the ACT, and the states in which (red color) more seniors took the ACT than the SAT.

Historically, the SAT was more widely used by students living in coastal states and the ACT was more widely used by students in the Midwest and South; in recent years, however, an increasing number of students on the East and West coasts have been taking the ACT. Since 2007, all four-year colleges and universities in the United States that require a test as part of an application for admission will accept either the SAT or ACT, and hundreds of colleges and universities do not require any standardized test scores at all for admission.

Structure

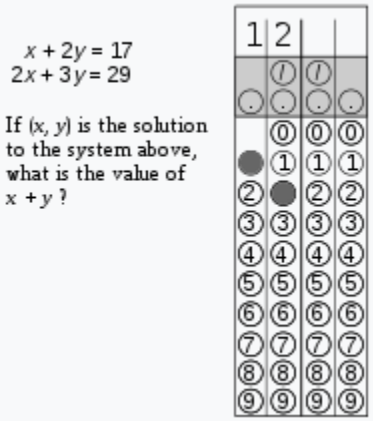
SAT consists of three major sections: Critical Reading, Mathematics, and Writing. Each section receives a score on the scale of 200–800. All scores are multiples of 10. Total scores are calculated by adding up scores of the three sections. Each major section is divided into three parts. There are 10 sub-sections, including an additional 25-minute experimental or "equating" section that may be in any of the three major sections. The experimental section is used to normalize questions for future administrations of the SAT and does not count toward the final score. The test contains 3 hours and 45 minutes of actual timed sections, most administrations (after accounting for orientation, distribution of materials, completion of biographical sections, and fifteen minutes of timed breaks) run for about four and a half hours. The questions range from easy, medium, and hard depending on the scoring from the experimental sections. Easier questions typically appear closer to the beginning of the section while harder questions are toward the end in certain sections. This is not true for every section (the Critical Reading section is in chronological order) but it is the rule of thumb mainly for math, grammar, and the 19 sentence-completions in the reading sections.

Critical Reading

The Critical Reading section of the SAT is made up of one section with 52 questions, which should be completed in 65 minutes, with varying types of questions, including sentence completions and questions about short and long reading passages. Critical Reading sections normally begin with 5 to 8 sentence completion questions; the remainder of the questions are focused on the reading passages. Sentence completions generally test the student's vocabulary and understanding of sentence structure and organization by requiring the student to select one or two words that best complete a given sentence. The bulk of the Critical Reading section is made up of questions regarding reading passages, in which students read short excerpts on social sciences, humanities, physical sciences, or personal narratives and answer questions based on the passage. Certain sections contain passages asking the student to compare two related passages; generally, these consist of shorter reading passages. The number of questions about each passage is proportional to the length of the passage. Unlike in the Mathematics section, where questions go in the order of difficulty, questions in the Critical Reading section go in the order of the passage. Overall, question

sets near the beginning of the section are easier, and question sets near the end of the section are harder.

Mathematics



An example of an SAT "grid-in" math question and the correctly gridded answer.

The mathematics portion of the SAT is divided into two sections: Math Test – Calculator and Math Test – No Calculator. In total, the SAT math test is 80 minutes long and includes 58 questions: 45 multiple choice questions and 13 grid-in questions. The multiple choice questions have four possible answers; the grid-in questions are free response and require the test taker to provide an answer.

- The Math Test – No Calculator section has 20 questions (15 multiple choice and 5 grid-in) and lasts 25 minutes.
- The Math Test – Calculator section has 38 questions (30 multiple choice and 8 grid-in) and lasts 55 minutes.

Several scores are provided to the test taker for the math test. A sub score (on a scale of 1 to 15) is reported for each of three categories of math content: "Heart of Algebra" (linear equations, systems of linear equations, and linear functions), "Problem Solving and Data Analysis" (statistics, modeling, and problem-solving skills), and "Passport to Advanced Math" (non-linear expressions, radicals, exponentials and other topics that form the basis of more advanced math). A test score for the math test is reported on a scale of 10 to 40, and a section score (equal to the test score multiplied by 20) is reported on a scale of 200 to 800.

Calculator use

All scientific and most graphing calculators, including Computer Algebra System (CAS) calculators, are permitted on the SAT Math – Calculator section only. All four-function calculators are allowed as well; however, these devices are not recommended. All mobile phone and smartphone calculators, calculators with typewriter-like (QWERTY) keyboards, laptops and other portable computers, and calculators capable of accessing the Internet are not permitted.

Research was conducted by the College Board to study the effect of calculator use on SAT I: Reasoning Test math scores. The study found that performance on the math section was associated with the extent of calculator use: those using calculators on about one third to one half of the items averaged higher scores than those using calculators more or less frequently. However, the effect was "more likely to have been the result of able students using calculators

differently than less able students rather than calculator use per sec. There is some evidence to suggest that the frequent use of a calculator in school outside of the testing situation has a positive effect on test performance compared to those who do not use calculators in school.

Writing

The writing part portion of the SAT, based on but not directly comparable to the old SAT II subject test in writing (which in turn was developed from the old Test of Standard Written English (TSWE), includes multiple choice questions and a brief essay. The essay sub score contributes about 28% to the total writing score, with the multiple choice questions contributing 70%. This section was implemented in March 2005 following complaints from colleges about the lack of uniform examples of a student's writing ability and critical thinking.

The multiple choice questions include error-identification questions, sentence-improvement questions, and paragraph-improvement questions. Error-identification and sentence-improvement questions test the student's knowledge of grammar, presenting an awkward or grammatically incorrect sentence; in the error identification section, the student must locate the word producing the source of the error or indicate that the sentence has no error, while the sentence improvement section requires the student to select an acceptable fix to the awkward sentence. The paragraph improvement questions test the student's understanding of logical organization of ideas, presenting a poorly written student essay and asking a series of questions as to what changes might be made to best improve it.

The essay section, which is always administered as the first section of the test, is 25 minutes long. All essays must be in response to a given prompt. The prompts are broad and often philosophical and are designed to be accessible to students regardless of their educational and social backgrounds. For instance, test takers may be asked to expand on such ideas as their opinion on the value of work in human life or whether technological change also carries negative consequences to those who benefit from it. No particular essay structure is required, and the College Board accepts examples "taken from [the student's] reading, studies, experience, or observations." Two trained readers assign each essay a score between 1 and 6, where a score of 0 is reserved for essays that are blank, off-topic, non-English, not written with a Number 2 pencil, or considered illegible after several attempts at reading. The scores are summed to produce a final score from 2 to 12 (or 0). If the two readers' scores differ by more than one point, then a senior third reader decides. The average time each reader/grader spends on each essay is less than 3 minutes.

In March 2005, Les Perelman analyzed 15 scored sample essays contained in the College Board's *Score Write* book along with 30 other training samples and found that in over 90% of cases, the essay's score could be predicted from simply counting the number of words in the essay. Two years later, Perelman trained high school seniors to write essays that made little sense but contained infrequently used words such as "plethora" and "myriad". All of the students received scores of "10" or better, which placed the essays in the 92nd percentile or higher.

Style of questions

Most of the questions on the SAT, except for the essay and the grid-in math responses, are multiple choice; all multiple-choice questions have four answer choices, one of which is correct. The questions of each section of the same type are generally ordered by difficulty. However, an important exception exists: Questions that follow the long and short reading passages are organized chronologically, rather than by difficulty. Ten of the questions in one of the math sub-sections are not multiple choice. They instead require the test taker to bubble in a number in a four-column grid.

The questions are weighted equally. For each correct answer, one raw point is added. No points are deducted for incorrect answers. The final score is derived from the raw score; the precise conversion chart varies between test administrations.

Section	Average Score[1]	Time (Minutes)	Content
Writing	484	35	Grammar, usage, and diction.
Mathematics	511	60	Number and operations; algebra and functions; geometry; statistics, probability, and data analysis
Critical Reading	495	65	Vocabulary, Critical reading, and sentence-level reading

Logistics

The SAT is offered four times a year in the United States; in October, December, March (or April, alternating) and May. The test is typically offered on the first Saturday of the month for the October, December, March (or April, alternating) and May administrations. In other countries, the SAT is offered on the same dates as in the United States except for the first spring test date (i.e., March or April), which is not offered. The test was taken by 1,698,521 high school graduates in the class of 2015.

Candidates wishing to take the test may register online at the College Board's website, by mail, or by telephone, at least three weeks before the test date.

The SAT costs $45 ($57 with the optional essay), plus additional fees if testing outside the United States) as of 2017. The College Board makes fee waivers available for low income students. Additional fees apply for late registration, standby testing, registration changes, scores by telephone, and extra score reports (beyond the four provided for free).

Candidates whose religious beliefs prevent them from taking the test on a Saturday may request to take the test on the following day, except for the October test date in which the Sunday test date is eight days after the main test offering. Such requests must be made at the time of registration and are subject to denial.

Students with verifiable disabilities, including physical and learning disabilities, are eligible to take the SAT with accommodations. The standard time increase for students requiring additional time due to learning disabilities or physical handicaps is time + 50%; time + 100% is also offered.

Raw scores, scaled scores, and percentiles

Students receive their online score reports approximately three weeks after test administration (six weeks for mailed, paper scores), with each section graded on a scale of 200–800 and two sub scores for the writing section: the essay score and the multiple choice sub score. In addition to their score, students receive their percentile (the percentage of other test takers with lower scores). The raw score, or the number of points gained from correct answers and lost from incorrect answers is also included. Students may also receive, for an additional fee, the Question and Answer Service, which provides the student's answer, the correct answer to each question, and online resources explaining each question.

The corresponding percentile of each scaled score varies from test to test—for example, in 2003, a scaled score of 800 in both sections of the SAT Reasoning Test corresponded to a percentile of 99.9, while a scaled score of 800 in the SAT Physics Test corresponded to the 94th percentile. The differences in what scores mean with regard to percentiles are due to the content of the exam and the caliber of students choosing to take each exam. Subject Tests are subject to intensive study (often in the form of an AP, which is relatively more difficult), and only those who know they will perform well tend to take these tests, creating a skewed distribution of scores.

The percentiles that various SAT scores for college-bound seniors correspond to be summarized in the following chart;

Percentile	Score, 1600 Scale (official, 2006)	Score, 2400 Scale (official, 2006)
99.93/99.98*	1600	2400
99+ **	≥1540	≥2280
99	≥1480	≥2200
98	≥1450	≥2140
97	≥1420	≥2100
93	≥1340	≥1990
88	≥1280	≥1900
81	≥1220	≥1800

72	≥1150	≥1700
61	≥1090	≥1600
48	≥1010	≥1500
36	≥950	≥1400
24	≥870	≥1300
15	≥810	≥1200
8	≥730	≥1090
4	≥650	≥990
2	≥590	≥890
* The percentile of the perfect score was 99.98 on the 2400 scale and 99.93 on the 1600 scale.		
** 99+ means better than 99.5 percent of test takers.		

The older SAT (before 1995) had a very high ceiling. In any given year, only seven of the million test-takers scored above 1580. A score above 1580 was equivalent to the 99.9995 percentile.

In 2015 the average score for the Class of 2015 was 1490 out of a maximum 2400. That was down 7 points from the previous class's mark and was the lowest composite score of the past decade.

SAT-ACT score comparisons

The College Board and ACT, Inc. conducted a joint study of students who took both the SAT and the ACT between September 2004 (for the ACT) or March 2005 (for the SAT) and June 2006. Tables were provided to concord scores for students taking the SAT after January 2005 and before March 2016.

In May, 2016, the College Board released concordance tables to concord scores on the SAT used from March 2005 through January 2016 to the SAT used since March 2016, as well as tables to concord scores on the SAT used since March 2016 to the ACT.

History

Mean SAT Scores by year

Year of exam	Reading /Verbal Score	Math Score
1972	530	509
1973	523	506
1974	521	505
1975	512	498
1976	509	497
1977	507	496
1978	507	494
1979	505	493
1980	502	492
1981	502	492
1982	504	493
1983	503	494
1984	504	497
1985	509	500
1986	509	500
1987	507	501
1988	505	501
1989	504	502
1990	500	501
1991	499	500
1992	500	501

1993	500	503
1994	499	504
1995	504	506
1996	505	508
1997	505	511
1998	505	512
1999	505	511
2000	505	514
2001	506	514
2002	504	516
2003	507	519
2004	508	518
2005	508	520
2006	503	518
2007	502	515
2008	502	515
2009	501	515
2010	501	516
2011	497	514
2012	496	514
2013	496	514
2014	497	513
2015	495	511
2016	494	508

Historical average SAT test scores of college-bound seniors.

Many college entrance exams in the early 1900s were specific to each school and required candidates to travel to the school to take the tests. The College Board, a consortium of colleges in the northeastern United States, was formed in 1900 to establish a nationally administered, uniform set of essay tests based on the curricula of the boarding schools that typically provided graduates to the colleges of the Ivy League and Seven Sisters, among others.

In the same time period, Lewis Terman and others began to promote the use of tests such as Alfred Binet's in American schools. Terman in particular thought that such tests could identify an innate "intelligence quotient" (IQ) in a person. The results of an IQ test could then be used to find an elite group of students who would be given the chance to finish high school and go on to college. By the mid-1920s, the increasing use of IQ tests, such as the Army Alpha test administered to recruits in World War I, led the College Board to commission the development of the SAT. The commission, headed by Carl Brigham, argued that the test predicted success in higher education by identifying candidates primarily on the basis of intellectual promise rather than on specific accomplishment in high school subjects. In 1934, James Conant and Henry Chauncey used the SAT as a means to identify recipients for scholarships to Harvard University. Specifically, Conant wanted to find students, other than those from the traditional northeastern private schools that could do well at Harvard. The success of the scholarship program and the advent of World War II led to the end of the College Board essay exams and to the SAT being used as the only admissions test for College Board member colleges.

The SAT rose in prominence after World War II due to several factors. Machine-based scoring of multiple-choice tests taken by pencil had made it possible to rapidly process the exams. The G.I. Bill produced an influx of millions of veterans into higher education. The formation of the Educational Testing Service (ETS) also played a significant role in the expansion of the SAT beyond the roughly fifty colleges that made up the College Board at the time. The ETS was formed in 1947 by the College Board, Carnegie Foundation for the Advancement of Teaching, and the American Council on Education, to consolidate respectively the operations of the SAT, the GRE, and the achievement tests developed by Ben Wood for use with Conant's scholarship exams. The new organization was to be philosophically grounded in the concepts of open-minded, scientific research in testing with no doctrine to sell and with an eye toward public service. The ETS was chartered after the death of Brigham, who had opposed the creation of such an entity. Brigham felt that the interests of a consolidated testing agency would be more aligned with sales or marketing than with research into the science of testing. It has been argued that the interest of the ETS in expanding the SAT in order to support its operations aligned with the desire of public college and university faculties to have smaller, diversified, and more academic student bodies as a means to increase research activities. In 1951, about 80,000 SAT tests were taken; in 1961, about 800,000; and by 1971, about 1.5 million SAT tests were being taken each year.

A timeline of notable events in the history of the SAT follows.

1901 essay exams

On June 17, 1901, the first exams of the College Board were administered to 973 students across 67 locations in the United States, and two in Europe. Although those taking the test came from a variety of backgrounds, approximately one third were from New York, New Jersey, or Pennsylvania. The majority of those taking the test were from private schools, academies, or endowed schools. About 60% of those taking the test applied to Columbia University. The test contained sections on English, French, German, Latin, Greek, history, mathematics, chemistry, and physics. The test was not multiple choice, but instead was evaluated based on essay responses as "excellent", "good", "doubtful", "poor" or "very poor".

1926 test

The first administration of the SAT occurred on June 23, 1926, when it was known as the Scholastic Aptitude Test. This test, prepared by a committee headed by Princeton psychologist Carl Campbell Brigham, had sections of definitions, arithmetic, classification, artificial language, antonyms, number series, analogies, logical inference, and paragraph reading. It was administered to over 8,000 students at over 300 test centers. Men composed 60% of the test-takers. Slightly over a quarter of males and females applied to Yale University and Smith College. The test was paced rather quickly, test-takers being given only a little over 90 minutes to answer 315 questions. The raw score of each participating student was converted to a score scale with a mean of 500 and a standard deviation of 100. This scale was effectively equivalent to a 200 to 800 scale, although students could score more than 800 and less than 200.

1928 and 1929 tests

In 1928, the number of sections on the SAT was reduced to seven, and the time limit was increased to slightly under two hours. In 1929, the number of sections was again reduced, this time to six. These changes were designed in part to give test-takers more time per question. For these two years, all of the sections tested verbal ability: math was eliminated entirely from the SAT.

1930 test and 1936 changes

In 1930 the SAT was first split into the verbal and math sections, a structure that would continue through 2004. The verbal section of the 1930 test covered a more narrow range of content than its predecessors, examining only antonyms, double definitions (somewhat similar to sentence completions), and paragraph reading. In 1936, analogies were re-added. Between 1936 and 1946, students had between 80 and 115 minutes to answer 250 verbal questions (over a third of which were on antonyms). The mathematics test introduced in 1930 contained 100 free response questions to be answered in 80 minutes, and focused primarily on speed. From 1936 to 1941, like the 1928 and 1929 tests, the mathematics section was eliminated entirely. When the mathematics portion of the test was re-added in 1942, it consisted of multiple choice questions.

1941 and 1942 score scales

Until 1941, the scores on all SAT tests had been scaled to a mean of 500 with a standard deviation of 100. Although one test-taker could be compared to another for a given test date, comparisons from one year to another could not be made. For example, a score of 500 achieved on an SAT taken in one year could reflect a different ability level than a score of 500 achieved in another year. By 1940, it had become clear that setting the mean SAT score to 500 every year was unfair to those students who happened to take the SAT with a group of higher average ability.

In order to make cross-year score comparisons possible, in April 1941 the SAT verbal section was scaled to a mean of 500, and a standard deviation of 100, and the June 1941 SAT verbal section was equated (linked) to the April 1941 test. All SAT verbal sections after 1941 were equated to previous tests so that the same scores on different SAT tests would be comparable. Similarly, in June 1942 the SAT math section was equated to the April 1942 math section, which itself was linked to the 1942 SAT verbal section, and all SAT math sections after 1942 would be equated to previous tests. From this point forward, SAT mean scores could change over time, depending on the average ability of the group taking the test compared to the roughly 10,600 students taking the SAT in April 1941. The 1941 and 1942 score scales would remain in use until 1995.

1946 test and associated changes

Paragraph reading was eliminated from the verbal portion of the SAT in 1946, and replaced with reading comprehension, and "double definition" questions were replaced with sentence completions. Between 1946 and 1957 students were given 90 to 100 minutes to complete 107 to 170 verbal questions. Starting in 1958 time limits became more stable, and for 17 years, until 1975, students had 75 minutes to answer 90 questions. In 1959 questions on data sufficiency were introduced to the mathematics section, and then replaced with quantitative comparisons in 1974. In 1974 both verbal and math sections were reduced from 75 minutes to 60 minutes each, with changes in test composition compensating for the decreased time.

1960s and 1970s score declines

From 1926 to 1941, scores on the SAT were scaled to make 500 the mean score on each section. In 1941 and 1942, SAT scores were standardized via test equating, and as a consequence, average verbal and math scores could vary from that time forward. In 1952, mean verbal and math scores were 476 and 494, respectively, and scores were generally stable in the 1950s and early 1960s. However, starting in the mid-1960s and continuing until the early 1980s, SAT scores declined: the average verbal score dropped by about 50 points, and the average math score fell by about 30 points. By the late 1970s, only the upper third of test takers were doing as well as the upper half of those taking the SAT in 1963. From 1961 to 1977, the number of SAT tests taken per year doubled, suggesting that the decline could be explained by demographic changes in the group of students taking the SAT. Commissioned by the College Board, an independent study of the decline found that most (up to about 75%) of the test decline in the 1960s could be explained by compositional changes in the group of students taking the test; however, only about 25 percent of the 1970s decrease in test scores could similarly be explained. Later analyses suggested that up to 40 percent of the 1970s decline in scores could be explained by demographic changes, leaving unknown at least some of the reasons for the decline.

1994 changes

In early 1994, substantial changes were made to the SAT. Antonyms were removed from the verbal section in order to make rote memorization of vocabulary less useful. Also, the fraction of verbal questions devoted to passage-based reading material was increased from about 30% to about 50%, and the passages were chosen to be more like typical college-level reading material, compared to previous SAT reading passages. The changes for increased emphasis on analytical reading were made in response to a 1990 report issued by a commission established by the College Board. The commission recommended that the SAT should, among other things, "approximate more closely the skills used in college and high school work". A mandatory essay had been considered as well for the new version of the SAT; however, criticism from minority groups as well as a concomitant increase in the cost of the test necessary to grade the essay led the College Board to drop it from the planned changes.

Major changes were also made to the SAT mathematics section at this time, due in part to the influence of suggestions made by the National Council of Teachers of Mathematics. Test-takers were now permitted to use calculators on the math sections of the SAT. Also, for the first time since 1935, the SAT would now include some math questions that were not multiple choice, instead requiring students to supply the answers. Additionally, some of these "student-produced response" questions could have more than one correct answer. The tested mathematics content on the SAT was expanded to include concepts of slope of a line, probability, elementary statistics including median and mode, and counting problems.

1995 recentering (raising mean score back to 500)

By the early 1990s, average total SAT scores were around 900 (typically, 425 on the verbal and 475 on the math). The average scores on the 1994 modification of the SAT I were similar: 428 on the verbal and 482 on the math. SAT scores for admitted applicants to highly selective colleges in the United States were typically much higher. For example, the score ranges of the middle 50% of admitted applicants to Princeton University in 1985 were 600 to 720 (verbal) and 660 to 750 (math). Similarly, median scores on the modified 1994 SAT for freshmen entering Yale University in the fall of 1995 were 670 (verbal) and 720 (math). For the majority of SAT test takers, however, verbal and math scores were below 500: In 1992, half of the college-bound seniors taking the SAT were scoring between 340 and 500 on the verbal section and between 380 and 560 on the math section, with corresponding median scores of 420 and 470, respectively.

The drop in SAT verbal scores, in particular, meant that the usefulness of the SAT score scale (200 to 800) had become degraded. At the top end of the verbal scale, significant gaps were occurring between raw scores and uncorrected scaled scores: a perfect raw score no longer corresponded to an 800, and a single omission out of 85 questions could lead to a drop of 30 or 40 points in the scaled score. Corrections to scores above 700 had been necessary to reduce the size of the gaps and to make a perfect raw score result in an 800. At the other end of the scale, about 1.5 percent of test takers would have scored below 200 on the verbal section if that had not been the reported minimum score. Although the math score averages were closer to the center of the scale (500) than the verbal scores, the distribution of math scores was no longer well approximated by a normal distribution. These problems, among others, suggested that the original score scale and its reference group of about 10,000 students taking the SAT in 1941 needed to be replaced.

Beginning with the test administered in April 1995, the SAT score scale was recentered to return the average math and verbal scores close to 500. Although only 25 students had received perfect scores of 1600 in all of 1994, 137 students taking the April test scored a 1600. The new scale used a reference group of about one million seniors in the class of 1990: the scale was designed so that the SAT scores of this cohort would have a mean of 500 and a standard deviation of 110. Because the new scale would not be directly comparable to the old scale, scores awarded on April 1995 and later were officially reported with an "R" (for example, "560R") to reflect the change in scale, a practice that was continued until 2001. Scores awarded before April 1995 may be compared to those on the recentered scale by using official College Board tables. For example, verbal and math scores of 500 received before 1995 correspond to scores of 580 and 520, respectively, on the 1995 scale.

1995 re-centering controversy

Certain educational organizations viewed the SAT re-centering initiative as an attempt to stave off international embarrassment in regards to continuously declining test scores, even among top students. As evidence, it was presented that the number of pupils who scored above 600 on the verbal portion of the test had fallen from a peak of 112,530 in 1972 to 73,080 in 1993, a 36% backslide, despite the fact that the total number of test-takers had risen over 500,000.

2002 changes – Score Choice

In October 2002, the College Board dropped the Score Choice Option for SAT-II exams. Under this option, scores were not released to colleges until the student saw and approved of the score. The College Board has since decided to re-implement Score Choice in the spring of 2009. It is described as optional, and it is not clear if the reports sent will indicate whether or not this student has opted-in or not. A number of highly selective colleges and universities, including Yale, the University of Pennsylvania, and Stanford, have announced they will require applicants to submit all scores. Stanford, however, only prohibits Score Choice for the traditional SAT. Others, such as MIT and Harvard, have fully embraced Score Choice.

2005 changes, including a new 2400-point score

In 2005, the test was changed again, largely in response to criticism by the University of California system. Because of issues concerning ambiguous questions, especially analogies, certain types of questions were eliminated (the analogies from the verbal and quantitative comparisons from the math section). The test was made marginally harder, as a corrective to the rising number of perfect scores. A new writing section, with an essay, based on the former SAT II Writing Subject Test, was added, in part to increase the chances of closing the opening gap between the highest and midrange scores. Other factors included the desire to test the writing ability of each student; hence the essay. The essay section added an additional maximum 800 points to the score, which increased the new maximum score to 2400. The "New SAT" was first offered on March 12, 2005, after the last administration of the "old" SAT in January 2005. The mathematics section was expanded to cover three years of high school mathematics. The verbal section's name was changed to the Critical Reading section.

Scoring problems of October 2005 tests

In March 2006, it was announced that a small percentage of the SATs taken in October 2005 had been scored incorrectly due to the test papers' being moist and not scanning properly, and that some students had received erroneous scores. The College Board announced they would change the scores for the students who were given a lower score than they earned, but at this point many of those students had already applied to colleges using their original scores. The College Board decided not to change the scores for the students who were given a higher score than they earned. A lawsuit was filed in 2006 on behalf of the 4,411 students who received an incorrect score on the SAT. The class-action suit was settled in August 2007 when the College Board and Pearson Educational Measurement, the company that scored the SAT tests, announced they would pay $2.85 million into a settlement fund. Under the agreement each student could either elect to receive $275 or submit a claim for more money if he or she felt the damage was greater. A similar scoring error occurred on a secondary school admission test in 2010-2011 when the ERB (Educational Records Bureau) announced after the admission process was over that an error had been made in the scoring of the tests of 2010 (17%) of the students who had taken the Independent School Entrance Examination for admission to private secondary schools for 2011. Commenting on the effect of the error on students' school applications in *The New York Times*, David Clune, President of the ERB stated "It is a lesson we all learn at some point—that life isn't fair."

2008 changes

In late 2008, a new variable came into play. Previously, applicants to most colleges were required to submit all scores, with some colleges that embraced Score Choice retaining the option of allowing their applicants not to have to submit all scores. However, in 2008, an initiative to make Score Choice universal had begun, with some opposition from colleges desiring to maintain score report practices. While students theoretically now have the choice to submit their best score (in theory one could send any score one wishes to send) to the college of their choice, some colleges and universities, such as Cornell, ask that students send all test scores. This had led the College Board

to display on their web site which colleges agree with or dislike Score Choice, with continued claims that students will still never have scores submitted against their will. Regardless of whether a given college permits applicants to exercise Score Choice options, most colleges do not penalize students who report poor scores along with high ones; many universities, such as Columbia and Cornell expressly promise to overlook those scores that may be undesirable to the student and/or to focus more on those scores that are most representative of the student's achievement and academic potential. College Board maintains a list of colleges and their respective score choice policies that is recent as of November 2011.

2012 changes

Beginning in 2012, test takers are required to submit a current, recognizable photo during registration. Students are required to present their photo admission ticket – or another acceptable form of photo ID – for admittance to their designated test center. Student scores and registration information, including the photo provided, are made available to the student's high school. In the event of an investigation involving the validity of a student's test scores, their photo may be made available to institutions to which they have sent scores. Any college that is granted access to a student's photo is first required to certify that they are all admitted students.

2016 changes, including the return to a 1600-point score

On March 5, 2014, the College Board announced its plan to redesign the SAT in order to link the exam more closely to the work high school students encounter in the classroom. The new exam was administered for the first time in spring 2016. Some of the major changes are: an emphasis on the use of evidence to support answers, a shift away from obscure vocabulary to words that students are more likely to encounter in college and career, a math section that is focused on fewer areas, a return to the 1600-point score scale, an optional essay, and the removal of penalty for wrong answers (rights-only scoring). To combat the perceived advantage of costly test preparation courses, the College Board announced a new partnership with Khan Academy to offer free online practice problems and instructional videos.

Name changes

The SAT has been renamed several times since its introduction in 1926. It was originally known as the Scholastic Aptitude Test In 1990, a commission set up by the College Board to review the proposed changes to the SAT program recommended that the meaning of the initialism SAT be changed to "Scholastic Assessment Test" because a "test that integrates measures of achievement as well as developed ability can no longer be accurately described as a test of aptitude". In 1993, the College Board changed the name of the test to SAT I: Reasoning Test; at the same time, the name of the Achievement Tests was changed to SAT II: Subject Tests. The Reasoning Test and Subject Tests were to be collectively known as the Scholastic Assessment Tests. According to the president of the College Board at the time, the name change was meant "to correct the impression among some people that the SAT measures something that is innate and impervious to change regardless of effort or instruction. The new SAT debuted in March 1994, and was referred to as the Scholastic Assessment Test by major news organizations. However, in 1997, the College Board announced that the SAT could not properly be called the Scholastic Assessment Test, and that the letters SAT did not stand for anything. In 2004, the Roman numeral in SAT I: Reasoning Test was dropped, making SAT Reasoning Test the new name of the SAT.

Math–verbal achievement gap

In 2002, Richard Rothstein (education scholar and columnist) wrote in *The New York Times* that the U.S. math averages on the SAT and ACT continued their decade-long rise over national verbal averages on the tests.

Reuse of old SAT exams

The College Board has been accused of completely reusing old SAT papers previously given in the United States. The recycling of questions from previous exams has been exploited to allow for cheating on exams and impugned the validity of some students' test scores, according to college officials. Test preparation companies in Asia have been found to provide test questions to students within hours of a new SAT exam's administration.

Perception

Association with culture

For decades many critics have accused designers of the verbal SAT of cultural bias as an explanation for the disparity in scores between poorer and wealthier test-takers. A famous (and long past) example of this bias in the SAT I was the oarsman–regatta analogy question. The object of the question was to find the pair of terms that had the relationship most similar to the relationship between "runner" and "marathon". The correct answer was "oarsman" and "regatta". The choice of the correct answer was thought to have presupposed students' familiarity with rowing, a sport popular with the wealthy. However, according to Murray and Herrnstein, the black-white gap is smaller in culture-loaded questions like this one than in questions that appear to be culturally neutral. Analogy questions have since been replaced by short reading passages.

Association with family income

Parents with a higher income can afford to spend money on their child's education if they are not satisfied with it, while families with a lower income cannot. Families that have a higher family income can purchase multiple prep books, more specifically the College Board one. They can also pay for tutors, prep classes, and other resources. These resources not only prepare the students for the knowledge required in the test, but also provide practice tests and sample problems that can be studied and possibly be seen on the real exam. Another source of this difference comes from the quality of schooling as family income increases. Those with higher income families, "tend to have better teachers, more resource-rich educational environments, more educated parents who can help them with school and, sometimes, expensive SAT tutoring." The quality of education the student receives greatly affect the success during the test because the school not only relates to the student's knowledge but also many times correlates with the students study habits and motivation to succeed. As with racial bias, this correlation with income could also be due to the social class of the makers of the test, although empirical research suggests that poorer students actually perform worse on 'neutral' than such 'privileged' questions.

Association with gender

The largest association with gender on the SAT is found in the math section where male students, on average, score higher than female students by approximately 30 points. In 2013, the American College Testing Board released a report boys outperformed girls on the mathematics section of the test.

Association with race and ethnicity

African American, Hispanic, and Native American students, on average, perform an order of one standard deviation lower on the SAT than white and Asian students.

Researchers believe that the difference in scores is closely related to the overall achievement gap in American society between students of different racial groups. This gap may be explainable in part by the fact that students of disadvantaged racial groups tend to go to schools that provide lower educational quality. This view is supported by evidence that the black-white gap is higher in cities and neighborhoods that are more racially segregated. It has also been suggested that stereotype threat has a significant effect on lowering achievement of minority students. For example, African Americans perform worse on a test when they are told that the test measures "verbal reasoning ability", than when no mention of the test subject is made. Other research cites poorer minority proficiency in key coursework relevant to the SAT (English and math), as well as peer pressure against students who try to focus on their schoolwork ("acting white"). Cultural issues are also evident among black students in wealthier households, with high achieving parents. John Ogbu, a Nigerian-American professor of anthropology, found that instead of looking to their parents as role models, black youth chose other models like rappers and did not put forth the effort to be a good student. However, they felt that racism was wrong.

One set of studies has reported differential item functioning - namely, some test questions function differently based on the racial group of the test taker, reflecting some kind of systematic difference in a group ability to understand certain test questions or to acquire the knowledge required to answer them. In 2003 Freedle published data showing that Black students have had a slight advantage on the verbal questions that are labeled as difficult on the SAT, whereas white and Asian students tended to have a slight advantage on questions labeled as easy. Freedle argued that these findings suggest that "easy" test items use vocabulary that is easier to understand for white middle class students than for minorities, who often use a different language in the home environment, whereas the difficult items use complex language learned only through lectures and textbooks, giving both student groups equal opportunities to acquiring it. The study was severely criticized by the ETS board, but the findings were replicated in a subsequent study by Santelices and Wilson in 2010.

There is no evidence that SAT scores systematically underestimate future performance of minority students. However, the predictive validity of the SAT has been shown to depend on the dominant ethnic and racial composition of the college. Some studies have also shown that African American students under-perform in college relative to their white peers with the same SAT test scores; researchers have argued that this is likely because white students tend to benefit from social advantages outside of the educational environment (for example, high parental involvement in their education, inclusion in campus academic activities, positive bias from same-race teachers and peers) which result in better grades.

Christopher Jencks concludes that as a group African Americans have been harmed by the introduction of standardized entrance exams such as the SAT. This, according to him, is not because the tests themselves are flawed, but because of labeling bias and selection bias; the tests measure the skills that African Americans are less likely to develop in their socialization, rather than the skills they are more likely to develop. Furthermore, standardized entrance exams are often labeled as tests of general ability, rather than of certain aspects of ability. Thus, a situation is produced in which African American ability is consistently underestimated within the education and workplace environments, contributing in turn to selection bias against them which exacerbates underachievement.

Dropping SAT

A growing number of colleges have joined the SAT optional movement. These colleges do not require the SAT for admission.

One example of a college that did this is Drew University in New Jersey. After they adopted an optional SAT policy, they had a 20% increase in applications. Dean of Admissions Mary Beth Carey says that "Our own research showed us that high school grade point average is by far the most

important predictor of success in college." The college reported that they accepted their most diverse class ever as a result of the policy.

In a 2001 speech to the American Council on Education, Richard C. Atkinson, the president of the University of California, urged dropping the SAT as a college admissions requirement:

Anyone involved in education should be concerned about how overemphasis on the SAT is distorting educational priorities and practices, how the test is perceived by many as unfair, and how it can have a devastating impact on the self-esteem and aspirations of young students. There is widespread agreement that overemphasis on the SAT harms American education.

In response to threats by the University of California to drop the SAT as an admission requirement, the College Entrance Examination Board announced the restructuring of the SAT, to take effect in March 2005, as detailed above.

In the 1960s and 1970s there was a movement to drop achievement scores. After a period of time, the countries, states and provinces that reintroduced them agreed that academic standards had dropped, students had studied less, and had taken their studying less seriously. They reintroduced the tests after studies and research concluded that the high-stakes tests produced benefits that outweighed the costs.

IQ studies

Frey and Detterman (2003) investigated associations of SAT scores with intelligence test scores. Using an estimate of general mental ability, or *g*, based on the ASVAB test battery, which can be best thought of as representing crystallized intelligence (learned abilities), they found SAT scores to be highly correlated with *g*(r=.82 in their sample, .857 when adjusted for non-linearity) in their sample taken from a 1979 national probability survey. Additionally, they investigated the correlation between SAT results, using the revised and recentered form of the test, and scores on the Raven's Advanced Progressive Matrices, a test of fluid intelligence (reasoning), this time using a non-random sample. They found that the correlation of SAT results with scores on the Raven's Advanced Progressive Matrices was .483. They estimated that this latter correlation would have been about 0.72 were it not for the restriction of ability range in the sample. They also noted that there appeared to be a ceiling effect on the Raven's scores which may have suppressed the correlation. Beaujean and colleagues (2006) have reached similar conclusions to those reached by Frey and Detterman.

Preparation

SAT preparation is a highly lucrative field[] and many companies and organizations offer test preparation in the form of books, classes, online courses, and tutoring. The test preparation industry began almost simultaneously with the introduction of university entrance exams in the U.S. and flourished from the start.

The College Board maintains that the SAT is essentially uncoachable and research by the College Board and the National Association of College Admission Counseling suggests that tutoring courses result in an average increase of about 20 points on the math section and 10 points on the verbal section. Other studies have shown significantly different results. A longitudinal study from Ohio State showed that taking private SAT prep classes correlated with scores higher by ~60 points. A study from Oxford showed that coaching courses boosted scores by an average of 56 points.

Montgomery and Lilly (2012) performed a systematic literature review of all published SAT coaching research in search of high quality studies (defined as those with randomized controlled trials). They

found that the randomized treatments resulted in V/M gains of +23/32 points for a total of +56; the high quality study that showed the highest score increase was Johnson (1984; San Francisco) which was based on a 30-hour prep course that showed an average increase of 178 points. The Johnson San Francisco study was also the only high quality study found on a prep course of 30 hours or more in length, although validity of this outlier study is uncertain due to the attrition of half the participants.

Use by high-IQ societies

Certain high IQ societies, like Mensa, the Prometheus Society and the Triple Nine Society, use scores from certain years as one of their admission tests. For instance, the Triple Nine Society accepts scores of 1450 on tests taken before April 1995, and scores of at least 1520 on tests taken between April 1995 and February 2005.

The SAT is sometimes given to students younger than 13 by organizations such as the Study of Mathematically Precocious Youth, Johns Hopkins Center for Talented Youth, Duke TIP, and other organizations who use the results to select, study and mentor students of exceptional ability.

Writing section

In 2005, MIT Writing Director Pavan Sreekireddy plotted essay length versus essay score on the new SAT from released essays and found a high correlation between them. After studying over 50 graded essays, he found that longer essays consistently produced higher scores. In fact, he argues that by simply gauging the length of an essay without reading it, the given score of an essay could likely be determined correctly over 90% of the time. He also discovered that several of these essays were full of factual errors; the College Board does not claim to grade for factual accuracy.

Perelman, along with the National Council of Teachers of English also criticized the 25-minute writing section of the test for damaging standards of writing teaching in the classroom. They say that writing teachers training their students for the SAT will not focus on revision, depth, accuracy, but will instead produce long, formulaic, and wordy pieces. "You're getting teachers to train students to be bad writers", concluded Perelman