

Transitioning to a Digital SAT

Ensuring Tests and Scores Are Valid, Reliable, and Fair

In Summary

- The digital SAT[®] will continue to measure the skills and knowledge that research shows are crucial to college readiness and success and at the same level of rigor as the current SAT.
- Digital SAT scores will have the same meaning and will be measured on the same scale as current SAT scores, meaning that no concordance table will be necessary. Digital scores will be comparable and reliable within and across test administrations and among students.
- Our test development and psychometric methods and research are built on well-established, deeply researched assessment science and design, including rigorous pretesting.
- We are conducting a robust, multiyear research program to ensure that the digital test questions and format are valid, reliable, and fair, and we will share detailed test information and research results publicly.

College Board will begin to administer the SAT digitally in spring 2023.

The SAT will be delivered digitally internationally beginning in spring 2023 and in the U.S. in spring 2024. The PSAT/NMSQT[®] and PSAT[™] 8/9 will be delivered digitally in fall 2023.

Going digital allows us to give every student a unique test form, so the digital SAT will be far more secure. The digital SAT will also be shorter - about two hours instead of three with more time per question, and will feature more texts, topics, and authors that reflect the diversity of our students. And if a student doesn't have a device, we'll provide one on test day.

While the transition to digital will bring great benefits for students taking the SAT, many important things about the test won't change. The SAT Suite will continue to measure the knowledge and skills that students are learning in high school and that matter most for college and career readiness. The SAT will still be scored on a 1600-point scale, and educators and students can continue to track growth across the SAT Suite of Assessments over time. The assessments will continue to be administered in a school or in a test center with a proctor present—not at home. Students will still have access to free practice resources on Khan Academy[®]. And students taking the SAT Suite will continue to connect to scholarships and the College Board National Recognition Programs.

Ensuring score reliability, validity, and fairness

In consultation with higher education members, we are designing the digital SAT so that a score on the digital SAT will represent the same level of achievement in math and reading

and writing as the corresponding score on the current SAT. This will allow current SAT and digital SAT scores to be evaluated side-by-side without a concordance table, facilitating a seamless transition from considering scores for the current SAT to the digital SAT so higher ed users have the same confidence in using digital SAT scores as one of many factors taken into consideration in a holistic admission or placement process.

Four key design elements will ensure that the scores on the current SAT and digital SAT will carry the same meaning.

1. First, the content of the current SAT and the digital SAT is highly congruent and comparable.
2. Second, we pretest digital SAT questions to establish that those questions are comparable in difficulty and other psychometric properties to questions on the current SAT.
3. Third, the multistage adaptive testing (MST) model employed in the digital SAT means that we are able to assess students on specific knowledge and skills in a manner highly comparable to that used for the current SAT, with shorter yet still highly reliable tests.
4. Fourth, College Board has conducted or planned numerous rigorous studies for the digital SAT prior to launch in 2023 that are comparable to the studies performed for the current SAT, including a series of concordance studies to confirm the direct relationship between digital SAT and current SAT scores. Test specifications and research findings will be published regularly and will be available for review.

1. The digital SAT maintains strong congruity and continuity to the current SAT in terms of the knowledge, skills, and content domains assessed.

The current version of the SAT and the digital SAT assess comparable skills, knowledge, and content domains in two areas: (1) Reading and Writing and (2) Math.

Digital SAT Reading and Writing Content Domains	Digital SAT Math Content Domains
<ul style="list-style-type: none">Information and IdeasCraft and StructureExpression of IdeasStandard English Conventions	<ul style="list-style-type: none">AlgebraAdvanced MathProblem-Solving and Data AnalysisGeometry and Trigonometry

We know from prior research, based on scores from more than 220,000 students attending 165 colleges and universities, that SAT scores are useful predictors of [first-semester college course performance in the matching academic domains](#).¹ Beyond these empirical links to college readiness in literacy and math, many studies have also demonstrated the strong validity of the SAT for predicting key college outcomes, including [first-year grade-point average \(FYGPA\)](#),² [retention through each year of college](#),³ and [degree completion](#).⁴

Extensive pretesting of digital SAT questions to date has demonstrated that the questions, and the tests subsequently built from them, exhibit levels of difficulty comparable to those exhibited by their current SAT counterparts. The transition to the digital SAT also preserves the same range of item difficulty.

2. We pretest digital SAT questions to establish comparability in difficulty and other psychometric properties to questions on the current SAT.

Pretesting is a standard industry and academic approach to collecting student responses on newly developed test questions. This process allows developers to build tests that meet the targeted level of difficulty, and distinguish sufficiently among test takers of varying achievement levels. First, students are matched for academic achievement, then their responses to individual items are reviewed. We remove items that show significant DIF (Differential Item Functioning) by race/ethnicity, gender, or English as first language in alignment with industry best standards.

3. The multistage adaptive testing (MST) model employed for the digital SAT permits test developers to control the skills and knowledge on which students are assessed with shorter yet still highly reliable tests.

The digital SAT Suite will utilize a multistage adaptive methodology. In a digital multistage adaptive SAT test, each test section (Reading and Writing, Math) is divided into two stages, called modules. Students answer a set of questions in the first module before moving on to the next. The questions in the second module are configured based on performance in the first module. This means that the test “adapts” to show questions that are more appropriate to a student’s performance level.

Adaptive testing has a firm foundation in both research and practice. The foundations of adaptive testing date to the early 20th century, with major theoretical advances being made in the 1960s and with operational adaptive tests being available since the 1980s. Today, adaptive testing is used successfully for a wide range of purposes, including assessment of secondary students’ skills and knowledge (SBAC, Virginia SOLs, NWEA’s MAP Suite), graduate admission, professional licensure, and aptitude assessment for the U.S. military.

Multistage adaptive testing strikes a balance between traditional linear (nonadaptive) paper-based testing and question-by-question adaptive models. The MST model for the digital SAT allows for a similar level of control over content as with the current SAT, which is important to ensuring the test scores can be used in the same way. The ability to control content is greater in multistage adaptive testing than in a question-by-question adaptive model that selects on the fly the content to be delivered. The multistage format also allows students to review and revise their answers to questions when working within a module, which is not possible in a question-by-question adaptive model.

At the same time, the MST model retains key benefits in common with question-by-question adaptive models, chiefly about ability to deliver shorter tests with high reliability. This is possible because adaptive models are able to target the content students encounter to their achievement levels, whereas linear tests must include more content to achieve the same results.

In short, the MST model used for the digital SAT

- Maintains assessment precision with fewer test questions as compared to the current SAT, allowing for shorter tests that nonetheless retain the reliability of longer ones.

¹ Westrick, P.A., Marini J.P., Shmueli, D., Young, L., Shaw E.J., & Ng, H. (2020). Validity of the SAT for predicting first-semester, domain-specific grades (College Board Research Report). New York: College Board.

² Westrick, P.A., Marini, J.P., Young, L., Ng, H., Shmueli, D., & Shaw, E.J. (2019). Validity of the SAT for predicting first-year grades and retention to the second year (College Board Research Report). New York: College Board.

³ Marini, J.P., Westrick, P.A., & Shaw, E.J. (2021). *Examining the stability of SAT predictive relationships across cohorts and over time* (College Board Research Brief). New York: College Board.

⁴ Mattern, K.D., & Patterson, B.F. (2014). *Synthesis of recent SAT validity findings: Trend data over time and cohorts* (College Board Research in Review). New York: College Board.

- Permits a higher level of control over administered content than question-by-question adaptive models.
- Covers content and addresses skills and knowledge in ways similar to what is found on the current SAT.
- Allows students to review (and possibly change) their responses to previously administered questions within the same stage.

4. Extensive research studies will validate the validity, comparability, and reliability of digital SAT scores (see below).

College Board has made or will conduct numerous rigorous studies related to the digital SAT, including a series of concordance studies to confirm the direct relationship between digital SAT and current SAT scores. Test specifications and key research findings will be published when available.

FAQs and Research Study Visual Timeline

If the test is shorter, how can it be as rigorous and as valid in predicting success as the current test?

Rigor is not tied to length of assessment. A benefit of using adaptive technology is to accurately assess the same skills and knowledge and arrive at the same score with fewer test questions than is possible with linear (nonadaptive) test forms. With adaptive technology the difficulty of question sets can be tailored to the individual student, which in turn allows us to accurately assess achievement in literacy and math with fewer test questions. Moreover, extensive, ongoing pretesting of digital SAT questions has established that these questions are as capable of eliciting the same levels of difficulty as current SAT questions.

Digital SAT Study	Description	Timeline
Preliminary Reading and Writing comparability analysis	Ascertained which prospective digital SAT question types were most highly correlated with existing question types and, all else being equal, should be prioritized for inclusion on the digital test to facilitate concordance ⁵ .	January–February 2021
Student postexperience surveys and focus groups	Systematically collect information about student perceptions of their digital SAT test-taking experience as a way to inform the test’s design.	Ongoing from February 2021
Timing study	Established what test lengths were necessary to give the digital SAT test takers sufficient time to thoughtfully respond to each test question without rushing to completion.	fall 2021
External review of test materials	Routinely engage independent experts to evaluate representative test materials for soundness and fairness and to inform future test development.	Ongoing from December 2021
Vertical scaling studies	Will establish the reporting scale for the entire digital SAT Suite by collecting student performance data on full-length exams under typical administration conditions.	spring 2022; fall/spring 2022–2023
Concordance studies	Will confirm the direct relationship between digital SAT and current SAT scores; will yield college-reportable scores for student participants.	spring 2022; fall 2022
Pilot predictive validity study and concurrent validity study	Will ascertain the extent to which digital SAT scores relate to (i.e., “agree” with) other established educational measures, such as FYGPA, high school GPA, and AP [®] Exam scores.	2022–2023
Cognitive labs (verbal protocol analysis studies)	Expected to confirm that, like those on the current SAT, the digital SAT test questions elicit from students the kinds of higher-order, cognitively complex skills and knowledge required for college and career readiness; will devote particular attention to English learners and students with disabilities.	2022–2023
Curriculum survey	Will collect up-to-date data from secondary and postsecondary educators regarding college and career readiness priorities and prerequisites.	2023–2024

⁵ This phase of research focused on Reading and Writing because the paper-based and digital SAT Math sections employ highly similar test question formats.

If the test is different for each student, how can it be considered standardized?

Standardization in large-scale assessment programs does not require that all students receive exactly the same test content. It is a common, industry-standard practice for test makers to develop multiple, parallel test forms to administer at different times to different test takers. This approach is used currently with paper-and-pencil SAT testing, in which students see different test forms on any given test day. Standardization using this approach is achieved first by building each test form according to the same set of content and statistical specifications and second by taking each test form through an equating process to ensure that scores reported for each and every test form are comparable—i.e., that a score of 540 means the same in terms of level of achievement across multiple forms. The digital SAT is no different in that every test form administered is tightly aligned to a specified set of content and statistical specifications and goes through an equating process to ensure that scores are comparable and interchangeable across forms.

If the test is shorter and adaptive, how can it be as valid in predicting student success as the current test?

The ability of the SAT to predict future student outcomes is not directly tied to its length or whether it is adaptive or linear (nonadaptive). It is, rather, a product of the overall test design,

which—whether we consider the current SAT or the digital SAT—is centrally focused on measuring students’ attainment of those skills and knowledge that best available evidence indicates are prerequisite for readiness and success in common first-year, credit-bearing postsecondary courses. Given that the designs of the paper-based and digital SAT tests are highly similar, there is strong reason to believe that the digital SAT will be as successful as the current SAT in predicting postsecondary success. Beginning in spring 2022, we will conduct a number of different studies formally assessing the validity of the digital SAT as a predictor of future outcomes, including a concurrent validity study to determine how well scores on the digital SAT relate to (“agree” with) other established educational measures and multiple predictive validity studies to determine how performance on the digital SAT relates to a student’s grades in the first year of college.

How do you explain significant differences in SAT scores across race/ethnicity?

Real inequities exist in U.S. education, and they’re reflected in every measure of academic achievement, including the SAT Suite and other standardized assessments. Performance differences across subgroups of students reflect an unequal K–12 system. That’s why the SAT should only be considered in the context of where students live and go to school, and an SAT score should never be a veto on a student’s plans or ambitions.

Research Agenda for the Digital SAT Suite

Study Purpose	Targets (Students)	2021				2022				2023	
		Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2
Item Pretesting: Test items weekly with 10k students (per subject) to evaluate item difficulty and viability	~2M	[Timeline showing activity from Q1 2021 through Q2 2023]									
Reading & Writing Study (February 2021): Test the new Reading and Writing questions versus the current SAT Evidence based on reading and writing questions.	2,500	[Timeline showing activity in Q1 2021]									
Timing Study (November 2021): Evaluate proposed timing for the digital SAT	5,000	[Timeline showing activity in Q4 2021]									
Vertical Scaling Study (spring 2022 and fall 2023) Vertically link the flexible PSAT/NMSQT and PSAT 8/9 tests to the concordant new digital SAT scale	21,000	[Timeline showing activity in Q1 2022 and Q1 2023]									
Concordance Study (spring 2022 and fall 2023) : Establish linking (“straight-line concordance”) between the new digital SAT and current SAT	15,000	[Timeline showing activity in Q1 2022 and Q4 2022]									
Concurrent Validity Study: Evaluate how the new digital SAT scores relate to other educational measures that are used by higher education and compare to current SAT	4,500	[Timeline showing activity in Q1 2022]									
Predictive Validity Study: Demonstrate predictive validity between the new digital SAT and college performance/outcomes.	2,000	[Timeline showing activity in Q3 2022 and Q1 2023]									

→ Admin

→→ Analysis

★ College reportable scores (August, December)