

Test Characteristics of the SAT®: Reliability, Difficulty Levels, Completion Rates



January 2012–December 2012

	Critical Reading (67 Questions)	Mathematics (54 Questions)	Writing MC (49 Questions)	Writing Composite* (with Essay)
Reliability coefficient	.91-.92	.91-.93	.89-.91	.89-.91
Standard error of measurement (SEM)	30–32	28–32	3.4–3.5	31–35
Standard error of difference (SED)	42–45	40–45	5	44–49
Difficulty (average percent correct)	.55–.61	.55–.62	.58–.66	

*Essay reliability statistics, which are the proportion of essays that are scored the same or within one point, range between .99 and .100.

Critical Reading Completion Rates	Critical Reading 1, 25 min. No. of questions			Critical Reading 2, 25 min. No. of questions			Critical Reading 3, 20 min. No. of questions	
	(23)	(24)	(25)	(23)	(24)	(25)	(19)	(20)
Percent completing the section	88–90	77–86	77–83	79–83	81–89	67–78	73–96	85
Percent of students completing 75 percent of the section	99–100	99	98–99	96–99	95–99	95–98	97–100	97
Number of questions reached by 80 percent of the students	23	23–24	24–25	22–23	24	23–24	18–19	20
Average number of questions not reached	0.2–0.3	0.3–0.6	0.5–0.6	0.5–0.7	0.3–0.7	0.7–1.2	.1-.7	0.5

Points to Note

- Data in this table refer only to tests administered from January 2012 through December 2012.
- Reliability: The data show that the SAT is a reliable test and that an individual test-taker would tend to earn similar scores on repeated testing.
- The standard error of measurement (SEM) is approximately 32 points, indicating that, on average, if a test-taker takes the test again, there is a 68 percent chance that his or her second score will fall within a range 32 points above or below the first score.
- When comparing scores between students for the same measure (critical reading, mathematics, writing), the standard error of the difference (SED) can be used to assess how much scores must differ in order to reflect true differences in ability. If two scores differ by at least SED times 1.5, it is unlikely that the two scores indicate that the two candidates are equal in ability, since this level difference would occur 10 percent of the time or less. For example, when the SED is 40 points, you can be reasonably confident that if the score difference between two test-takers is greater than 60 points (40 x 1.5), the two test-takers are not likely to be equal in true ability.
- Difficulty: The SAT is shown to be appropriately difficult for the intended test-taking population.
- While the total number of items in the critical reading section of the test is 67, the number of items in the individual critical reading sections can vary slightly from form to form.

Test Characteristics of the SAT®: Reliability, Difficulty Levels, Completion Rates



January 2012–December 2012

Mathematics Completion Rates	Mathematics 1, 25 min. No. of questions (20)	Mathematics 2, 25 min. No. of questions (18)	Mathematics 3, 20 min. No. of questions (16)
Percent completing the section	50–79	40–71	55–92
Percent of students completing 75 percent of the section	96–99	92–98	97–99
Number of questions reached by 80 percent of the students	18–19	15–17	14–16
Average number of questions not reached	0.6–1.1	0.8–1.5	0.2–1.0

Writing Completion Rates	Writing 1, 25 min. No. of questions (35)	Writing 2, 10 min. No. of questions (14)
Percent completing the section	74–86	88–94
Percent of students completing 75 percent of the section	100	98–100
Number of questions reached by 80 percent of the students	34–35	14
Average number of questions not reached	0.4–0.6	0.1–0.2

Points to Note

- Data in this table refer only to tests administered from January 2012 through December 2012.
- Reliability: The data show that the SAT is a reliable test and that an individual test-taker would tend to earn similar scores on repeated testing.
- The standard error of measurement (SEM) is approximately 32 points, indicating that, on average, if a test-taker takes the test again, there is a 68 percent chance that his or her second score will fall within a range 32 points above or below the first score.
- When comparing scores between students for the same measure (critical reading, mathematics, writing), the standard error of the difference (SED) can be used to assess how much scores must differ in order to reflect true differences in ability. If two scores differ by at least SED times 1.5, it is unlikely that the two scores indicate that the two candidates are equal in ability, since this level difference would occur 10 percent of the time or less. For example, when the SED is 40 points, you can be reasonably confident that if the score difference between two test-takers is greater than 60 points (40 x 1.5), the two test-takers are not likely to be equal in true ability.
- Difficulty: The SAT is shown to be appropriately difficult for the intended test-taking population.
- While the total number of items in the critical reading section of the test is 67, the number of items in the individual critical reading sections can vary slightly from form to form.