

Interpreting CAHSEE Scores 2004

In July 2003, the State Board of Education approved revised blueprints for both portions (i.e., English-language arts [ELA] and math) of the California High School Exit Examination (CAHSEE) beginning with the February 2004 administration. The Board also directed the California Department of Education (CDE) to reduce the test length and testing time. Beginning in February 2004, the CAHSEE administration spans two days rather than three.

Summary of Test Changes Related to the Revised Blueprints

To reduce the ELA testing time from two days to a single day, the ELA portion of the CAHSEE has been revised in several ways. The revised ELA blueprint maintains the equal weighting of content standards related to reading and writing. The revisions include reducing the number of essays from two to one and reducing the number of multiple-choice (MC) questions from 82 to 72. In addition, the relative numbers of questions in other ELA strands have been revised and the relative weight of the essay in relation to the MC questions has been changed so that the weight of writing in the total score remains at about 50 percent. Specifically, the MC questions will each carry a weight of 1.0 and the four-point essay score will be weighted 4.5 per score point. In addition, 27 questions in the ELA test will assess writing standards. This results in a 90-point raw score that consists of 45 points based on MC reading questions, 27 points based on MC writing questions, and 18 points based on the essay.

While the math portion of CAHSEE remains at 80 questions, the math blueprint was slightly revised in the content standards being assessed for Mathematical Reasoning, and Statistics, Data Analysis, and Probability. In addition, the test developer worked to remove any extraneous information from the math questions and refined the difficulty specifications for the math questions used on the test. Because of these changes, and the additional access to standards-based instruction, students in the class of 2006 performed better on the new version of the math portion of the CAHSEE than the classes of 2004 and 2005 performed on the previous version.

Based on these changes to the blueprints, a new CAHSEE standard setting session was held in September 2003. As a result, a new CAHSEE scale was defined beginning with the February 2004 administration. This scale **is not** the same as the scale used for CAHSEE administrations between March 2001 and May 2003. The new CAHSEE scale ranges from 275 to 450. Passing scores for the ELA and math portions of the CAHSEE remain at 350 on the new CAHSEE scale, which is equal to approximately 60% and 55% correct for the ELA and math portions, respectively. Please note that 350 on the new CAHSEE scale is not equivalent to 350 on the old CAHSEE scale; therefore, CAHSEE results beginning with the February 2004 administration **should not** be compared with CAHSEE results based on the previous blueprint.

Important Testing Concepts

To adequately interpret ELA and math test scores across administrations of the CAHSEE, the following important testing and statistical concepts need to be understood:

- ✓ Standard Error of Measurement (SEM)
- ✓ Conditional Standard Error of Measurement (CSEM)
- ✓ Raw Score to Scale Score Conversion
- ✓ Weighting of Examination Portions

For each administration of the CAHSEE, the statistics may vary slightly. Text describing each of the above testing and statistical concepts and how they apply to the CAHSEE is listed below.

Standard Error of Measurement

As with every test score, a student's score on the CAHSEE includes some uncertainty. While uncertainty can come from a variety of sources, the amount of uncertainty can be described by a statistic called the Standard Error of Measurement (SEM). Statisticians define the "error of measurement" as the difference between the score a student obtains on a test (an observed score) and the hypothetical "true score" that the same student would obtain if a test could measure the student's achievement level with perfect accuracy. Statistical theory indicates that a student will have an observed score within one SEM of the student's true score about 68 percent of the time and within two SEMs of the student's true score about 95 percent of the time.

Conditional Standard Error of Measurement

The SEM is not the same at all score levels. The Conditional Standard Error of Measurement (CSEM) is the SEM at a specific score level. The CSEM for scores near the top and bottom of the CAHSEE scale, for example, are typically larger than the CSEM near the middle of the scale around the passing score of 350. Stated simply, the scores in the middle of the scale are generally more accurate measures of student performance than the scores at the lower or higher ends of the scale. It is critical to have accuracy at the passing score because the CAHSEE is a high-stakes examination.

To illustrate the CSEM principle, consider the following example. If a student achieves a score of 410 on the ELA portion of the CAHSEE and the CSEM for that score is 12 points, we would be about 68 percent confident that the student's true score lies between 422 and 398 (i.e., the student's score plus or minus 12 points). We would be 95 percent confident that the student's true score lies between 434 and 386, which is a band around the student's score equal to two CSEMs (i.e., the student's score plus or minus 24 points).

Raw Score to Scale Score Conversion

Students have multiple opportunities to pass the ELA and math portions of the CAHSEE. When administering multiple forms of a test there is a need for a "constant scale." This means that the passing score must represent essentially the same level of achievement on all forms of the CAHSEE. To maintain comparability of scores across multiple test forms, number correct or raw scores are converted to scale scores. The CAHSEE scale scores for ELA and math range from 275 to 450, with 350 being the score needed to pass each portion of the exam. The raw score to scale score conversion reflects the relationship between difficulty of individual test questions in each test form and the constant measure of achievement indicated by the reported scale scores. For different test forms, the expected number correct score for a given level of achievement may vary somewhat due to (usually small) differences in the average difficulty of the questions in one form compared to the average difficulty of questions in other test forms. This is why the conversion tables for each test administration will differ slightly in relating raw scores to scale scores. The procedure of converting the raw scores to scale scores involves scaling and equating.

Weighting ELA Examination Portions

The original High School Exit Examination Standards Panel recommended that the reading and writing sections of the ELA portion of the spring 2001 CAHSEE be assigned equal weights (fifty percent each) in the calculation of each student's total ELA scale score. The Panel also recommended that the writing applications (essays) be weighted 30 percent and the multiple-choice questions be weighted 70 percent of each student's total ELA scale score.

Under the revised ELA blueprint described earlier in this document, the reading and writing weights will remain equal. The single writing application (essay) will now be weighted 20 percent and the multiple-choice questions will be weighted 80 percent of the student's total ELA scale score. To accomplish this technically in terms of the raw to scale score conversion, the following procedures were used:

1. The reading and writing multiple-choice questions are each weighted one point:
 $72 \times 1.0 = 72$.
2. The weight of 4.5 is applied to the essay. The maximum score on the essay is four, therefore the weight is multiplied by four: 4.5×4 possible score points = 18.
3. The sum of the multiple-choice questions and weighted essay score is rounded to the nearest whole number. The weighted raw score is transformed to the ELA scale score.

The sum of steps 1 and 2 represent the range of the weighted ELA raw score, that is, 90. For the February 2004 administration, a student needed a weighted ELA raw score of 54 to achieve a minimum passing score of 350. Over time, different conversion tables may equate different weighted ELA raw scores to the minimum passing CAHSEE scale score due to variations in test difficulty across forms.

For some administrations, one or more multiple-choice questions may be removed from scoring. If this occurs for ELA, the weight assigned to the multiple-choice question score is adjusted so the product of the weight and the score remains 72. For example, if one multiple-choice question is removed, the remaining 71 multiple-choice questions are multiplied by 1.0141.

Baseline Conversions

After each administration of the CAHSEE, a link to the score conversion table for that administration will be added to this Web site. Beginning with the February 2004 administration, a new reporting scale for the CAHSEE was established. The February 2004 CAHSEE serves as the baseline to which all future forms will be equated. For example, the math raw score of 43 questions answered correctly on the February 2004 test converts to the 350 scale score that reflects the minimum passing performance approved by the State Board of Education.

The CAHSEE was designed to be an accurate measure of achievement in the score range from about 300 to 400 (350 being the passing score). This accuracy around the passing score is sufficient to equate test scores on one test form to another correctly and to reasonably interpret the “distance to passing.”

Use of the CAHSEE for No Child Left Behind (NCLB) Reporting

As part of the reporting requirements for the NCLB Act, cut scores defining “proficient” and “advanced” performance on the CAHSEE were defined on the old CAHSEE score scale. The use of the revised CAHSEE blueprints beginning in February 2004 required that these cut scores be reset on the CAHSEE scale. Following the scaling of CAHSEE to the new score scale based on the February 2004 administration, the NCLB proficient cut score was set at 380 for both ELA and math. The advanced cut scores were set at different places on the ELA and math scales: for ELA the advanced cut score was set at 403, and for math the advanced cut score was set at 422. These values will be used to classify 10th grade students taking CAHSEE into the “proficient and above” category as part of California’s assessment of Adequate Yearly Progress (AYP).

The new value of 380 on Math is equivalent to the old NCLB proficient cut score of 373 on the old CAHSEE Math scale and the new value of 380 on ELA is equivalent to the old NCLB proficient cut score of 387 on the old CAHSEE ELA scale. It is not true that the 380 on the new Math scale represents a more difficult standard on the new test than the 373 did on the old scale, nor is it true that the new value of 380 on ELA represents an easier standard than the 387 did on the old scale. There should not be much difference in the percentages of students at or above the proficient level this year compared with last year.

CAHSEE Reporting Strands

In addition to total scores in ELA and math, CAHSEE also reports the number of questions and percent correct for specified reporting strands within each subject area (e.g., number sense or reading comprehension). Reporting strands can help teachers and instructional leaders pinpoint areas of student strengths and weaknesses. However, reporting strands should be interpreted cautiously, and two very important limitations of reporting strands should always be kept in mind:

1. Reporting strands are based on different numbers of questions and, in some cases, the number of questions that makes up a reporting strand may be quite small. The smaller number of questions results in scores that are less accurate than the overall test scores.
2. Reporting strand scores are reported in terms of number correct and percent correct. The difficulty of the questions tested for each strand may vary from one administration to the next. This variability is adjusted in the scale score, but not in the number and percent correct. Comparison by strands across administrations by number and percent correct may be imprecise.

Tables 1 and 2 present the strands within each portion of the CAHSEE and the number of questions for which percent correct is calculated.

Table 1: CAHSEE ELA Reporting Strands

ELA	Number of Questions
Word Analysis	7
Reading Comprehension	18
Literary Responses & Analysis	20
Writing Strategies	12
Writing Conventions	15
Writing Applications	1

Table 2: CAHSEE Math Reporting Strands

Math	Number of Questions
Probability & Statistics	13
Number Sense	17
Algebra & Functions	20
Measurement & Geometry	18
Algebra I	12
Mathematical Reasoning	(8*)

*Each Mathematical Reasoning question is also linked to one other strand. For reporting purposes only, these questions are counted within the other reporting strands to which the question is linked.

A useful benchmark for interpreting strand scores is the performance on the strand for students who scored exactly at the passing score, 350, on the CAHSEE. For each CAHSEE administration, the average percent correct scores for students who scored exactly at the passing score is calculated for each CAHSEE content area. Caution should be used in making these comparisons when the strand scores are based on relatively few questions (e.g., fewer than 15).