

Chapter 8 End-of-Course Assessments



Overview

EOC Participation Requirements

Test Development

Test Administrations

Testing Accommodations

Scores and Reports

Standard Setting

Scaling

Equating

Reliability

Validity

Measures of Student Progress

Sampling

Overview

On December 16, 2005, Governor Rick Perry issued Executive Order RP53, which called for increased college readiness programs in Texas public schools and authorized “the development of a series of voluntary end-of-course assessments in Science, Mathematics, and other subjects, currently assessed by the 11th grade Texas Assessment of Knowledge and Skills, to measure student performance....” As a result of Executive Order RP53, the Texas Education Agency’s (TEA’s) Student Assessment Division began the development of end-of-course (EOC) assessments in geometry, biology, chemistry, physics, and U.S. history and reestablished the development of Algebra I.

In May 2007, the Texas Legislature enacted Senate Bill (SB) 1031, expanding the role of the EOC assessment program. The bill phased out the Texas Assessment of Knowledge and Skills (TAKS) assessments for grades 9–11 and replaced them with the EOC assessments as a component of the new high school graduation requirements, beginning with the incoming freshman class of 2011–2012. The bill required the development of six additional EOC assessments:

- Algebra II
- English I
- English II



- English III
- world geography
- world history

Additionally, SB 1031 required that EOC assessments include items to measure college readiness. Performance at the highest cut score will indicate a strong application of knowledge and skills, and will indicate college readiness for Algebra II and English III. It will indicate advanced course readiness for Algebra I, English I, and English II, and it will indicate advanced performance for the remaining courses.

The high school, grade-based testing represented by TAKS will be replaced with the course-based EOC assessments in Algebra I, geometry, Algebra II, world geography, world history, U.S. history, biology, chemistry, physics, and English I, II, and III under the new STAAR testing program beginning in spring 2012.

In the process of implementing HB 3, enacted in June 2009, TEA will set the standards for the STAAR EOC assessments prior to the first mandatory administrations in 2011–2012 for graduation purposes. The standards will be externally validated by means of several studies that will ensure the rigor and alignment of the STAAR EOC assessments.

TEA’s Student Assessment Division is implementing the new legislation. Table 47 reflects the schedule used to field-test and implement the 12 EOC assessments over the course of several years.

Table 47. EOC Assessments—Implementation Plan

EOC Assessment	Spring 2007	Spring 2008	Spring 2009	Spring 2010	Spring 2011	Spring 2012	Spring 2013
Algebra I	Operational	→→→	→→→	→→→	→→→	→→→	→→→
Geometry	Field Test	Operational	→→→	→→→	→→→	→→→	→→→
Biology	Field Test	Operational	→→→	→→→	→→→	→→→	→→→
Chemistry		Field Test	Operational	→→→	→→→	→→→	→→→
US History		Field Test	Operational	→→→	→→→	→→→	→→→
Physics			Field Test	Operational	→→→	→→→	→→→
World Geography			Field Test	Operational	→→→	→→→	→→→
English I				Field test	Operational	→→→	→→→
Algebra II				Field test	Operational	→→→	→→→
English II					Field test	Operational	→→→
World History					Field test	Operational	→→→
English III					Field test	Operational	→→→



EOC Participation Requirements

In 2010–2011, any student enrolled in and completing a course in the spring semester for which an EOC assessment was offered was eligible to participate in testing, regardless of grade level. In the 2010–2011 school year, mandatory sampling occurred to support the development of the English II, English III, and world history tests and to support studies being conducted in Algebra I, geometry, Algebra II, biology, chemistry, physics, English I, world geography, and U.S. history. In addition to the sampling, participation in operational tests—those already implemented—was voluntary in 2010–2011 for each district. If a district chose to participate in testing, it had the flexibility to select participation on a district, campus, and individual student basis.

Test Development

Maintaining a student assessment system of the highest quality involves completing a set of tasks during the test development process. The procedures described in [chapter 2](#) outline the steps used to develop a framework for each EOC assessment and explain the ongoing development. TEA involves educators at each step of the development process because an equitable and accurate measure of learning can be achieved only if development is a shared responsibility.

Recent EOC development activities are summarized below.

- Using the established test development processes, items were developed and field-tested in the following courses: Algebra I, geometry, Algebra II, biology, chemistry, physics, English I, English II, English III, world geography, world history, and U.S. history.
- With the input of national- and state- level writing experts, scoring rubrics were developed to guide the scoring of the short answer reading questions and the written compositions for the English I, English II, and English III assessments. Following the application of the draft rubrics in scoring field test responses, TEA convened rubric validation committees composed of Texas writing educators to review the field test results and validate the rubrics. Rubrics for reading short answer, expository, and literary writing were validated in summer 2010. Rubrics for persuasive and analytical writing were validated in summer 2011.
- In 2010 and 2011, TEA invited advisory groups composed of curriculum specialists, teachers, and professors to meet and provide input and guidance about which of the content standards eligible for assessment were critical for student success and should be emphasized on the assessments. From this input, TEA developed a set of readiness and supporting standards designed to focus the assessment at each grade and course, and to provide a vertical link between the assessments from grade to grade and course to course. From this, a draft set of assessable student expectations and a draft test blueprint were developed for the new State of Texas Assessments of Academic Readiness (STAAR) EOC assessments. The draft assessed curricula and test blueprints were subsequently approved and published.



- An additional design aspect of the new assessments is a focus on preparedness for success in subsequent grades and courses, and ultimately for postsecondary education and career. This focus is reflected in the development of items in 2010–2011 that have a higher level of cognitive complexity and that closely align with the cognitive complexity evident in the TEKS. Additional open-ended items that require students to derive an answer independently have been developed for science and mathematics courses. In social studies and science courses, items that measure process skills in context rather than in isolation have been developed. In reading, greater emphasis will be given to critical analysis rather than literal understanding. In writing, prompts have been developed to support expository, analytical, and persuasive writing.
- Test items were developed so they could be delivered in both paper and online test formats. A plan was created to establish the comparability of items delivered in paper mode to items delivered in online mode. The plan also sought to minimize formatting issues during item development, such as scrolling, that might create a different student experience when testing online as opposed to testing on paper. The following 2011 EOC assessments were delivered in both paper and online modes: Algebra I, physics, English II, English III, world history, and U.S. history. A comparability study was conducted to identify field-test items that performed differently between the two modes.

Test Administrations

Each EOC assessment measures a student’s mastery of the TEKS for that specific course. Because the assessments are designed to be administered at the end of the course, the majority of the student expectations are eligible for testing. Students are given an EOC assessment during the published testing window and upon completing the course of study.

All 2010–2011 EOC operational tests and field tests were offered as online and paper administrations.

Further information about the online system, including an overview of the system, information on delivery and reporting, and a list of frequently asked questions, is available in the [Texas Assessment Management System](#).

Table 48. EOC Assessments Administered Online and On Paper in 2010–2011

Test Administration	Online Tests Administered	Paper Tests Administered
Field-Test Administrations		
Spring 2011 English II Field Test	13,658	44,089
Spring 2011 English III Field Test	12,563	44,198
Spring 2011 World History Field Test	13,050	24,580
Total	39,271	112,867
Operational Administrations		
Spring 2011 Algebra I	117,397	46,658
Spring 2011 Geometry	82,641	23,005
Spring 2011 Algebra II	28,380	45,334
Spring 2011 Biology	50,820	83,172
Spring 2011 Chemistry	29,256	60,723
Spring 2011 Physics	23,776	148,855
Spring 2011 English I	26,938	100,290
Spring 2011 World Geography	48,632	70,559
Spring 2011 U.S. History	23,858	71,821
Total 2011 EOC Assessments	431,698	656,412



Testing Accommodations

For students who met the eligibility criteria, an oral administration of the Algebra I, geometry, Algebra II, biology, chemistry, physics, world geography, and U.S. history EOC assessments was allowed during spring 2011. An oral administration was not available for the English I assessment or the field tests in English II, English III, or world history. Directions for test administrators conducting an oral administration were included in the test administrator manuals.

Accommodation Request Forms were not required for EOC assessments in the 2010–2011 school year; the use of accommodations was determined at the local level.

Scores and Reports

The various reports available for each EOC assessment are described in this section.

Description of Scores

For a detailed description of how test scores are derived, refer to [chapter 2](#).

**RAW SCORE**

The raw score is the number of items answered correctly on an EOC assessment (for example, geometry). By itself the raw score has limited utility. It can be interpreted only in reference to the total number of items on a test, and raw scores should not be compared across tests or administrations.

SCALE SCORE

Because the EOC assessments do not yet have established standards, the EOC Confidential Student Reports (CSR) provided a student's raw scores but not scale scores. Scale scores for all STAAR EOC assessments will be established after standards are set in February 2012. Refer to [chapter 3](#) for detailed information about scaling.

Report Formats

Two types of reports are typically provided for the various testing programs: standard and optional. Standard reports are provided automatically to districts. Information contained in standard reports is sufficient to satisfy mandatory reporting requirements. The EOC assessment program did not provide optional reports for 2010–2011 because the assessment is currently voluntary and not part of the reporting for accountability purposes.

Reports that include “Confidential” in the title contain student-level results. These reports are available for authorized users. All other reports present test results in an aggregated format and are considered public information.

Standard Reports

The standard reports available for the EOC assessment program include the Confidential Student Report (CSR) and Confidential List of Student Results (CLSR). These reports were provided for the operational tests within a 24-hour timeframe for online administrations, and on a seven-day turnaround for paper test administrations. In addition, Summary Reports and an EOC Data File (Confidential) are posted online for superintendent access following the close of the testing window.

Additional Reports in 2010–2011

To provide additional performance information to districts, three reports were designed for the operational EOC assessments for which performance standards have not yet been established (Algebra I, geometry, Algebra II, biology, chemistry, physics, English I, world geography and U.S. history). These new reports were provided at the district and campus levels and contained aggregated information about students for whom an online record or a paper answer document were submitted. A separate report was provided for each EOC assessment.

The reports for 2010–2011 included the Raw Score Frequency Distribution—All Students, Raw Score Distribution by Reporting Category Summary—All Students, and Cumulative Raw Score Frequency Distribution—All Students.

Standard Setting

Performance standards for all 12 assessments will be set in spring 2012 prior to the first mandatory administrations for graduation purposes in the 2011–2012 school year. Planning for setting performance standards on all 12 STAAR EOC assessments continued during the 2010–2011 school year. A process and timeline for setting the performance standards was established and presented to the Texas Technical Advisory Committee (TTAC) for its feedback. According to the current plan, the performance standards will be set prior to the first high-stakes STAAR EOC administration in spring 2012. Research studies, as mandated in Texas Education Code, §39.0242, will provide information for the standard-setting process. Data collection for these studies also took place during the 2010–2011 school year. This was done through the mandatory testing of sampled campuses for all EOC assessments.

Refer to [chapter 3](#) for detailed information about standard setting.





Scaling

As with many other tests in the Texas assessment program, the EOC assessment program uses the Rasch Partial-Credit Model (RPCM) to place test items on the same scale across administrations for a given EOC assessment. Once performance standards have been set for an assessment, its initial scale is then transformed to a more user-friendly metric to facilitate interpretation of the test scores. Details of the RPCM scaling method used in Texas are provided in [chapter 3](#).

Scale Score

Refer to [chapter 3](#) for detailed information about scale scores. TEA will establish performance standards for all STAAR EOC assessments during standard setting meetings in February 2012. Using the RPCM scaling procedures described in [chapter 3](#), a unique scale transformation will be developed for the STAAR EOC assessments so the resulting set of scale scores will have the panel-recommended Level II: Satisfactory Academic Performance and Level III: Advanced Academic Performance cut scores. The linear transformation of the underlying Rasch proficiency level estimate is as follows:

$$SS_j = (\theta_j \times T1) + T2$$

where SS_j is the scale score for student j , θ_j is the Rasch partial credit model proficiency level estimate for student j , and $T1$ and $T2$ are scale score transformation constants that establish the scale score system.

Raw Score

In 2011 the EOC assessment program reported results for several assessments using raw scores. This score is the number of items that a student answers correctly. Because no performance standards have been set for EOC assessments, only raw scores for these tests were reported.



Equating

During the 2010–2011 school year, field-test equating was conducted in the EOC assessment program. In addition, comparability analyses were conducted for the following EOC assessments that were administered to students both online and on paper in spring 2011: Algebra I, physics, English II, English III, world history, and U. S. history. English II, English III, and world history were first-time stand-alone field tests in spring 2011. Refer to [chapter 3](#) for detailed information about equating and comparability analyses.

Field-Test Equating

To replenish the item bank as new tests are created each year, newly developed items must be field-tested and equated to the scale of the original form. During 2010–2011, field-test equating was conducted for all EOC assessments. The field-test equating process for each test depended on the model in which field-test items were placed on the test form—either through embedded field testing or stand-alone field testing.

Because nine of the EOC assessments (Algebra I, geometry, Algebra II, biology, chemistry, physics, English I, world geography, and U.S. history) were operational assessments with embedded field-test items, live test items common to each form of the test were used to place the embedded field-test items onto the baseline scale of each EOC assessment. [Chapter 3](#) provides more details about the procedures for equating embedded field-test items.

For English II, English III, and world history, where no operational test forms existed, newly constructed items were placed in stand-alone field-test forms. For these three assessments, a set of linking items common across all field-test forms was used to equate the field-test items to each other. The specific procedure for equating stand-alone field-test items is described in further details in [chapter 3](#).

Comparability Analyses

In spring 2011, comparability analyses were conducted for six EOC assessments. These assessments were administered in both online and paper modes. Of these assessments, English II, English III, and world history were stand-alone field tests, while Algebra I, physics, and U.S. history were operational tests. In order to evaluate whether the mode of presentation affected item difficulty, comparability analyses were conducted at the item level for all of the items on these assessments. For the English II and English III EOC assessments, the comparability analyses were conducted separately for the reading and writing components of the assessments.



Two types of item-level analyses were conducted across testing modes: (a) comparison of item p-values, and (b) comparison of Rasch item difficulties. The results of these comparability studies indicated that, for all of the assessments except English II and English III, very few items showed differences between the online and paper versions of the tests. The percentage of items that demonstrated a mode effect for Algebra I, physics, world history, and U.S. history were 1.0%, 1.9%, 0.9%, and 5.0%, respectively. For English II, more reading items were identified as showing a mode effect than writing items. The percentages of items that demonstrated a mode effect for English II reading was 8.6% as compared to 4.9% for English II writing. For English III, more writing items were identified as showing a mode effect than reading items. The percentage of items that demonstrated a mode effect for English III writing was 16.7% as compared to 12.2% for English III reading.

The specific procedures used to evaluate comparability as well as the detailed results of the analyses are available in the “[2011 End-of-Course Comparability Study Report](#)” in the TEA technical report series. The results from the comparability study were used in field-test equating and will be used in future test construction. Refer to [chapter 3](#) for detailed information about comparability analyses.

Reliability

During the 2010–2011 school year, reliability for the nine operational EOC assessments was estimated through several reliability indices, including internal consistency and classical standard error of measurement. Refer to [chapter 3](#) for detailed information about reliability.

Internal Consistency

The Kuder-Richardson Formula 20 (KR20) was used to calculate the reliability estimates for all EOC assessments. As a general rule, reliability coefficients from 0.70 to 0.79 are considered adequate, 0.80 to 0.89 are considered good, and above 0.90 are considered excellent. However, appropriate levels of reliability depend on how an assessment is being used. For the EOC assessment program, the internal consistency estimates range from 0.88 to 0.94. The internal consistency estimates for the EOC assessments are available in [Appendix F](#). Refer to [chapter 3](#) for detailed information about internal consistency.

Classical Standard Error of Measurement

For the EOC assessments, Standard Error of Measurement (SEM) values range from 2.32 to 3.70. The SEM values for the EOC assessments are provided in [Appendix F](#). Refer to [chapter 3](#) for detailed information about standard error of measurement.



Validity

The sections that follow describe how validity evidence was collected for the EOC assessments in 2010–2011. Refer to [chapter 3](#) for detailed information about validity.

Evidence Based on Test Content

Evidence based on test content is information that shows the relationship between content of the test and the test constructs that are intended to be measured by the test. The EOC assessments have been developed to align with the content defined by the TEKS. Content validity evidence has been collected at all stages of the test development process.

Nationally established test development processes for the Texas assessment program were followed in developing the EOC assessments to support the use of EOC scores in making inferences about students' knowledge and understanding of the TEKS.

The following activities took place during the 2010–2011 school year to support the content validity of EOC assessments.

RELATIONSHIP TO THE STATEWIDE CURRICULUM

As part of the transition to a high-stakes graduation program in 2012, teachers, curriculum specialists, test development specialists, college educators, and TEA staff members worked together in advisory committees to identify appropriate assessment objectives for new STAAR EOC assessments in Algebra I, geometry, Algebra II, biology, chemistry, physics, English I, English II, English III, world geography, world history, and U.S. history. The input of the advisory committees is reflected in the assessed curricula and test blueprints.

Early in the development process, prototype items were developed for the new EOC assessments. As part of the item development process, these prototypes were reviewed by advisory committees and TEA staff to identify how well they measured the student expectations to which they were aligned and to provide information for item-development guidelines and test-item types. Item-development guidelines continue to be refined through the test development process, as STAAR item review committees share their feedback about how the student expectations can be effectively assessed.

EDUCATOR INPUT

Following item development for each EOC assessment, committees of Texas educators met to review test items and confirm that each item appropriately measured the TEKS to which it is aligned. The committees also reviewed and edited the items for content and bias. These reviews occurred for Algebra I, geometry, Algebra II, biology, chemistry, physics, English I, English II, English III, world geography, world history, and U.S. history. The educator review committees revised and edited items, as appropriate, before field testing occurred.



As EOC assessments transition from the low-stakes environment in which they were developed and administered beginning in 2005 to the high-stakes graduation measure under which they are now developed and will be administered in 2012, evidence of curricular and instructional validity is being gathered. The evidence will determine (1) if there is a match between the tested content drawn from the state curriculum standards and classroom curricular materials, and (2) if there is a match between the tested content and what is being taught in the classrooms. This evidence has been collected as part of the item-by-item judgments made by educators during item review.

DATA REVIEW

Following field testing, TEA conducted a review of the performance data of each item and made a judgment about whether the item appropriately measured the construct and was eligible for placement on an operational test.

TEST DEVELOPER INPUT

Item writers and reviewers followed test development guidelines that informed how the content of the assessed TEKS should be measured. At each stage of development, writers and reviewers verified the alignment of the test items with the assessed objectives.

TEST EXPERT INPUT

TEA, in conjunction with Pearson, receives ongoing input from the Texas Technical Advisory Committee (TTAC), a panel of national testing experts, related to plans for collecting validity evidence for Texas testing programs, including EOC assessments.

In February 2009, the TTAC provided input on validity studies designed to examine the relationship between students' EOC scores with performance on college readiness tests such as the ACT and SAT, and longitudinal studies that examine the relationship between the college-readiness performance standard and performance in college courses in the same content area. In February 2010, the TTAC provided input on conducting studies that link performance on EOC assessments with commonly taken college placement tests, such as ACCUPLACER, COMPASS, and THEA. In March 2011, the TTAC provided input on conducting studies that link college students' performance on EOC assessments with their performance in entry-level college classes.



Evidence Based on the Response Process

Response processes refer to the cognitive behaviors required to respond to a test item. For the EOC assessments, Texas collects validity evidence based on response process through the different item types on each test and the mode in which tests were administered.

Multiple-choice was the primary item type to which students were asked to respond on all EOC assessments. One of the reasons this type of item was used is because it most closely resembles what students typically experience in classroom testing. The multiple-choice items were developed so students were required to recall and apply what they had learned about the course, thereby supporting an accurate measurement of the construct being assessed. For the Algebra I, geometry, Algebra II, chemistry, and physics assessments, griddable items that required the students to determine a numerical answer and then grid in (or type in, for online tests) their answers, were also administered. These item types facilitated the assessment of the students' knowledge and skills at an even deeper level by requiring students to generate answers independently without being influenced by answer choices provided with questions. These skills were essential in assessing the construct in the mathematics and science content areas. In addition, passage-based items were used for the English I, English II, and English III assessments. Passage-based items are a group of multiple-choice and short-answer items associated with a common stimulus, such as literary or informational reading selections. The use of passage-based items requires students to apply their knowledge and skills within the context introduced by the stimulus in order to respond correctly to the items. This skill is important in assessing the construct in reading and writing.

Evidence Based on Internal Structure

Texas collects evidence that shows the relationship among test questions and test objectives to demonstrate that the parts of a test conform to the test construct. When tests are designed to measure a single construct, the internal components of the test should exhibit a high level of homogeneity, which can be evaluated in terms of the internal consistency estimates of reliability. Refer to the “[Reliability](#)” section for descriptions and estimates of internal consistency in the EOC assessments during 2010–2011.



Evidence Based on Relationship to Other Variables

Another source of validity evidence is the relationship between test performance and performance on some other measure, sometimes called criterion-related validity. Planning for validity studies that will correlate student performance between EOC assessments in the mathematics and English content areas occurred during the 2010–2011 school year. In addition, studies that evaluate the relationship between the EOC assessments and other external measures such as the high school TAKS assessments; college course performance; and ACCUPLACER, THEA, AP, IB, SAT, PSAT, and PLAN assessments are being planned. Results for these validity studies can be used to provide criterion-related validity evidence for the STAAR EOC assessments and may also be used in the standard-setting process to support the validity of the performance standards.

Evidence Based on the Consequences of Testing

Another way to provide validity evidence is by documenting the intended and unintended consequences of administering an assessment. Validity evidence indicating the impact of EOC testing on students was collected through a question on the online survey at the end of each operational EOC assessment (Algebra I, geometry, biology, chemistry, and U.S. history) during the 2008–2009 school year. The survey was voluntary for students and asked them the question “For what percent of your course grade does this test count?” The survey results showed that some students (about 12–21% across the five tests) responded that the EOC test scores did affect their course grades. This provided evidence that the results from the EOC assessments had academic consequences for some students in 2008–2009.

Measures of Student Progress

Student growth and projection measures track a student’s performance across time. Growth measures track student performance from year to year, while projection measures use current student performance to predict future performance. No student growth or projection measures are currently used with the EOC assessment program. During the 2010–2011 school year, TEA and Pearson started planning for the development of student growth and projection measures in the STAAR EOC assessments to meet federal and state accountability requirements.



Sampling

During the 2010–2011 school year, sampling was used to select campuses to participate in an assessment's first-time stand-alone field test or to collect data for research studies that will provide information for the STAAR EOC standard-setting process in February 2012. Stand-alone field tests were conducted for the three new EOC assessments: English II, English III, and world history. Data for standard-setting research studies were collected for nine operational EOC assessments: Algebra I, geometry, Algebra II, biology, chemistry, physics, English I, world geography, and U.S. history. The sampling for Algebra I, physics, English II, English III, world history, and U.S. history was conducted in both testing modes (i.e., on paper and online), while the sampling for Algebra II, biology, chemistry, English I, and world geography was conducted in the paper mode only. Geometry was sampled online only. Campuses were specifically assigned to be part of each of the EOC samples. Campuses selected for one of the dual-mode EOC assessments were specifically assigned to test in one of the two modes: paper or online. Participation in the EOC assessments by all selected campuses was mandatory. Campuses not selected to participate in one of the EOC assessments could volunteer to test in one of the available modes.

EOC Sampling Process

A stratified sampling design was used for the EOC assessment program in which the campus was the sampling unit, but the student was the observation unit. Each campus was classified into one of five strata based on its campus size, or estimated student count for each EOC assessment. Because the campus was the sampling unit, it was necessary to obtain the student course enrollment from each campus as an estimate of the number of students that would participate in the corresponding EOC assessment. The estimated student counts for each campus were based on the number of students who were enrolled in each of the sampled courses in fall 2009.

The following factors were considered in determining each EOC sample for 2011:

- The sample was chosen to be representative of the overall population of Texas high school students taking the course in terms of ethnic composition and campus size.
- The sample was selected to include a minimum of 280 students per form from each major ethnic and gender subgroup (i.e., African American, Hispanic, white, male, and female groups).
 - Campuses were not assigned to more than four total EOC assessments, with each English assessment counting as two assessments.
 - For any online EOC assessments, each sampled campus was required to test all of its enrolled students in the course, regardless of grade, up to a maximum of 350 students during the testing window.
 - For any paper EOC assessments, each selected campus was required to test all enrolled students in the course, regardless of grade, during the window.
 - Campuses were not assigned to more than two online EOC assessments.



- To reduce the district and campus field-testing burden, eligibility criteria were used to eliminate the following campuses from the sample:
 - Campuses with fewer than 15 students enrolled in the course for the EOC assessment.
 - Campuses that are part of the Juvenile Justice Alternative Education Program (JJAEP), Disciplinary Alternative Education Programs (DAEP), or Texas Youth Commission (TYC).

The process for selecting campuses to be included in each EOC sample was as follows:

1. All eligible campuses were divided into five even-sized strata based on campus size (i.e., strata 1, 2, 3, 4, and 5). If a number of campuses of equal size appeared around the threshold between strata, the placement in the upper or lower stratum was done randomly.
2. Campuses were sorted randomly within each stratum.
3. One campus was randomly selected from each stratum. Each campus was chosen into the sample in ascending and descending order of strata (e.g., 5-4-3-2-1-1-2-3-4-5-5-4-3-2-1-...). For example, the first campus was chosen from stratum 5, the second campus from stratum 4, the third campus from stratum 3, the fourth campus from stratum 2, the fifth campus from stratum 1, etc. Thus, one campus was selected in one stratum first before moving to the next stratum.
4. The number of students in the sample was evaluated relative to the target total number of students after the campus had been selected. Step 3 was repeated until the target number of students was reached.
5. Once the final sample was determined, it was regenerated using the appropriate random number seed so additional detailed output descriptive statistics for this sample could be generated.

The final sample was determined after evaluating four key elements: fit to statewide ethnic percentages, number of campuses, number of students, and distribution of campus size strata within the sample. A summary of the number of campuses and students selected for the 2010–2011 EOC samples is provided in Table 49.

Table 49. Sampling Summary for 2011 EOC Assessments

EOC Assessment	Online		Paper	
	Number of Sampled Campuses	Number of Expected Students at the Time of Sampling	Number of Sampled Campuses	Number of Expected Students at the Time of Sampling
Algebra I	285	40,252	49	8,329
Geometry	407	54,618	0	0
Algebra II	0	0	200	40,145
Biology	0	0	271	54,237
Chemistry	0	0	235	54,072
Physics*	84	10,580	761	101,527
World Geography	0	0	171	41,110
World History	39	6,095	93	19,425
US History	47	6,308	476	75,215
English I	0	0	329	80,010
English II	74	11,429	208	43,612
English III	84	11,693	249	43,713

* All campuses offering the physics course participated in the Physics EOC Assessment.



