Chapter 6 STAAR Alternate

Overview

STAAR Alternate is an alternate assessment based on alternate academic achievement standards and is designed for students with the most significant cognitive disabilities receiving special education services who also meet the participation requirements for STAAR Alternate. This assessment was designed to meet federal requirements mandated under No Child Left Behind (NCLB). STAAR Alternate is not a traditional paper or multiple-choice test. Instead, it involves test administrators observing students as they complete standardized, state-developed assessment tasks that link to the grade-level Texas Essential Knowledge and Skills (TEKS). Teachers then evaluate student performance based on the components of the STAAR Alternate rubric and submit student results through the Assessment Management System.

The assessments included in STAAR Alternate are shown in Table 6.1. STAAR Alternate was administered during the window of January 7, 2013, through April 19, 2013 for all subject areas and grades.

Table 6.1. 2012–2013 STAAR Alternate Assessments

2012–2013 STAAR Alternate Assessments					
Grade	**Assessed Subject Area/Course**				
3	Mathematics	Reading			
4	Mathematics	Reading	Writing		
5	Mathematics	Reading		Science	
6	Mathematics	Reading			
7	Mathematics	Reading	Writing		
8	Mathematics	Reading		Science	Social Studies
High School	Algebra I Geometry	English I English II English III		Biology	World Geography World History U.S. History

Participation Requirements

STAAR Alternate has specific participation requirements that an admission, review and dismissal (ARD) committee must carefully consider when recommending these assessments for students receiving special education services. Prior to reviewing the eligibility criteria for STAAR Alternate, the ARD committee must understand all assessment options, including the characteristics of each assessment and the potential implications of each assessment choice.

If STAAR Alternate is being considered, the ARD committee must review the five criteria below and indicate whether the description is applicable to the student. For a student to be eligible to participate in STAAR Alternate, the answer to all five questions below must be "Yes." If the answer to any question is "No," the student is not eligible to participate in STAAR Alternate and must participate in one of the other statewide assessments. Each "Yes" answer must be justified by evidence that the student meets the criterion.

1. **Does the student have a significant cognitive disability?** A significant cognitive disability is determined by the ARD committee and must be based on evaluation information performed by a qualified evaluation team. The significant cognitive disability must affect the student's intellectual potential and be documented as such in the student's individualized education program (IEP). A student with a significant cognitive disability has limited potential to reach grade-level expectations; whereas,

a student with a learning disability has the potential to reach grade-level expectations, but has difficulty doing so due to his or her disability.

2. **Does the student require specialized supports to access the grade-level curriculum and environment?** Federal regulations mandate that all students have access to and be assessed on grade-level curriculum. To access the state-mandated grade-level or course curriculum, the Texas Essential Knowledge and Skills (TEKS), a student with a significant cognitive disability needs specialized academic instruction as well as support throughout the day in areas such as expressing his or her needs, getting from place to place, eating lunch, negotiating social situations, and/or taking care of personal needs.

3. **Does the student require intensive, individualized instruction in a variety of instructional settings?** The student needs specialized academic instruction and techniques over a period of time to ensure that he or she can learn, retain information, and transfer skills to other settings.

4. **Does the student access and participate in the grade-level TEKS through prerequisite skills?** Access to the grade-level curriculum is mandated by the federal government. A student with a significant cognitive disability requires access to the TEKS through prerequisite skills that are linked to the grade-level curriculum.

5. **Does the student primarily demonstrate knowledge and skills through performance tasks?** The student may be able to perform some literacy skills (e.g., tracing words, copying spelling words, completing simple worksheets, writing simple phrases or sentences). However, the student is typically evaluated by methods other than paper and pencil, such as observation of student performance while the student manipulates items, verbalizes responses, eye gazes, or activates an augmentative communication device. A one-day, multiple-choice test would not be an appropriate assessment format to effectively show what the student has learned.

Additional information and resources are available for use by the ARD committee in making assessment decisions regarding students receiving special education services. This information can be found at the STAAR Alternate Participation Requirements page on the TEA's Student Assessment Division website.

Testing Requirements for Graduation

With the passage of House Bill (HB) 3, the relationship between high school courses and participation in the STAAR Alternate EOC assessments is now linked to a student's graduation plan. HB 5 reduced the total number of tests to five: Algebra I, biology, U.S. history, English I and English II. However, the ARD committee makes final determinations on the graduation requirements for students receiving special education services and who are eligible to take STAAR Alternate.

Test Development

As much as possible, STAAR Alternate follows the same test development procedures as other STAAR assessments. However, the test development process does reflect the unique characteristics of STAAR Alternate, specifically its reliance upon performance-based assessment tasks and the needs of the STAAR Alternate population.

Assessment Content

Like other STAAR assessments, STAAR Alternate is linked to grade-level TEKS and student expectations for STAAR. To link the assessment to the content and expectations, TEA worked with experts in test development, special education, and content to develop curriculum frameworks and vertical alignment documents. The curriculum frameworks outline the grade-level TEKS and the associated prerequisite skills for each grade and subject area. The vertical alignment documents link skills and knowledge across grades within the same subject area. After the initial creation of the curriculum frameworks and vertical alignment documents, TEA sought additional input from the educator committees and the STAAR Alternate steering committee, which is a statewide advisory group that includes state experts, parents, advocacy group representatives, related service providers, administrators, and Texas regional Education Service Center (ESC) professionals. The curriculum frameworks and vertical alignment documents can be found at the STAAR Alternate Resources page of TEA's Student Assessment Division website.

The next step in developing STAAR Alternate was to generate essence statements that summarize the TEKS and student expectations and link the expectations to the prerequisite skills and assessment performance categories. Typically, each grade and subject area contains 10–20 essence statements. From these, four essence statements are identified for inclusion in the STAAR Alternate assessment each year. The 2012–2013 assessed essence statements were made available to teachers in spring 2012 to allow time for planning and developing standards-based individualized education programs (IEPs) for the following school year. The essence statements for all grades and subject areas can be found at the STAAR Alternate Essence Statements page of TEA's Student Assessment Division website.

Complexity Levels

Three assessment tasks of varying complexity levels are developed for each essence statement to allow for the accessibility and flexibility of the assessment for the diverse STAAR Alternate population. To establish the verbs that define the complexity levels for the assessment tasks, Bloom's work on learning taxonomies (Bloom, Englehart, Furst, Hill, & Krathwohl, 1956) was consulted. To develop the three complexity levels of the assessment tasks, Webb's depth of knowledge (Webb, 1997), Cook's extended depth of knowledge (Cook, 2008), and Browder and Flowers' depth of knowledge scales (Flowers, Wakeman, Browder, & Karvonen, 2007) were also referenced. Using a combination of cognitive scales, verbs were selected that define each complexity level and show how the student demonstrates knowledge. Each verb is specifically

defined and suggestions for possible ways the student can respond are provided in order to further standardize task implementation. The three complexity levels are described as follows.

Level 1: Beginning Awareness

Level 1 assessment tasks are the least complex and involve responding with knowledge at the beginning awareness level. Skills that students at this level are expected to demonstrate might include acknowledging features, responding to stimuli, participating in processes, exploring materials, or anticipating outcomes.

Level 2: Basic Recall

Level 2 assessment tasks are moderately complex and involve recalling or reciting information at a basic level. Skills that students at this level are expected to demonstrate might include identifying or sorting elements, assisting in procedures, choosing options, examining features, or matching or replicating components.

Level 3: Application

Level 3 assessment tasks are the most complex and involve applying knowledge beyond basic recall. Skills that students at this level are expected to demonstrate might include determining distinguishing features, organizing information, comparing components, generating ideas, making inferences, or justifying answers.

Assessment Task Criteria

In addition to the procedures outlined in chapter 2, "Building a High-Quality Assessment System," nationally accepted criteria provide guidance during the development of the STAAR Alternate assessment tasks. Specifically, the following criteria are directly referenced during development activities:

- Standard 3.6 of the Standards for Educational and Psychological Testing (AERA, APA, NCME, 1999), which states:

  > The type of items, the response formats, scoring procedures, and test administration procedures should be selected based on the purposes of the test, the domain to be measured, and the intended test takers. To the extent possible, test content should be chosen to ensure that intended inferences from test scores are equally valid for members of different groups of test takers. The test review process should include empirical analyses and, when appropriate, the use of expert judges to review items and response formats. The qualifications, relevant experiences, and demographic characteristics of expert judges should also be documented (p. 44).

- Universal design, with particular attention given to (1) students' response modes, allowing students to show what they know and can do; (2) differentiated supports and materials, allowing students to access the content of the assessment; and (3) multiple means of engagement to allow students more

time to complete the task, meaningful activities, and context (Center for Applied Special Technology, 2002). According to the principles of universal design, each item has precisely defined constructs, has maximum legibility, has maximum readability and comprehensibility, is amenable to accommodations, is accessible and non-biased, and takes into consideration special populations.

Review of Assessment Tasks

During development, educator committees meet to complete reviews of every STAAR Alternate assessment task. The committees are made up of educators from across Texas, specifically special education experts, special education classroom teachers (including teachers from the Texas School for the Blind and Visually Impaired and the Texas School for the Deaf), teachers of English language learners, and general education teachers.

The educator committees focus on the relationship between the grade-level content and the assessment tasks. Based on guidance from the National Alternate Assessment Center (2005), they consider the following questions in regard to each assessment task:

- Does the assessment task cover academic content?
- Does the assessment task reflect the grade-level curriculum?
- Does the assessment task access the grade-level STAAR reporting category and knowledge and skills statements?
- Is the assessment task meaningful to the student?
- Will the assessed skills be useful to the student in the immediate future?

In addition to these questions, educator committees are asked, “Is the assessment task free from bias on the basis of students’ personal characteristics, such as gender, ethnicity, or disability?”

Feedback from the educator committees is used to revise the STAAR Alternate assessment tasks as needed. The 2012–2013 assessment tasks were made available to teachers in fall 2012 to allow time for instruction prior to the assessment window.

Training

Because STAAR Alternate includes unique materials and requirements and the use of the Assessment Management System, it is essential that test administrators receive proper and effective training. TEA developed four web-based training modules to standardize training across the state and assist test administrators in effectively implementing STAAR Alternate. Personnel who plan to administer the STAAR Alternate assessment are required to complete each of the four modules and pass the associated qualification activities with a score of 80 percent or higher. The four modules and qualification activities must be successfully completed prior to any

STAAR Alternate administrations. Once completed, the teacher receives a qualification certificate confirming the teacher's status as a test administrator.

Test administrators have secure access to the online training modules on the training center website. From this site, they can access additional online training resources, track their training status, and access training certificates. In addition, the training center tracks and reports the completion and qualification status for each user on the four STAAR Alternate training modules, allowing for monitoring by district testing coordinators and TEA. This ensures that each test administrator is adequately trained and prepared to assess students taking STAAR Alternate.

The structure of the STAAR Alternate training modules follows the steps of test administration. The following section outlines the content of each of the four training modules.

Module 1: Selecting the Assessment Task

Module 1 provides an overview of the STAAR Alternate assessment, including the role of the ARD committee, the participation requirements, and the strategies for selecting assessment tasks, which is Step 1 of the test administration process.

Module 2: Implementing the Assessment Task

Module 2 explains how to make assessment tasks accessible at each complexity level by using appropriate presentation supports/materials and response modes, and how to use the resource documents for planning and implementing the assessment observation for Step 2 of the test administration process.

Module 3: Observing/Documenting Student Performance

Module 3 clarifies the difference between instruction and assessment, between cues and prompts, between primary and generalization observations, and explains how to prepare for and document the observation to complete Step 3 of the test administration process.

Module 4: Evaluating Student Performance

Module 4 provides information about the scoring rubric, the evaluation process for both the primary and generalization observations, and the appropriate maintenance of the documentation forms. This module provides practice using the Assessment Management System and includes a flowchart to assist in completing Step 4 of the test administration process.

After completion of the modules and qualification activities, test administrators are encouraged to reference training resources as needed. In 2012–2013 TEA also offered additional training via the Texas Education Telecommunication Network and PowerPoint presentations on TEA's website that can be downloaded and used for individual or group training sessions.

Test Administrations

More than 92,000 STAAR Alternate assessments were administered in 2012–2013 to approximately 30,000 students. Table 6.2 further describes the 2012–2013 STAAR Alternate administrations by grade and subject area.

Table 6.2. Students Tested in 2012–2013 STAAR Alternate Assessments

STAAR Alternate Assessments	Students Tested
Grade 3 reading	4,286
Grade 3 mathematics	4,286
Grade 4 reading	4,257
Grade 4 mathematics	4,258
Grade 4 writing	4,257
Grade 5 reading	4,062
Grade 5 mathematics	4,062
Grade 5 science	4,062
Grade 6 reading	3,940
Grade 6 mathematics	3,940
Grade 7 reading	3,496
Grade 7 mathematics	3,495
Grade 7 writing	3,493
Grade 8 reading	3,438
Grade 8 mathematics	3,439
Grade 8 science	3,440
Grade 8 social studies	3,438
English I	3,142
Algebra I	3,159
English II	2,967
Geometry	2,986
English III	2,732
Biology	3,370
U.S. history	2,817
World geography	3,036
World history	2,557

Administration Procedures

The STAAR Alternate assessment process is designed to mirror the instructional process for a student with a significant cognitive disability. The essence statements, upon which the 2012–2013 STAAR Alternate assessment tasks were based, were made available in spring 2012 so that they could be included in ARD committee meetings and other planning related to the 2012–2013 school year. The STAAR Alternate assessment tasks were made available at the beginning of the 2012–2013 school year so that teachers could provide related instruction and prepare students for the assessment.

The assessment administration window extended from January 7, 2013, to April 19, 2013, which allowed teachers ample time to select appropriate assessment tasks, determine appropriate implementation of the tasks, evaluate and document student performance, and enter results in the Assessment Management System. Because of the heterogeneity of the population of students who take STAAR Alternate, flexibility is built into the assessment and its administration, allowing teachers to shape the tasks to fit the individual needs of each student. Once test administrators have completed the required STAAR Alternate training, these test administrators then take the following steps as part of the administration of the assessment.

1. Select the assessment tasks appropriate for each student. One of three possible assessment tasks is selected for each of the four assessed essence statements. As part of this step, teachers are asked to instruct the student on the prerequisite skills associated with the task and document the materials and supports that are needed as part of instruction.

2. Implement the assessment task in a manner appropriate for each student. Using the information obtained during instruction, the test administrator then documents the supports and materials needed by the student to complete the task as well as the response mode used by the student. This documentation must be completed prior to the assessment observations.

3. Observe and document student performance. Each task is administered in the manner documented in Step 2. The test administrator then records student performance for each of the three predetermined criteria for each assessment task, including information about cueing and prompting. If applicable, the generalization observation is also conducted and documented.

4. Evaluate student performance within the Assessment Management System. Test administrators enter the results from the student observation into the Assessment Management System so that each student's performance can be scored. In addition, documentation forms must be completed and securely maintained.

If a student is unable to display any observable change in affect or movement due to either an ongoing medical condition or the severity of the student's disability, that student may receive a No Response Observed (NRO) score designation.

Test administrators are able to deliver the assessment and submit assessment results at any time during the STAAR Alternate assessment window.

Testing Accommodations

Students being assessed with STAAR Alternate can use the accommodations and supports that are routinely and successfully used as instructional accommodations. These accommodations become the specific materials and supports that are provided during the assessment observation and allow the student access to the task. After the observation begins, if the student needs additional assistance beyond the pre-planned supports, the test administrator might provide cues and prompts to continue the task; however, the use of cueing and prompting will affect the student's score and should only be provided after sufficient wait time in order to allow the student an opportunity to respond without the additional support.

The difference between a cue and a prompt is related to the degree of assistance provided to the student. A cue is a hint and does not lead the student to a direct answer. A prompt is more directive, as it takes the student step-by-step through the task, leading to a specific answer.

Because of the design of STAAR Alternate, linguistic accommodations are not necessary for limited English proficient (LEP)-exempt immigrants receiving special education services. The STAAR Alternate assessment can be administered using any language or other communication method routinely used by the student.

Additional information about accommodations can be found at the Accommodations Resources page of TEA's Student Assessment Division website.

Student Success Initiative

The Student Success Initiative (SSI) provides a system of academic support to help students achieve on grade level success in reading and mathematics. The SSI incorporates a grade advancement component adopted by the Texas Legislature in 1999. The instructional processes used with students who have significant cognitive disabilities serve as the basis of the STAAR Alternate assessments. Test administrators can observe student performance on several occasions if necessary, and provide remediation as needed throughout the academic year. Because of this relationship between STAAR Alternate and instructional practice, students who participate in these assessments are not subject to the SSI requirements. Each student's promotion is determined by that student's ARD committee rather than being based on STAAR Alternate scores.

Scores and Reports

Scoring STAAR Alternate Assessments

STAAR Alternate scoring is based on an application of the scoring rubric to the student performance evaluation information that the test administrators submit electronically.

SCORING RUBRIC

The STAAR Alternate Scoring Rubric outlines the way various components are used to compute each assessment task score. The STAAR Alternate scoring components are: (1) Demonstration of Skill, (2) Level of Support, and (3) Generalization of Skill. Task complexity level is used as an additional factor in determining students' scores by weighting the Demonstration of Skill component. The inclusion of complexity level in the scoring allows students who successfully complete tasks that have higher complexity levels to receive higher scores than students who successfully complete tasks that have lower complexity levels.

Table 6.3 outlines the scoring used for the primary observation of each assessment task.

Table 6.3. Scoring of the STAAR Alternate Primary Observation

Predetermined Criteria	Demonstration of Skill	Level of Support
	Did the student demonstrate the skill?	How did the student perform the skill?
1	Yes – 2 points No – 0 points Yes, but needed prompting – 0 points	Independently – 2 points Needed Cueing – 1 point Needed Prompting – 0 points N/A – 0 points
2	Yes – 2 points No – 0 points Yes, but needed prompting – 0 points	Independently – 2 points Needed Cueing – 1 point Needed Prompting – 0 points N/A – 0 points
3	Yes – 2 points No – 0 points Yes, but needed prompting – 0 points	Independently – 2 points Needed Cueing – 1 point Needed Prompting – 0 points N/A – 0 points
	Level 3 task weighted by 1.5 Level 2 task weighted by 1.2 Level 1 task weighted by 1.0	No weighting
Total Points Possible	**9 points**	**6 points**

Performance on the primary observation determines whether a student has the opportunity to generalize the skill. The student is eligible for Generalization of Skill if all the following criteria are met:

- The student is assessed with a complexity level 2 or 3 assessment task.
- The skill was successfully demonstrated for all three predetermined criteria.
- There was no prompting on any of the three predetermined criteria.

To receive points for the generalization observation, the student must be assessed using materials different from those that were used during the primary observation. A total of six points are possible for Generalization of Skill. The student receives two points for each predetermined criterion completed independently. The student receives one point for each predetermined criterion completed with cueing. The student receives zero points for each predetermined criterion completed with prompting or that was not completed.

Each assessment task score is computed by summing the Demonstration of Skill, Level of Support, and Generalization of Skill scores. Assessment task scores range from 0 to 21 points. Students' total scores are computed by summing the four assessment task scores and rounding to the nearest whole number. STAAR Alternate total scores range from 0 to 84 points.

Description of Scores

There are a variety of reports that show a student's performance on STAAR Alternate. The information below describes the types of scores given on reports and the types of reports available.

Raw Score

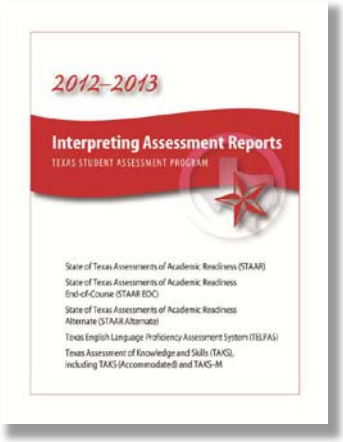
For STAAR Alternate, a raw score is based on the student's performance on the four assessment tasks and the points assigned to that performance based on the scoring rubric. Unlike other STAAR assessments, scaling is not used for STAAR Alternate (refer to the Scaling section of this chapter).

Additional Performance Information

STAAR Alternate reports include each of the component scores (Demonstration of Skill, Level of Support, and Generalization of Skill) that were assigned for each assessment task to provide as much information as possible about student performance. The assessment task scores, called reporting categories on reports, provide information about a student's relative strengths or weaknesses. Individual student test scores might be used in conjunction with other performance indicators to make decisions regarding student placement and instruction. While STAAR Alternate scores can contribute to these decisions, as much additional information as possible about each student should be taken into account.

Report Formats

Two types of reports are provided for the various testing programs: standard and optional. Standard reports are provided automatically to districts, and the information in the standard reports satisfies mandatory reporting requirements. To receive optional reports that detail student performance data in additional formats, a district must select the corresponding optional reports in the *Administration Details* screen in the Assessment Management System. Generally, districts are required to pay a nominal fee for each optional report requested.



For more information about scoring and reporting for STAAR Alternate, refer to the TEA publication "Interpreting Assessment Reports."

Use of Test Results

Reports of STAAR Alternate students are used in the following ways:

- helping parents monitor the progress their children make
- informing instructional planning for individual students
- reporting results to local school boards, school professionals, and the community
- evaluating programs, resources, and staffing patterns
- evaluating district effectiveness in accountability measures

Parent Brochure

TEA's Student Assessment Division produces the brochure Understanding Your Child's Confidential Student Report (CSR): A Guide for Parents (English). The brochure includes a sample CSR with explanations of each element of the report to help parents better understand their child's score report. Reporting categories for each subject area assessed with STAAR Alternate are summarized. The guide, developed in both English and Spanish, was distributed with 2012–2013 individual student results in January 2013.

Audits

TEA conducts periodic audits of the STAAR Alternate assessment as one means of collecting reliability and validity evidence. Audits enable the collection of information from test administrators and school districts that can be used to evaluate the training, administration, and scoring of STAAR Alternate. The first audit of STAAR Alternate was conducted in 2011–2012. Because STAAR Alternate is being redesigned as a result of HB 5, the audit was not conducted in 2012–2013 and no data were collected.

Performance Standards

Performance standards relate levels of test performance directly to what students are expected to learn as described in the statewide curriculum.

Performance Levels and Policy Definitions

For the STAAR Alternate assessments, the performance levels are:

- Level I: Developing Academic Performance
- Level II: Satisfactory Academic Performance
- Level III: Accomplished Academic Performance

More detailed descriptions, known as policy definitions, of these performance levels are given below.

Level I: Developing Academic Performance

Performance in this category indicates that students are insufficiently prepared for the assessment tasks at the next grade or course even with instructional supports for accessing the curriculum through prerequisite skills. They demonstrate insufficient knowledge and skills that are linked to content measured at this grade or course. Performance on the tested skills required cueing and prompting. Students in this category are in need of significant intervention in addition to continued supports to show progress for the assessment tasks at the next grade or course.

Level II: Satisfactory Academic Performance

Performance in this category indicates that students are sufficiently prepared for the assessment tasks in the next grade or course with instructional supports for accessing the curriculum through prerequisite skills. They demonstrate sufficient understanding of the knowledge and skills that are linked to content measured at this grade or course. Performance on the tested skills required little or no cueing. Students in this category have a reasonable likelihood of showing progress for the assessment tasks at the next grade or course with continued supports.

LEVEL III: ACCOMPLISHED ACADEMIC PERFORMANCE

Performance in this category indicates that students are well prepared for the assessments tasks in the next grade or course with instructional supports for accessing the curriculum through prerequisite skills. They demonstrate consistent understanding of the knowledge and skills by generalizing the skills to a different context. Students in this category have a high likelihood of showing progress and generalization of knowledge for the assessment tasks at the next grade or course with supports.

Standard-Setting Process for STAAR Alternate

Standard setting for STAAR Alternate involved a process of combining considerations regarding policy, the TEKS content standards, educator knowledge about what students should know and be able to do, and information about how student performance on statewide assessments aligns with student performance on other assessments.

TEA used an evidence-based standard-setting approach (O'Malley, Keng, & Miles, 2012) for the STAAR program. Using this approach, TEA defined and implemented a nine-step process to establish performance standards for all the STAAR Alternate 3–8 and EOC assessments. The nine steps were:

1. Conduct validity and linking studies
2. Develop performance labels and policy definitions
3. Convene a policy committee and/or develop reasonable ranges for performance standards
4. Develop grade- and course-specific performance level descriptors (PLDs)
5. Convene standard-setting committees
6. Review performance standards for reasonableness
7. Approve performance standards
8. Implement performance standards
9. Review performance standards

Table 6.4 provides high-level descriptions and timelines for the steps in the STAAR Alternate standard-setting process.

Table 6.4. Overview of the STAAR Alternate Standard-Setting Process

Standard-Setting Step	Description	Timeline
1. Conduct validity and linking studies	Scores on each assessment are linked to performance on the Texas Assessment of Knowledge and Skills–Alternate (TAKS–Alt).	Spring 2012
2. Develop performance labels and policy definitions	A committee was convened jointly by the Texas Education Agency (TEA) and the Texas Higher Education Coordinating Board (THECB) to recommend performance categories, performance category labels, and general policy definitions for each performance category. The STAAR Alternate performance labels and policy definitions were adapted from those created by the committee.	September 2010
3. Convene a policy committee and/or develop reasonable ranges for performance standards	The committee considered the policy implications of performance standards and validity and linking study results and made recommendations to identify reasonable ranges for performance standards ("neighborhoods"). The STAAR EOC recommendations served as the foundation for decisions made regarding STAAR Alternate.	February 2012
4. Develop grade-/course-specific performance level descriptors (PLDs)	TEA created draft specific PLDs and educator committees reviewed and edited the PLDs. A goal of the development and review of the specific PLDs was to create an aligned system, describing a reasonable progression of skills within each subject area (mathematics, reading, science, and social studies).	July 2012
5. Convene standard-setting committees	Committees consisting of general education and special education experts with experience in grades 3–12 used performance labels, policy definitions, specific PLDs, and predetermined ranges within which to recommend cut scores for each STAAR Alternate assessment. These committees also provided comments to assist TEA with finalizing the specific PLDs.	September 2012
6. Review performance standards for reasonableness	TEA reviewed the recommendations across subject areas.	October 2012
7. Approve performance standards	The commissioner of education approved the STAAR Alternate performance standards.	December 2012
8. Implement performance standards	Once established, performance standards were reported to students for the spring 2012 administration. The process for adjusting cut scores for the 2011–2012 school year was also determined.	January 2013
9. Review performance standards	Performance standards are reviewed at least once every three years.*	If applicable

* In June 2013, the 83rd Texas Legislature enacted HB 5, which removed the requirement to review performance standards (Step 9). Prior to this legislation, Step 9 was scheduled for fall 2014. TEA may review the performance standards if deemed applicable.

More details about each of the steps in the STAAR Alternate standard-setting process are provided in the State of Texas Assessments of Academic Readiness (STAAR)

Alternate Standard Setting Technical Report available on the STAAR Alternate Standard Setting Information page of TEA's Student Assessment Division website.

Standard-Setting Committees

The goal of each standard-setting committee was to recommend two cut scores that would define the three performance levels for each of the STAAR Alternate assessments. The standard-setting committees were made up of K–12 educators. When selecting standard-setting committee members, TEA placed an emphasis on experience with the population of students for whom STAAR Alternate is appropriate, as well as content knowledge and classroom experience. Standard-setting committees also included educators who had ELL and general education expertise.

In September 2012, educator committees were convened to recommend performance standards for all STAAR Alternate assessments. Committees reviewed STAAR Alternate assessment tasks, policy definitions, PLDs, the scoring rubric, and score profiles. The panelists also received training in the evidence-based standard-setting process that incorporated aspects of the extended Angoff process (Angoff, 1971; Hambleton & Plake, 1995) and modified performance profiles (Morgan, 2003) process as well as external validity data. Committee members were provided reasonable ranges within which performance standards should be set. The ranges were determined by two guiding principles: the STAAR Alternate cut score should be more rigorous than the TAKS–Alt cut score, and students needed to complete at least one task at Complexity Level 2 in order to reach Level II: Satisfactory Academic Performance. Ranges were also considered with respect to the policy definitions for the performance levels. With this information in mind, committee members were asked to provide recommendations for where the cut scores should be placed in order to create the three STAAR Alternate performance levels. Three rounds of recommendations were provided, with time for discussion and feedback between rounds. Committee members also participated in an articulation round where they could look at the third-round recommendations across grades/courses and suggest adjustments. TEA used the third round and articulation recommendations in making final decisions about the performance standards.

Implementation of Performance Standards

A phase-in period was implemented for performance standards throughout the Texas assessment program. This phase-in is intended to provide time to adjust instruction, to provide new professional development, to increase teacher effectiveness, and to close knowledge gaps. Although STAAR Alternate underwent the same process used on other state assessments to develop more rigorous assessments and standards, student performance on STAAR Alternate did not indicate the same need to provide school districts with additional time to adjust instruction and close knowledge gaps. Specifically, the difference in difficulty between TAKS–Alt and STAAR Alternate, in terms of student performance, was smaller than the difference in difficulty observed between the STAAR or STAAR Modified assessments and TAKS.

During the standard-setting meetings, educator input also indicated that it was necessary to make changes to the STAAR Alternate administration guidelines. One suggested change was to allow a combination of Complexity Level 1 and Complexity Level 2 tasks, which was not allowed in 2011–2012. This change would allow a path to proficiency for all students taking the STAAR Alternate assessment when the final cut score was implemented in 2012–2013. However, because the combination of Complexity Level 1 and Complexity Level 2 tasks was not allowed in 2011–2012, there was a need for an adjusted cut score for the first year of the new STAAR Alternate (2011–2012) that would allow students completing only Complexity Level 1 tasks to reach proficiency or Level II: Satisfactory Academic Performance.

Students who take STAAR Alternate will graduate under the Minimum High School Plan (MHSP). In contrast to the general STAAR EOC program, STAAR Alternate does not have the requirement that students must achieve Level III: Accomplished Academic Performance in order to graduate under the Distinguished Achievement Program. Therefore, there is no adjusted cut score for the STAAR Alternate Level III: Accomplished Academic Performance.

Outcome of Standard Setting

The purpose of the standard-setting process is to establish cut scores that reflect the level of performance a student must demonstrate in order to be classified into a performance level on each STAAR Alternate assessment. These performance standards were approved by the commissioner of education in December 2012. Table 6.5 presents the approved performance standards, both adjusted and final, for STAAR Alternate.

Table 6.5. STAAR Alternate Performance Standards

Performance Level	Implementation Year	
	2011–2012 (Adjusted)	2012–2013 (Final)
Level I: Developing Academic Performance	0 – 47	0 – 49
Level II: Satisfactory Academic Performance	48 – 77	50 – 77
Level III: Accomplished Academic Performance	78 – 74	78 – 84

Review of Performance Standards

In June 2009, Texas Education Code §39.0242 required that performance standards for the STAAR program be reviewed at least once every three years. Step 9 of the standard-setting process, review the STAAR performance standards, was scheduled for fall 2014. In June 2013, the 83rd Texas Legislature enacted HB 5, which removed the requirement to review performance standards. TEA may review the performance standards if deemed applicable.

Scaling

STAAR assessment scores are commonly reported as both raw and scale scores. With performance-based assessments such as STAAR Alternate, scoring is based on a rubric. A single rubric is used to score student performance across all STAAR Alternate tasks, forms, and administrations, and because of the training provided, test administrators are able to apply this rubric in a consistent way. Given the use of a rubric and the consistency of its application, raw scores for STAAR Alternate can be compared across forms and administrations. Therefore, scale scores are not computed for STAAR Alternate.

For each student who participates in STAAR Alternate, several raw scores are reported. These include the four assessment task scores and the total test score. For each assessment task, the Demonstration of Skill score, Level of Support score, and Generalization of Skill score are summed to obtain the total assessment task score. Assessment task scores range from 0 to 21. The overall total test score is computed by summing the four assessment task scores. Thus, the total test score for STAAR Alternate ranges from 0 to 84. For more details about STAAR Alternate scores, refer to the Scores and Reports section of this chapter.

Equating

Equating activities are not conducted for STAAR Alternate. The difficulty level of the assessments and assessment tasks is taken into consideration through the differential weighting of the complexity level of each task. In addition, consistency across administrations is maintained through the requirement of training and qualification procedures that are completed by test administrators before delivering the assessment. Due to the training and qualification activities, it can be assumed that test administrators are able to apply the STAAR Alternate rubric consistently and maintain the integrity of the STAAR Alternate raw score scale across assessment tasks and administrations without the need for equating.

Reliability

Assessments that are not traditional paper-and-pencil or multiple-choice tests may require a different approach to gather reliability evidence. Interrater reliability for STAAR Alternate is evaluated by having two raters simultaneously observe the same student performing a specific assessment task. Both raters evaluate the student's performance using the assessment's performance evaluation questions, and then the two independent ratings are compared to determine the reliability, or consistency of scoring, for STAAR Alternate. The first and second ratings for this study took place during the 2011–2012 STAAR Alternate assessment window (January 7, 2013 to April 19, 2013). The following grades and subjects were included in the study:

- Grades 5 and 8 mathematics
- Grades 5 and 8 reading
- Grade 7 writing

A target sample size of 300 students for each grade/subject was selected to obtain the desired statistics for the study, for a total of 1,500 students, taking the return rate into account. The interrater reliability study was sampled at the campus level. Not all students who were selected were able to participate in the study, resulting in a final sample size of 1,283 students.

When a student was selected for the interrater study, the student's primary teacher was notified by TEA of the subject area and essence statement to be observed by the second rater. To standardize the qualifications for selection of the second rater, TEA developed and published the following guidelines.

The second raters for STAAR Alternate should be professionals or under the supervision of professionals who hold valid education credentials such as Texas teacher certificates or permits. Those selected may include the following:

- teachers (including general, special education, and teachers for the visually and auditorily impaired)
- paraprofessionals (when appropriate)
- assessment specialists
- speech therapists
- occupational therapists
- physical therapists

The second rater for the study should be someone who knows the student well enough to provide a knowledgeable rating and is available to observe the assessment task. The second rater is required to complete all the STAAR Alternate training modules and successfully complete the online training qualification activities before taking part in the study.

Once the second rater qualified for the study, the primary teacher was able to plan a task and schedule when the task would be administered to the student.

The primary teacher implemented the assessment task while the second rater observed and both raters documented their own observations on a data observation sheet, which was used to complete the student performance evaluation questions on the STAAR Alternate online system by the primary teacher. Second raters entered their data into the STAAR Alternate interrater reliability online tool by answering similar student performance evaluation questions. Once the observation was complete, the raters gave their separate, independent ratings.

The second rater was asked to respond to the performance evaluation questions for demonstration of skill and level of support. The generalization of skill dimension was not included in the study because the student may demonstrate this skill at a different time and the second rater may not be present when generalization occurs.

The results for the interrater reliability study are provided in the following tables. There are several statistics calculated for the study's results. The results are shown for each dimension of the rubric (except for generalization of skill) and the score combining Demonstration of Skill and Level of Support.

Correlation coefficients, which describe the relationship between two variables, are calculated between the two sets of ratings as shown in Table 6.6. Complexity Level correlations show the strongest relationship between first and second raters (0.96 to 0.98). Correlations between the raters are not as strong for Demonstration of Skill (ranging from 0.64 to 0.92), Level of Support (ranging from 0.60 to 0.85) and combined score (0.56 to 0.89) but still indicate a high level of agreement between the two ratings. The correlation for mathematics ranges from 0.66 to 0.98; ranges from 0.56 to 0.97 for reading; and ranges from 0.68 to 0.98 for writing. The patterns of correlations across subject areas are similar with the highest correlation found in Complexity Level and the lowest correlation found in Level of Support. The exception is grade 8 reading and writing. The correlation between the two ratings for Combined Score is slightly lower than the correlation for Level of Support.

Table 6.6. Correlations Between First and Second Ratings

	Grade	Complexity Level	Demonstration of Skill	Level of Support	Combined Score	Sample Size
Mathematics	5	0.98	0.77	0.60	0.66	297
	8	0.98	0.92	0.76	0.85	267
Reading	5	0.97	0.90	0.85	0.89	257
	8	0.96	0.64	0.60	0.56	275
Writing	7	0.98	0.70	0.69	0.68	277

In addition to the above correlation, the weighted Kappa (Cohen, 1968) is calculated. The Kappa statistic produces an estimate of the reliability of the score classifications between the two sets of ratings. It is a measure of reproducibility for categorical data. A common stumbling block in evaluating scoring reliability or consistency is the basic concept of agreement beyond chance and, in turn, the importance of correcting for chance agreement. The Kappa statistic corrects for this chance agreement and tells us how much of the possible agreement over and above chance the scorers have achieved. By using the weighted Kappa, larger differences between two ratings are given smaller weights. It can differ from the agreement measure in certain conditions. Therefore, the weighted Kappa was evaluated only as one of the many pieces of evidence supporting the reliability of STAAR Alternate assessments.

Table 6.7 lists the guidelines for the evaluation of the weighted Kappa agreement coefficients (Altman, 1991):

Table 6.7. Guidelines for Weighted Kappa Agreement Coefficients

Value of K_w,	Strength of Agreement
< 0.20	Poor
0.21 – 0.40	Fair
0.41 – 0.60	Moderate
0.61 – 0.80	Good
0.81 – 1.00	Very Good

The weighted Kappa coefficients are listed in Table 6.8. The results show a high degree of consistency between the two sets of ratings. In particular, Complexity Level had a weighted Kappa statistic that denoted very good reproducibility. Demonstration of Skill, Level of Support and Combined Score produced a weighted Kappa statistic denoting good reproducibility.

Table 6.8. Weighted Kappa Agreement Coefficients Between First and Second Ratings

	Grade	Complexity Level	Demonstration of Skill	Level of Support	Combined Score	Sample Size
Mathematics	5	0.97	0.82	0.60	0.71	297
	8	0.97	0.93	0.73	0.85	267
Reading	5	0.96	0.89	0.77	0.83	257
	8	0.96	0.76	0.61	0.65	275
Writing	7	0.98	0.79	0.63	0.69	277

The percent of agreement between two ratings are calculated for Complexity Level, Demonstration of Skill, Level of Support and Combined Score, as shown in Table 6.9. Table 6.9 shows a high degree of agreement between two sets of ratings. Specifically, the percentage of agreement was high for Complexity Level (97.45% to 98.92%) and Demonstration of Skill (91.27% to 96.63%). The percentage of agreement is slightly lower for Level of Support (80.51% to 90.26%) and the Combined Score (79.06% to 88.76%). Results are similar across subject areas. The agreement rate for Complexity level is highest and for Combined Score is lowest among those four variables.

Table 6.9. Percent Agreement Between First and Second Ratings

	Grade	Complexity Level	Demonstration of Skill	Level of Support	Combined Score
Mathematics	5	98.32%	92.93%	82.83%	82.49%
	8	98.13%	96.63%	90.26%	88.76%
Reading	5	97.67%	94.55%	85.60%	83.27%
	8	97.45%	91.27%	84.00%	82.55%
Writing	7	98.92%	92.42%	80.51%	79.06%

The range of the correlation coefficient, the percent of agreement, the strength of agreement level of the kappa coefficients indicate that the relationships between the first and second ratings are high for STAAR Alternate. This trend occurred for all three grade levels and across all three subjects for Complexity Level, Demonstration of Skill, Level of Support, and for the score combining Demonstration of Skill and Level of Support—supporting the reliability of the STAAR Alternate.

Validity

STAAR Alternate scores are used to make inferences about student achievement. In support of these inferences, evidence is continually collected throughout the development and administration of STAAR Alternate to demonstrate that the assessments measure the intended content. This validity evidence can be categorized as being based on test content, response processes, internal structure, relations to other variables, and the consequences of testing. This validity evidence supports multiple uses of test scores. Texas follows national standards of best practice to continue to build its body of validity evidence for the STAAR assessments. The Texas Technical Advisory Committee (TTAC) provides ongoing input to TEA about STAAR Alternate validity evidence. The following sections describe the validity evidence that has been collected for STAAR Alternate.

Evidence Based on Test Content

Validity evidence based on test content refers to evidence of the relationship between tested content and the construct the assessment is intended to measure. All STAAR assessments, including STAAR Alternate, have been designed to align with the content defined by the TEKS. The STAAR Alternate test development process played an integral role in providing validity evidence based on test content for the assessment. The test development process and the evidence collected related to test content support the use of STAAR Alternate scores in making inferences about students' knowledge and understanding of the TEKS.

Relationship to the Statewide Curriculum

The revised TEKS curriculum was adopted for reading and writing in 2008 and for mathematics, science, and social studies in 2009. To link the STAAR Alternate

assessment to the revised Texas grade-level content standards, vertical alignment and curriculum framework documents for STAAR Alternate were developed for the reading/English language arts, writing, mathematics, science, and social studies TEKS curriculum. The STAAR Alternate vertical alignment and curriculum framework documents help teachers access the grade-level TEKS for students with significant cognitive disabilities. The STAAR Alternate steering committee, as well as educator advisory committees, also reviewed and provided feedback on the alignment of the TEKS to STAAR Alternate.

EDUCATOR INPUT

Professional judgments from educator review meetings provided additional content-validity evidence. Educators from across the state reviewed the content of every assessment task to validate that each task matched the appropriate content standard. The educator committees included special education experts, special education classroom teachers, teachers from the Texas School for the Blind and Visually Impaired and the Texas School for the Deaf, teachers of English language learners, and general education teachers.

As part of the review meetings, educators considered each assessment task and were asked, “Does this assessment task measure the reporting category, student expectation, essence statement, and prerequisite skills it was designed to measure?” To respond to this question, educators referenced resources such as the TEKS curriculum documents to verify the match of the reporting category, student expectation, essence statement, and prerequisite skills to each assessment task. Across STAAR Alternate tasks, educator review committees affirmed the relationship between the assessment tasks and the TEKS. Additional committee input also confirmed that students are provided opportunities to learn the content before the assessment.

Another important source of content validity is evidence related to bias. In order to be valid, an assessment must not only assess the intended content, but also must prevent issues of bias. To provide this validity evidence, educator committees were asked the following question regarding each assessment task: “Is this assessment task free from bias on the basis of students’ personal characteristics such as gender, ethnicity, or disability?” Committee members affirmed that STAAR Alternate tasks are free from bias.

TEST DEVELOPER INPUT

Item writers and reviewers follow test development guidelines that explain how the content of the assessed TEKS should be measured. At each stage of development, writers and reviewers verify the alignment of the test items with the assessed reporting categories.

Evidence Based on Response Processes

TEA also collects evidence about the way students respond to test questions on the STAAR Alternate assessments to determine whether scores are accurate measures of

the construct. To gather this evidence, TEA conducted an audit of STAAR Alternate in 2011–2012. As part of the audit, materials such as documentation forms and student scores for a sample of students were submitted by test administrators for review. These materials were reviewed by a group of teachers to determine whether the scores were supported by the documentation and whether complexity levels were maintained throughout the administration of tasks. Auditors felt that most of the sampled student scores were appropriate and were supported by documentation across each of the scoring components. Auditors also felt that nearly all of the complexity levels were maintained for the audit sample.

Additional information about the STAAR Alternate audit can be found in the 2011-2012 Technical Digest Chapter 6 Audit section.

Evidence Based on Internal Structure

Texas collects evidence that shows the relationship between test items and reporting categories to demonstrate that the parts of an assessment conform to the overall test construct. A measure of internal consistency is used to provide evidence of the internal structure of a test; however, this measure is currently not available for STAAR Alternate, given the limited number and unique characteristics of performance-based items included on the assessment.

Evidence Based on Relationships to Other Variables

Another source of validity evidence is the relationship between test performance and performance on other assessments, that is, criterion-related validity. Several analyses were conducted to show that STAAR Alternate scores are related to each other as expected and related weakly, if at all, to irrelevant characteristics.

For each STAAR Alternate assessment task, three scores are recorded: Demonstration of Skill, Level of Support, and Generalization of Skill. Additionally, as noted earlier, the complexity level of the task also impacts the score. These four scoring components would be expected to be related because they are all based on the same assessment task and student. Correlations of the four STAAR Alternate scoring components were computed to determine the relationships among them. These correlations were based on 92,329 completed assessment tasks, of which 61,544 included generalization of skill. As shown in Table 6.10, correlations were moderate to strong for the relationships among Complexity Level and Demonstration of Skill, Demonstration of Skill and Level of Support, and Level of Support and Generalization of Skill. The correlation between Complexity Level and Generalization of Skill is likely reduced because students who complete Complexity Level 1 tasks are not eligible for Generalization of Skill. Although the correlation is significant, which is likely due to having a large sample of students, it is too small to indicate any practical significance. The low correlation between Demonstration of Skill and Generalization of Skill indicates that these scores represent different skills, which is reasonable given the way these two scores are defined. The negative correlation between Complexity Level and Level of Support is likely related to the way these components are scored. That is, students who need more support are more likely to be assessed with tasks of a lower

complexity level, meaning that high scores on Level of Support are related to low values of Complexity Level. This inverse relationship is suggested by the negative correlation.

Table 6.10. Correlations Among STAAR Alternate Scoring Components

	Complexity Level	Demonstration of Skill	Level of Support	Generalization of Skill
Complexity Level	1.00	–	–	–
Demonstration of Skill	0.71*	1.00	–	–
Level of Support	-0.10*	0.45*	1.00	–
Generalization of Skill	0.02*	0.02*	0.56*	1.00

*Indicates that correlations were significant at the p ≤ .01 level.

Correlations among STAAR Alternate subject area scores were also calculated. As shown in Table 6.11, the correlations among subject area scores were high as expected because the same rubric is used for scoring across all subject areas. The correlations among subject areas ranged from 0.86 to 0.90. This finding is a strong source of validity evidence because the empirical results matched the relationships that were expected among these constructs. Correlations between writing and science and writing and social studies are not included because students do not take these assessments in the same year.

Table 6.11. Overall STAAR Alternate Correlation Between Subject Area Scores

Subject Areas Compared (Using Total Scores)	N-Count	Correlation
Mathematics & Reading	29,562	0.86*
Science & Reading	10,831	0.86*
Social Studies & Reading	11,688	0.87*
Writing & Reading	7,750	0.90*
Science & Mathematics	10,569	0.89*
Social Studies & Mathematics	9,002	0.86*
Writing & Mathematics	7,750	0.87*
Social Studies & Science	6,604	0.89*

*Indicates that correlations were significant at the p ≤ .01 level.

Additional validity evidence was gathered in the form of discriminant validity analyses, which demonstrated that the STAAR Alternate test scores were unrelated to demographic variables. Theoretically, student characteristics such as ethnicity and gender should not relate to their performance on the assessment; therefore, the lack of meaningful empirical relationships among these measures is to be expected.

To investigate the relationship between STAAR Alternate and demographic variables, correlations were computed specifically for gender and ethnicity. The correlation between STAAR Alternate scores and gender was -0.01, and the correlation between STAAR Alternate scores and ethnicity was 0.01. Both the gender and ethnicity correlations are very small and do not indicate a meaningful relationship between STAAR Alternate scores and either demographic variable.

Evidence Based on Consequences of Testing

Another way of providing validity evidence is by documenting the intended and unintended consequences of administering an assessment. Some of the intended consequences of the STAAR Alternate assessment, based on the requirements in federal and state statutes, are:

- Students with the most severe cognitive disabilities can receive challenging instruction that is linked to state content standards.
- Students with the most severe cognitive disabilities can be included in state assessment programs.
- STAAR Alternate assessments can assess the achievement of students with the most severe cognitive disabilities.
- Performance on STAAR Alternate assessments can be used to track the academic progress of students across years.

Unintended consequences have not yet been identified given that 2012–13 was only the second year of STAAR Alternate administrations and because the assessment is also under redesign efforts resulting from legislation passed in 2013.

Measures of Student Progress

While progress measures were calculated and reported for STAAR in 2012–2013, these measures were not yet available for STAAR Alternate. The development of these measures continues, and they will be reported for the first time in 2013–2014.

Because of the unique characteristics of STAAR Alternate and the students who take it, progress for STAAR Alternate will be measured differently from that for STAAR and STAAR Modified. While STAAR and STAAR Modified use a gain score approach for progress, STAAR Alternate will use a transition table methodology (for more information about types of progress measures see the Measures of Student Progress section in chapter 3, "Standard Technical Processes"). While the approach used to measure progress is different, student progress on STAAR Alternate, like that on STAAR and STAAR Modified, will be classified as *Did Not Meet*, *Met*, or *Exceeded*.

As the development of the STAAR Alternate progress measure proceeds additional information will be made available on the STAAR Alternate Resource page of TEA's Student Assessment Division website.

Sampling

In 2012–2013, stratified sampling without replacement was used to generate a sample for the STAAR Alternate interrater reliability study (refer to the Sampling section of chapter 3, "Standard Technical Processes," for descriptions of sampling methods). The objective of the audit was to include one percent of the STAAR Alternate population, or 1,500 students.

To identify the sample, all students who take STAAR Alternate were first stratified or grouped by campus. Then within each campus a simple random sample was generated. In consideration of the additional work associated with the interrater reliability procedures, the number of students selected from each campus was limited to two for grades 3–8 and four for high school campuses.

Demographics of the sample were then considered and compared to those of the entire STAAR Alternate population. Specifically, the sample was checked for representativeness in terms of gender, ethnicity, economic disadvantage status, and geographic region. Because the sample was carefully selected to represent the entire population of STAAR Alternate students, the results of the audit could be generalized beyond the students sampled.

Test Results

Appendix D provides STAAR Alternate score information based on 2012–2013 administrations. The following data are included: summary statistics, such as the mean and standard deviation; score distributions for each STAAR Alternate assessment; the number of assessment tasks administered at each complexity level; assessment task score distributions by complexity levels; and the distribution of assessment task complexity level combinations. Table 6.12 shows spring 2013 pass rates for STAAR Alternate.

Table 6.12. STAAR Alternate Spring 2013 Pass Rates (at the Recommended Standard)

Subject Area	Grade/Course	Pass Rate
Mathematics	Grade 3	72%
	Grade 4	75%
	Grade 5	75%
	Grade 6	73%
	Grade 7	72%
	Grade 8	72%
	Algebra I	65%
	Geometry	69%
Reading/English Language Arts	Grade 3	71%
	Grade 4	71%
	Grade 5	70%
	Grade 6	71%
	Grade 7	70%
	Grade 8	69%
	English I	68%
	English II	68%
	English III	66%
Writing	Grade 4	72%
	Grade 7	70%
Science	Grade 5	73%
	Grade 8	73%
	Biology	68%
Social Studies	Grade 8	73%
	World geography	67%
	World history	66%
	U.S. history	69%