

# Chapter 5 STAAR Alternate 2

Overview

Participation Requirements

Testing Requirements for Graduation

Test Development

Training

Test Administrations

Testing Accommodations

Scores and Reports

Performance Standards

Scaling

Equating

Reliability

Validity

Measures of Student Progress

Sampling

Test Results

## Overview



developed assessment questions that align to the grade-level [Texas Essential Knowledge and Skills \(TEKS\)](#). Teachers evaluate student performance based on



cognitive disability must follow the Texas definition of a significant cognitive disability. A student with a significant cognitive disability has limited potential to reach grade-level expectations, and evidence must be documented in the Participation Requirement information as supported by the student's most recent Full and Individual Evaluation (FIE) report.

2.





After prototype items were developed, cognitive labs were conducted to gather information on student performance, engagement, and interaction with the redesigned STAAR Alternate 2 items. Test administrators were interviewed regarding the proposed test design and the feasibility of the assessment for students. The next step was a pilot test to gather further student performance data and a survey of test administrators regarding the STAAR Alternate 2 test items. The data from the cognitive labs, pilot tests, and test administrator surveys were used to develop items for the first operational assessment in spring 2015.

### Assessment



on7l rmop ional mtewional2.7P(A)2 (R)2.6]( and , (es)8. N( )7)1a (t)-6.6 ( i)2. (onal)2.6 ( )]TJ C(t)-6



populations.

## Review of Items

During the item development process for STAAR Alternate 2, educator committees met to complete reviews of every item. The committees were comprised 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 EB students, and general education teachers.

The educator committees focused on the relationship between the grade-level content and the items. Each committee member completed an item judgment form with the following questions regarding each item:

Does this item measure the reporting category/student expectation/essence statement/prerequisite skill it was designed to measure?

Is this item an appropriate measure of the TEKS student expectation/essence statement/prerequisite skill?

Is this item free from bias based on students' personal characteristics such as gender, ethnicity, or disability?

Would you expect students in your district to have received sufficient instruction by the end of the grade/course to enable them to answer this item correctly?

Feedback from the educator committees was used to revise the STAAR Alternate 2 items as needed.

## Training

Resources available for training are listed in the Technical Digest.

## Test Administrations

To ensure equity for all students, a student had to be present at a monitored testing session and be supervised by a trained test administrator. The STAAR Alternate 2 testing window was extended to seven weeks to provide districts more flexibility in scheduling assessments and more opportunity to test all eligible students. Additionally, the testing window was extended by one week for a small subset of larger districts to ensure that ratings were properly entered into DEI during the test administration window.

### Test Administration Procedures

The STAAR Alternate 2 assessment process is designed with scripted test administrator presentation instructions that mirror instructional techniques for a student with a significant cognitive disability. The essence statements, upon which the 2021–2022 STAAR Alternate 2 items were based, were made available in fall 2021 so that



The student is unable to respond to test questions due to a terminal or degenerative illness.

The student receives extensive short-term medical treatment due to a medical emergency or serious injury in an accident.

The student is unable to interact with peers or staff without risk of infection or contamination to himself/herself or others.

The student is unable to receive sufficient or consistent homebound services due to medical issues.

#### NO AUTHENTIC ACADEMIC RESPONSE

Students who are unable to produce an authentic academic response due to level of cognition rather than a medical condition may qualify for a NAAR exception under the following circumstances:

Because of multiple impairments, the student is unable to receive information during instruction and assessment. For example, the student may have a combination of sensory impairments, such as hearing, vision, and/or tactile.

The student is consistently unable to provide an authentic academic response during instruction. His or her behavior may be described by one or more of the following characterizations:

The student is unable to demonstrate a meaningful, observable reaction to a specific stimulus.

The student exhibits only startle responses.

The student tracks or fixates on objects at random and not for a specific purpose.

The student moves or responds only to internal stimuli.

The student vocalizes intermittently regardless of changes in environment around him or her.

## Testing Accommodations

STAAR Alternate 2 is a standardized assessment intended to be appropriate for eligible students in its original intact form. However, ARD committees and test administrators may elect to provide appropriate allowable accommodations to students whose disability precludes them from participating meaningfully in the assessment.

Test administrators may use accommodations only if they are routinely provided in classroom instruction and listed in the student's IEP. Some accommodations provided





during classroom instruction may not be allowed during testing, as certain accommodations used in the classroom would invalidate the content being assessed or compromise the security and integrity of the test. Figure 5.1 provides examples of accommodations that can be considered for STAAR Alternate 2 and guidelines on how such accommodations should be applied.

Figure 5.1. STAAR Alternate 2 Allowable Accommodations

Allowable Accommodations
<ul style="list-style-type: none"> <li>■ Color or highlight images or text</li> <li>■ Place color overlays on images or text</li> <li>■ Pair images or text with photographs, picture representations, or <b>real objects of the same content</b> <ul style="list-style-type: none"> <li>• photographs, pictures, or real objects that have close to the original appearance</li> </ul> </li> <li>■ Attach textured materials to images or text</li> <li>■ Demonstrate concepts or relationships in images or text</li> <li>■ Raise or darken the outline in images or text</li> <li>■ Enlarge images or text                     <ul style="list-style-type: none"> <li>• magnification devices, photocopying, or computer magnification</li> </ul> </li> <li>■ Add braille labels to images or provide text in braille</li> <li>■ Describe images for students with visual impairments                     <ul style="list-style-type: none"> <li>• descriptions of images that only include details of what can be seen in the images without the overall impression of the image</li> </ul> </li> <li>■ Provide images on a non-surface paper presented one at a time                     <ul style="list-style-type: none"> <li>• images must be presented in the same order or configuration as they appear in the test booklet</li> </ul> </li> <li>■ Cover or isolate images or text</li> <li>■ Use routine picture representations or keywords in verbal directions to the student                     <ul style="list-style-type: none"> <li>• only what is visually presented, stated in text, or supplied in the test administrator instructions can be provided</li> </ul> </li> <li>■ Use calculator, manipulatives, or math tools to arrive at response                     <ul style="list-style-type: none"> <li>• fraction pieces, geometric shapes, number lines, number charts, money, base-ten blocks, counters</li> </ul> </li> <li>■ Reread sections of the text                     <ul style="list-style-type: none"> <li>• Follow the guidelines in the “Presentation Instructions” section of the test administrator manual for guidance on repeating presentation instructions and rereading sections of the text</li> </ul> </li> <li>■ Provide structured reminders                     <ul style="list-style-type: none"> <li>• personal timers, token systems, color cards or manipulation reminders, or visual cues</li> </ul> </li> </ul>

## Scores and Reports

### Scoring STAAR Alternate 2 Assessments

STAAR Alternate 2 is scored polytomously using a standard scoring rubric that follows the same process of item administration across all items and is applied to the student performance evaluation information that test administrators submit electronically via DEI. Each item is scored according to the level of independence with which a student responds. The scoring rubric is as follows:

If a student responds correctly to the first presentation of an item, he or she receives a score point of 2. If the student does not respond or responds incorrectly, the item is presented again with allowable teacher assists.

If the student responds correctly to the second presentation of the item, he

or she receives a score point of 1.

If the student does not respond or responds incorrectly to the second presentation, he or she receives a score point of 0.

Each item is scored in the same manner. Item scores range from 0 to 2. There are 20 scored items per test, resulting in a total test score range of 0 to 40 points.

## Description of Scores

Scores for the STAAR Alternate 2 assessments consist of the number of points earned (raw scores), scale scores, and the resulting performance level associated with the student's score.

### RAW SCORE

The number of points that a student earns on a STAAR Alternate 2 assessment is the student's raw score. The raw score can be interpreted only in terms of the specific set of test items on that test form. However, because the difficulty of items might vary among test forms over time, raw scores alone cannot be used to compare performance across tests or administrations. To make these comparisons, raw scores must be converted to scale scores.

### SCALE SCORE

A scale score is a conversion of the raw score onto a scale that is common to all test forms for that assessment. Scale scores allow for direct comparisons of student performance between specific sets of test items from different test administrations.

The scale score is used to determine whether a student attained Level II: Satisfactory Academic Performance or Level III: Accomplished Academic Performance.

Performance-level cut scores are discussed in the







Alternate 2 test booklets, policy definitions, and PLDs. The panelists also received training in the evidence-



Subject Area	Grade/Course	Level II: Satisfactory	Level III: Accomplished
Sct(ti)4.o (t)e f* EMC /Artifac eh(S)33[(S)0 9450mHTc 0 Twc eh(.72 7216(236MC BT /P9c e1rtifac)-Y e)9.4 (2)-2.4 g			







previously on one or more test forms. This permits the difficulty level of the newly developed form to be closely determined even prior to its administration. Thus, the



## Classical Standard Error of Measurement

Classical SEM represents the amount of variance in a score that results from random factors other than what the assessment is intended to measure. The SEM is helpful for quantifying the margin of uncertainty that occurs on every test.

## Conditional Standard Error of Measurement

It is important to note that the SEM index provides only an estimate of the average test score error for all students regardless of their individual levels of proficiency. By comparison, CSEM provides a reliability estimate at each score point on a test. More specifically, CSEM is an estimate of the average test score measurement error that is conditional on the proficiency or scale score estimate.

## Classification Accuracy and Consistency

Classification accuracy and consistency provide estimates of the accuracy and consistency of student classifications into performance categories based on current test results.

## Validity

STAAR Alternate 2 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 2 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. TEA follows national standards of best practice to continue to build its body of validity evidence for all the STAAR assessments. The Texas Technical Advisory Committee (TTAC) provides ongoing input to TEA about STAAR Alternate 2 validity evidence. The following sections describe the validity evidence that has been collected for STAAR Alternate 2.

## Evidence Based      Based

## RELATIONSHIP TO THE STATEWIDE CURRICULUM

At the inception of the STAAR Alternate 2 assessments, a steering committee was convened to review and provide feedback on the alignment of STAAR Alternate 2 assessment tasks to the TEKS. Educator reviews and focus group meetings continue to be a part of ongoing content development with revisions to the STAAR Alternate 2. Both focus groups and educator review meetings have occurred to review and provide feedback on alignment of items and content standards as well as to review and provide feedback on items themselves.

In 2015–2016, an independent third-



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

Response processes refer to the cognitive behaviors that are required to respond to a test item. Texas collects evidence to show that the way students respond to items on the STAAR Alternate 2 assessments reflects accurate measurement of the construct.

### ITEMS

Texas gathers theoretical and empirical evidence that support the expectation that the way students respond to test items does not add construct-irrelevant variance. Every year, during item reviews, educators evaluate whether the content for a given item is being appropriately assessed and whether students will be able to accurately demonstrate their knowledge of the construct given the items' planned format. When items are field tested, additional student response data are gathered. Data such as item difficulty, item-total correlations, and item fit are all evaluated. For additional information, see the Item Analyses section of [Chapter 3, "Standard Technical Processes."](#)

### SCORING PROCESS

The process used to score items can provide additional validity evidence based on response processes. This type of validity evidence is predicated on accurate scoring. Within the test administrator booklet, test administrators are provided exact scoring rules and scripted instructions for how to present every item to a student. Test administrators are provided resources to prepare for a STAAR Alternate 2 test administration, including a period of time directly prior to the testing window in which they can preview the test booklet to prepare for a valid test administration.

Evidence Based on Internal

relationships are consistent with the expected relationships. STAAR Alternate 2 correlation estimates, which evaluate the strength of the relationship (or the lack of one) between scores on the STAAR Alternate 2 assessments across different content areas (for example, grade 4 mathematics and grade 4 reading, or biology and U.S. history) are calculated.

## Evidence Based on Consequences of Testing

Another way of providing validity evidence is by documenting the intended and unintended consequences of administering an assessment. Possible unintended negative consequences can include narrowing of curriculum or instruction to focus on specific learning outcomes that are assessed, or the inappropriate use of test scores by stakeholders. Some of the intended consequences of the STAAR Alternate 2 assessment, based on the requirements in federal and state statutes, are as follows:

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 2 can assess the achievement of students with the most severe cognitive disabilities.

Performance on STAAR Alternate 2 assessments can be used to track the academic progress of students across years.

## Measures of Student Progress

Student progress measures provide information beyond performance level by considering performance over time. Whereas performance-





Table 5.5. STAAR Alternate 2 Spring 2022 Pass Rates

Subject Area	Grade/Course	Pass Rate
Mathematics	Grade 3	92%
	Grade 4	95%
	Grade 5	94%
	Grade 6	93%
	Grade 7	95%
	Grade 8	95%
	Algebra I	93%
Reading/English Language Arts	Grade 3	85%
	Grade 4	88%
	Grade 5	89%
	Grade 6	91%
	Grade 7	92%
	Grade 8	93%
	English I	91%
	English II	94%
Science	Grade 5	94%
	Grade 8	97%
	Biology	96%
Social Studies	Grade 8	95%
	U.S. History	95%

