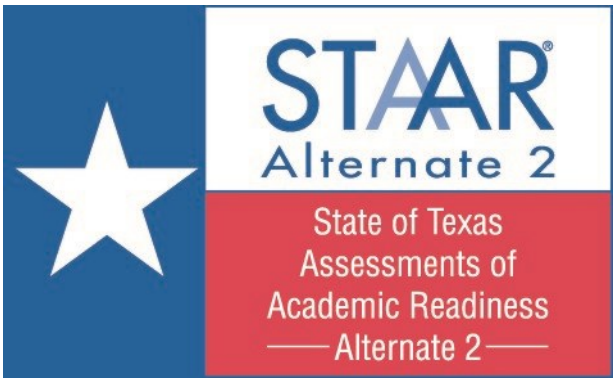


State of Texas Assessments of Academic Readiness Alternate 2 (STAAR® Alternate 2)

Standard-Setting Technical Report

September 7, 2015

Table of Contents

Chapter 1: The State of Texas Assessments of Academic Readiness Alternate 2 (STAAR® Alternate 2) 4
Overview of STAAR Alternate 2 4
TEKS Curriculum Frameworks for STAAR Alternate 2 5
STAAR Alternate 2 Performance Standards 6
Chapter 2: Overview of STAAR Alternate 2 Standard-Setting Process 8
Goals of Setting Performance Standards 8
Evidence-Based Standard Setting 8
TEKS Curriculum Standards 9
Assessment 10
Policy Considerations and External Validation 10
Expertise and Knowledge about Students and Subject Matter 10
Standard Setting 10
The STAAR Alternate 2 Standard-Setting Process 10
Chapter 3: Preparation for Standard Setting 12
Purpose of Analyses Performed in Preparation for Standard Setting 12
STAAR Alternate 2 Cut Score Reasonable Ranges 12
Data Used during Standard Setting 13
Chapter 4: Performance Labels, Policy Definitions, PLDs, and Borderline Descriptors 14
Purpose of Performance Labels and Policy Definitions 14
Development of Performance Labels and Policy Definitions 14
Level III: Accomplished Academic Performance 15
Level II: Satisfactory Academic Performance 15
Level I: Developing Academic Performance 16
Purpose of the Specific PLDs 16
Development of the Specific PLDs 16
Texas Educator Review of the Specific PLDs 17
Development of Borderline Descriptors 17
Review and Approval Process for the PLDs 18
Chapter 5: Standard Setting for STAAR Alternate 2 19
Purpose of Standard–Setting Meetings 19

Committee Composition and Attendees 19
Implementing the Standard-Setting Process 20
Review Assessment Booklets 21
Review Specific PLDs and BDs 21
Round 1 Judgments 21
Round 1 Feedback 21
Round 2 Judgments 22
Round 2 Feedback 22
Round 3 Judgments 22
Cross-Course Articulation 22
Evaluation Survey 23
Standard-Setting Process Evaluation Results 23
Meeting Success 23
Usefulness of Activities and Information 23
Adequacy of the Meeting Elements 23
Confidence in Specific PLDs. 23
Confidence in Round 3 Ratings. 24
Opportunities to Collaborate. 24
Respect. 24
Chapter 6: Reasonableness Review, Approval, and Implementation of Performance Standards 25
Purpose of Reasonableness Review 25
Rationale for Adjustments Made During Reasonableness Review 25
Implementation of Standards 25
References 27
Appendix 1: STAAR Alternate 2 Standard-Setting Committee Composition 28
Appendix 2: STAAR Alternate 2 Standard-Setting Results by Round 37
Appendix 3: STAAR Alternate 2 Impact Data Using Final Cut Scores 47
Appendix 4: STAAR Alternate Process Evaluation Survey and Results 51
Appendix 5: STAAR Alternate 2 Standard-Setting Process Evaluation Results 56

Chapter 1: The State of Texas Assessments of Academic Readiness Alternate 2 (STAAR® Alternate 2)

This chapter provides an overview of the STAAR Alternate 2 program and includes the following sections:

- Overview of STAAR Alternate 2
- TEKS Curriculum Frameworks for STAAR Alternate 2
- STAAR Alternate 2 Performance Standards

Overview of STAAR Alternate 2

As a result of House Bill 5 of the 83rd Texas Legislative Session, TEA has redesigned the STAAR Alternate assessment to meet the diverse needs of students with significant cognitive disabilities. The legislation, as quoted below, indicates that the assessment must not require teachers to prepare tasks or materials.

> "TEA, in conjunction with appropriate interested persons, shall redevelop assessment instruments adopted or developed under Subsection (b) for administration to significantly cognitively disabled students in a manner consistent with federal law. An assessment instrument under this subsection may not require a teacher to prepare tasks or materials for a student who will be administered such an assessment instrument."

To meet requirements of the legislation and maintain an appropriate assessment for students with significant cognitive disabilities, a question-based approach to the assessment has been implemented for the redesign of STAAR Alternate. The assessment consists of 24 scripted questions. The test materials include a test administrator manual with the scripted questions and guidelines for how the test is administered. A student booklet is provided that contains stimulus images and text needed for the student to select answers. This design allows for standardization of the assessment and eliminates the need for teachers to prepare tasks or materials.

The STAAR Alternate 2 development process mirrors the procedures used for all state assessments in Texas. The issues of validity, reliability, fairness, accessibility, and consistency within item types were carefully considered as part of the development for the assessment. When developing STAAR Alternate 2, attention was also given to the criteria of fairness, principles of alignment, and universal design. These principles were considered from the beginning to bridge the gap between the grade-level content and the learning styles of students with significant cognitive disabilities. In incorporating universal design for STAAR Alternate 2, consideration was also given to students' response modes, which allow students to show what they know and are able to do. Accommodations were also considered to allow students access to individualization that is unique to their special needs.

The STAAR Alternate 2 assessment replaces the previous STAAR Alternate assessment. Although there have been significant changes to create the new STAAR Alternate 2 assessment, the intent of STAAR Alternate 2 remains much the same as it was under the original STAAR Alternate assessment. The changes implemented for STAAR Alternate 2 were intended to standardize the administration but maintain the same level of content coverage and rigor as the original STAAR Alternate assessment.

STAAR Alternate 2 is an alternate assessment designed to measure student progress on skills aligned with academic grade-level content standards. STAAR Alternate 2 is designed as an evaluation of the student's knowledge and skills demonstrated through selected responses by the student. The student is scored according to their answer choices while participating in state-developed assessment items that are linked to the Texas Essential Knowledge and Skills (TEKS) curriculum through prerequisite skills. For the STAAR Alternate 2 assessment, teachers follow a standard administration script and scoring protocol. Students assessed with STAAR Alternate 2 are those with the most significant cognitive disabilities who cannot participate in other components of the Texas student assessment program, even with accommodations. More than one percent of students with significant cognitive disabilities may take STAAR Alternate 2, although only one percent can be counted as proficient (based on alternate performance standards) for federal Adequate Yearly Progress (AYP) reporting purposes.

On the STAAR Alternate 2 test, each question measures a targeted prerequisite skill. Each essence statement has four questions that form a cluster and test a common skill or concept at varying levels of difficulty. Six clusters comprise a test form resulting in 24 questions per test. Each item is scored 0 to 2, and the total test score is based on 20 operational items for a maximum possible score of 40. Each student is required to complete all items on the assessment.

The question clusters have the following characteristics:

- The range of abilities of students taking the assessment is factored in across all questions within a cluster.
- The four questions are scaffolded based on the grade level of the prerequisite skill, the difficulty of the skill, and what the student is being asked to do.
- Each of the question types within a cluster vary in difficulty from question to question and essence statement to essence statement.
- The first question is always the easiest of the four questions in a cluster, moving to the last and most cognitively complex question.
- The cluster design requires the student to make six concept transitions throughout the test rather than one per item.

TEKS Curriculum Frameworks for STAAR Alternate 2

Alignment with the state curriculum through federal and state laws is a critical requirement for STAAR Alternate 2. The No Child Left Behind Act (NCLB, 2001) mandates that alternate assessments must be aligned with the state's challenging academic content standards and challenging academic achievement standards. The Texas Education Code (TEC), Chapter 39 .023, Subsection A lists the subject areas and grades to be tested in the statewide student assessment program. Considering the elements of federal and state law, the Texas Education Agency (TEA) developed vertical alignment and curriculum framework documents to help students with significant cognitive disabilities access the grade-level TEKS curriculum. The two alignment resource documents help ensure that all students eligible to take an alternate assessment based on alternate achievement standards are instructed and assessed on curriculum that is linked to grade-level content. Through the processes illustrated in Figure 1.1 and described below, TEA aligned the STAAR Alternate 2 assessment to the grade-level TEKS curriculum.

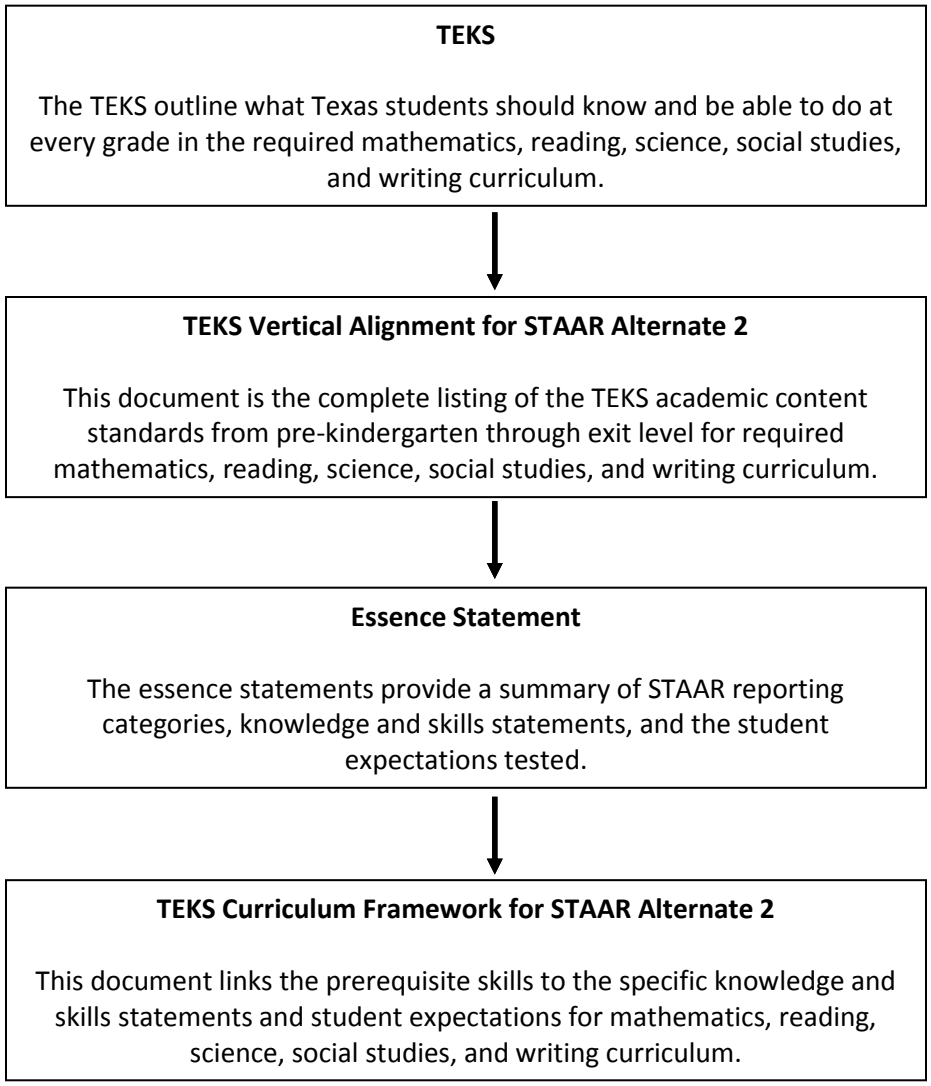


Figure 1.1 Access to the Grade-Level TEKS Academic Content Standards for Students with Significant Cognitive Disabilities

STAAR Alternate 2 Performance Standards

Federal statute requires any statewide assessment used for accountability (AYP) purposes to include at least three performance levels. In order to obtain at least three performance levels, any STAAR Alternate 2 assessment used for accountability needs to have at least two cut scores, or performances standards: one that distinguishes the "Level I" and "Level II" performance levels and one that distinguishes the "Level II" and "Level III" performance levels. In addition, TEC requires the establishment of specific performance standards for all STAAR assessments. For all assessments, there must be a cut score indicating satisfactory performance.

The STAAR Alternate 2 performance levels are:

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

Level II and Level III scores represent passing scores whereas Level I represents a failing score for purposes of accountability.

Chapter 2: Overview of STAAR Alternate 2 Standard-Setting Process

This chapter provides an overview of the STAAR Alternate 2 standard-setting process and includes the following sections:

- Goals of Setting Performance Standards
- Evidence-Based Standard Setting
- The STAAR Alternate 2 Standard-Setting Process

Goals of Setting Performance Standards

A critical aspect of any statewide testing program is the establishment of performance levels that provide a frame of reference for interpreting test scores. Once an assessment is administered, students, parents, educators, administrators, and policymakers want to know, in clear language, how students performed on that assessment. In general, by relating test performance directly to the student expectations expressed in the state curriculum in terms of what content and skills students are expected to demonstrate upon completion of each grade or course, performance standards describe the level of competence students are expected to exhibit.

Evidence-Based Standard Setting

Standard setting for STAAR Alternate 2 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 performance on other assessments. Standard-setting advisory panels composed of diverse groups of stakeholders considered the interaction of these elements for each STAAR Alternate 2 assessment.

Figure 2.1 illustrates the critical elements of this evidence-based standard-setting approach used by Texas to establish the STAAR performance standards (O'Malley, Keng, & Miles, 2012).

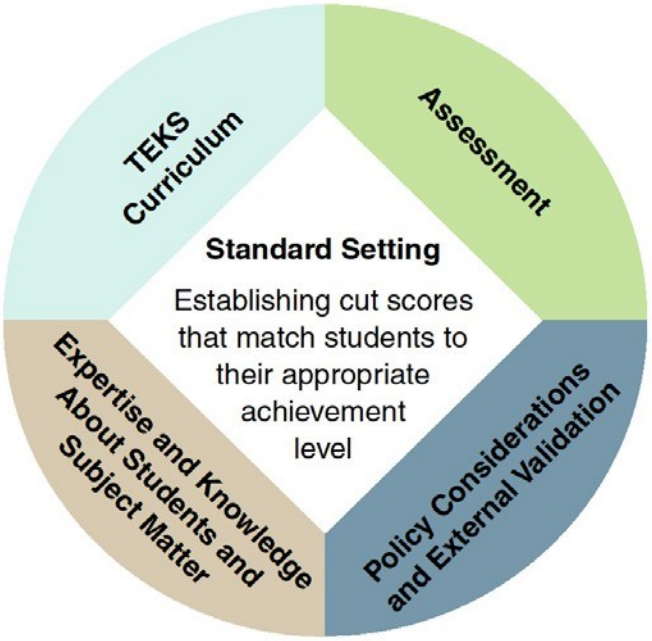


Figure 2.1: Critical Elements of the Evidence-Based Standard-Setting Approach

Each element of the evidence-based standard-setting approach as it relates to the STAAR Alternate 2 assessments is described below.

TEKS Curriculum Standards

The STAAR Alternate 2 TEKS Curriculum Framework documents list the reporting categories, knowledge and skills statements, and student expectations assessed on the STAAR Alternate 2. Each reporting category statement, knowledge and skills statement, and student expectation have been summarized into an essence statement that serves as the connection between the grade-level TEKS and the STAAR Alternate 2 assessments. The curriculum framework document can be used in conjunction with the STAAR Alternate 2 TEKS Vertical Alignment documents. The vertical alignment documents provide a complete listing of the TEKS curriculum from pre-kindergarten through high school. These documents provide a total overview of the knowledge and skills statements and align student expectations across grades.

The student expectations provide access points to the general education curriculum by serving as prerequisite skills for the STAAR Alternate 2. The curriculum framework documents also list access points in the form of prerequisite skills that link to the student expectations on the vertical alignment documents for each grade/subject or high school course. The student expectations provide the underlying basis for several key components of the standard-setting process, including the performance labels, policy definitions, and specific performance level descriptors (PLDs).

Assessment

Each STAAR Alternate 2 assessment is based on the student expectations and reporting categories specified in the STAAR Alternate 2 TEKS Curriculum Frameworks document.

Policy Considerations and External Validation

The STAAR Alternate 2 standard-setting process incorporated information from pilot administrations of STAAR Alternate 2, content judgments of student performance and expectations, policy considerations, and operational data from the first administration of STAAR Alternate 2.

Expertise and Knowledge about Students and Subject Matter

Texas educators, including classroom teachers and curriculum specialists from special education and general education backgrounds, brought content knowledge and classroom experience to the standard-setting process. They played an integral role in recommending the performance standards and served as advisers in reviewing the performance labels, policy definitions, and specific performance level descriptors.

Standard Setting

Within the framework of evidence-based standard setting an established standard-setting method known as the Extended Angoff approach (Hambleton & Plake, 1995) was used to recommend the cut scores, or performance standards.

The STAAR Alternate 2 Standard-Setting Process

To fulfill legislative requirements, a standard process was followed in order to establish performance standards for the STAAR Alternate 2 assessment program. To learn more about the standard-setting process for the general STAAR program, please reference the STAAR Standard-Setting Technical Report. For STAAR Alternate 2, the STAAR policy committee discussions served as a foundation for setting standards on the alternate assessment; however, STAAR Alternate 2 stakeholders familiar with the tested population and the new assessment were also convened to review information gathered from STAAR stakeholder meetings and determine its applicability to the STAAR Alternate 2 standard-setting process.

Table 2.1 provides high-level descriptions and timelines for the steps in the STAAR Alternate 2 standard-setting process. Each step is described in detail in the remaining chapters of this report.

Table 2.1 STAAR Alternate 2 Assessment Program Standard-Setting Process Overview

Standard-Setting Steps	Description	Timeline
1. Conduct empirical studies	Evaluate data from pilot and operational administrations of STAAR Alternate 2	Spring 2015
2. Develop performance labels and policy definitions	At the inception of the STAAR assessment program the TEA and the Texas Higher and Education Coordinating Board (THECB) jointly convened a committee 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. The performance categories and their labels were carried forward to the new STAAR Alternate 2 assessment. The policy definitions were reviewed and revised by a committee of stakeholders convened by TEA in January 2015.	January 2015
3. Develop grade/course specific PLDs	TEA created the first draft of 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 content area (mathematics, reading, science, and social studies).	January 2015
4. 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 2 assessment. These committees also provided comments to assist TEA with finalizing the specific PLDs.	April 2015
5. Review performance standards for reasonableness	TEA reviewed the cut-score recommendations across content areas.	April 2015
6. Approve performance standards	The Commissioner of Education approved the STAAR Alternate 2 performance standards.	April 2015
7. Implement performance standards	Performance standards were reported to students based on the spring 2015 administration.	May 2015

Chapter 3: Preparation for Standard Setting

This chapter provides details about preparation for the STAAR Alternate 2 standard-setting process. The first part of this chapter outlines analyses of operational data as well as preparation of information related to reasonable ranges wherein cut scores might be recommended. The second part of this chapter outlines the types of information presented to panelists during each round of standard setting. The sections in this chapter include:

- Purpose of Analyses Performed in Preparation for Standard Setting
- STAAR Alternate 2 Cut Score Reasonable Ranges
- Data Used during Standard Setting

Purpose of Analyses Performed in Preparation for Standard Setting

Because STAAR Alternate 2 was redesigned as a standardized item-based assessment, as opposed to the previous performance-based assessment, it was necessary to consider the impact the change might have on student performance and to determine what the expectations were for student performance. In updating the test design for STAAR Alternate 2, there was no intent to increase the rigor of the assessment beyond the intended change in rigor with the implementation of the STAAR assessment program in 2012. More specifically, the purpose of the redesign was to standardize the test administration and reduce the burden on test administrators to create assessment materials; the purpose was not to increase the rigor of the content or performance standards. Therefore, in preparation for standard setting, the percentage of students in each performance level from the 2014 administration was used as one of the data points informing the development of guiding principles for reasonable ranges within which cut scores might fall. Impact data from 2014 were compared to student performance from 2015 in order to devise guiding principles that did not result in significant shifts in the percentage of students in each performance level across the grades and subjects.

STAAR Alternate 2 Cut Score Reasonable Ranges

The standard-setting process used for most of the STAAR assessments included the development of reasonable ranges wherein the cut score recommendations might fall. Through the use of impact data from 2014 and student performance in 2015, as well as judgments from content experts about the expectations for student performance, a reasonable range (or neighborhood) was identified in which it made sense for the passing standard to be set. Panelists use this additional information while making their cut score recommendations.

Expert judgments of the test and knowledge of students' abilities and expectations for performance were used in establishing reasonable ranges. A committee of general education and special education experts familiar with the curriculum and the characteristics of the student population was convened prior to the standard-setting meeting. These experts were trained in the standard-setting methodology and completed one round of cut score recommendations in order to establish reasonable ranges. This committee completed one round of judgements for all 22 assessments in the STAAR Alternate 2 assessment program. Recommendations from this committee served to establish a potential lower boundary cut score for each performance level.

Information regarding item difficulty and student performance from the operational administration was combined with the committee recommendations for the reasonable range in order to establish the final set of reasonable ranges for all tests in the program. In addition, a comparison to the previous passing standard for

STAAR Alternate in 2014 was considered in an effort to prevent a standard being set that far exceeded or fell far short of the 2014 standard. A set of guiding principles was used to ensure that reasonable ranges were established uniformly across all assessments. The guiding principles are shown below.

Level II: Satisfactory

1. The *Satisfactory* lower boundary should be
    a. at the minimum *Satisfactory* raw score from the expert committee; or
    b. at the raw score aligned with 2014 STAAR Alternate impact data for Level II, whichever is the higher raw score.
2. The *Satisfactory* upper boundary will be below the *Accomplished* lower boundary.

Level III: Accomplished

3. The *Accomplished* lower boundary should be
    a. at the median *Accomplished* raw score from the expert committee; or
    b. at the raw score between the median *Accomplished* raw score from content review and the raw score aligned with 2014 STAAR Alternate impact data for Level III.
4. The *Accomplished* upper boundary should be below raw scores with a standard error of theta greater than or equal to 1.00.

As previously mentioned, the reasonable ranges were presented to panelists; however, the reasonable ranges were not presented until round 2 recommendations. As such, panelists' initial recommendations might have fallen outside of the reasonable range. Panelists were encouraged to make their recommendations within the ranges; however, they were permitted to recommend cut scores anywhere along the raw score scale continuum if they could provide a content justification.

Data Used during Standard Setting

Reasonable ranges were one piece of information used during the standard-setting meetings. In addition, panelists were presented with information regarding item difficulty as well as impact data from the 2015 operational administration showing how students performed on the test overall. Item difficulty was presented during the second round of judgments and was presented in terms of the average score per item obtained by students during the 2015 operational administration. Performance was presented during the third round of judgments and was presented in terms of the percent of students who would fall into each performance level if the cut scores were set at the recommendation resulting from round 2 judgments.

Chapter 4: Performance Labels, Policy Definitions, PLDs, and Borderline Descriptors

This chapter provides details about the development of performance labels and policy definitions, as well as the development of grade/course specific PLDs and borderline descriptors (BDs). This chapter includes the following topics:

- Purpose of Performance Labels and Policy Definitions
- Development of Performance Labels and Policy Definitions
- Purpose of the Specific PLDs
- Development of the Specific PLDs
- Texas Educator Review of the Specific PLDs
- Development of BDs
- Review and Approval Process for the Specific PLDs

Purpose of Performance Labels and Policy Definitions

The purpose of the performance labels and policy definitions is to describe the general level of knowledge and skills evident at each performance level across all content areas and grades/courses. The policy definitions are explicit definitions of what students must do to demonstrate proficiency at a specific level on the content standards. During the standard-setting committee meetings, these labels and definitions provided the panelists with a consistent baseline as they developed recommendations for the cut scores associated with each performance standard.

Development of Performance Labels and Policy Definitions

As part of the implementation of the STAAR assessment program, TEA, in conjunction with the THECB, convened a Performance Descriptor Advisory Committee (PDAC) to make recommendations for the performance labels and policy definitions that define the performance standards for STAAR. The committee was comprised of individuals representing the diversity of stakeholders in public education and higher education in Texas. Following the meeting, TEA and THECB staff members considered the committee's recommendations for performance labels and used the recommendations to draft the policy definitions. The PDAC reviewed these performance labels and policy definitions prior to final review and approval by the commissioner of education. The policy labels for the general STAAR assessments are Level III: Advanced Academic Performance, Level II: Satisfactory Academic Performance, and Level I: Unsatisfactory Academic Performance.

A team of special education content experts from TEA and Pearson used the approved STAAR performance labels listed above and policy definitions to develop the STAAR Alternate 2 performance labels and policy definitions. The STAAR labels and definitions were reviewed and carefully considered by the team who made one change for the labels—for STAAR Alternate Level III would be changed to Accomplished Academic Performance instead of Advanced Academic Performance. Performance expectations are defined differently for STAAR Alternate 2 because it is an assessment based on alternate performance standards, is linked to grade-level content, and covers a different depth, breadth, and complexity from the general assessment performance standards.

In addition to the different performance label at Level III, a change from the STAAR Level I: Unsatisfactory Label resulted from the previous meeting of the STAAR Alternate standard-setting committees in 2012. During the first STAAR Alternate standard-setting meetings, the panelists strongly recommended changing the "Unsatisfactory" performance label for the STAAR Alternate assessments. The panelists' concerns were two-fold: 1) labeling the performance of students with the most significant cognitive disabilities as "Unsatisfactory" is biased against a group of students who, through no fault of their own, have the most severe intellectual disabilities; and 2) reporting "Unsatisfactory" on the Confidential Student Reports to parents of students with the most significant cognitive disabilities may be extremely distressing, as these results communicate that their child's performance is inadequate and unacceptable, even though that performance is due to limitations imposed by the severity of the disability.

Perie, Hess, and Gong (2008) state that PLDs can greatly influence instruction and teacher and parent expectations. They recommend thinking holistically about how a student moves from one performance level to the next and one grade to the next. They also indicate that although most states use the same performance level labels for the alternate assessments as they do for their general assessments, other states have created new performance labels. As a result of the standard-setting committees' recommendation, PLD literature review, and research of other states' PLDs, Level I for STAAR Alternate was changed to Developing Academic Performance. These changes carried through with the redesign of the alternate assessment.

For STAAR Alternate 2 assessments, the performance labels are

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

The policy definitions from the previous version of the alternate assessment were also revised to more accurately reflect the new standardized assessment rather than a performance-based assessment. The revised policy definitions were reviewed and revised by a committee of general and special education experts prior to the standard-setting meeting. Below are the policy definitions for STAAR Alternate 2.

Level III: Accomplished Academic Performance

Performance in this category indicates that students are well prepared for the next grade or course with instructional supports for accessing the curriculum through prerequisite skills. Students demonstrate a strong understanding of the knowledge and skills that are linked to content measured at this grade or course. Students exhibit the ability to use higher-level thinking and more complex skills, which includes making inferences, comparisons, and solving multi-step problems. With support, students in this category have a high likelihood of showing progress in the next grade or course.

Level II: Satisfactory Academic Performance

Performance in this category indicates that students are sufficiently prepared for the next grade or course with instructional supports for accessing the curriculum through prerequisite skills. Students demonstrate sufficient understanding of the knowledge and skills that are linked to content measured at this grade or course. Students exhibit the ability to determine relationships, integrate multiple pieces of information, extend details, identify concepts, and match concepts that are similar. With continued support, students in this category have a

reasonable likelihood of showing progress in the next grade or course.

Level I: Developing Academic Performance

Performance in this category indicates that students are insufficiently prepared for the next grade or course and need additional instructional support for accessing the curriculum through prerequisite skills. Students are able to acknowledge concepts, but they demonstrate a minimal or inconsistent understanding of the knowledge and skills that are linked to content measured in this grade or course. Even with continued support, students in this category are in need of significant intervention to show progress in the next grade or course.

Purpose of the Specific PLDs

Specific PLDs outline the degree of knowledge and skills required of each performance level of a test given for a specific grade or course. The PLDs developed for STAAR Alternate 2 provide a snapshot of students' academic characteristics based on performance on a given STAAR Alternate 2 assessment and reflect the content and skills evident in the curriculum standards, the TEKS. The STAAR Alternate 2 specific PLDs were written prior to and used for standard setting.

As a component of the standard-setting process, specific PLDs served to anchor training activities and guide committee members by establishing a common understanding of expected performance on each STAAR Alternate 2 assessment. PLDs helped ground standard-setting committee members in the content standards and guide them as they made their recommendations for the scores needed to achieve Level II and Level III on each STAAR Alternate 2 assessment. In addition to their use in standard setting, specific PLDs have been published to serve as a tool for classroom instruction and to help educators interpret student performance on the assessments. PLDs can enhance parents' understanding of their child's academic strengths and weaknesses and can help the community at large better understand state test scores and the level of performance required of students on STAAR Alternate 2 PLDs are also a requirement of the U.S. Department of Education (USDE) in their review and approval of state assessments.

Development of the Specific PLDs

The same TEA and Pearson team that reviewed the STAAR Alternate 2 performance labels and policy definitions also drafted and reviewed the STAAR Alternate 2 specific PLDs for each content area and grade/course. Other resources that were used to draft the STAAR Alternate 2 PLDs were the essence statements and the curriculum framework documents which include the prerequisite skills for each essence statement. The specific PLDs were developed to link to grade-level content standards and show a progression of performance across the grades and achievement levels. The guiding criteria for developing the draft specific PLDs were as follows.

- The specific PLDs should focus on and reflect the expectations outlined in the performance labels and policy definitions.
- The specific PLDs link to the grade-level knowledge and skill statements and student expectations through prerequisite skills.
- Each reporting category should be represented in the specific PLDs.
- Each essence statements identified for assessment should be reflected in the specific PLDs.
- Each performance level should reflect a range of skills within the specific PLD as well as reflecting an increased expectation moving from Level I: Developing to Level III: Accomplished.

- The content skill for each descriptor should be accessible to all students in the tested population regardless of the disability.

Texas Educator Review of the Specific PLDs

Two educator committees reviewed the draft specific PLDs for each grade and subject and high school course. The PLD committee was convened to review and revise the specific PLDs in an effort to have a more final draft of the PLDs for the standard-setting meeting. This committee was a subset of the larger standard-setting committee. The standard-setting committee had an additional opportunity to review and provide feedback on the specific PLDs. Both committees were comprised of general and special educators. The general education teachers provided content expertise and verified that the performance level descriptors were aligned to the TEKS grade-level standards. The special education teachers provided knowledge of the learning needs of students with the most significant cognitive disabilities and verified that the specific PLDs were appropriate for the students. In the process of reviewing the specific PLDs, committee members considered the performance labels and policy definitions and the assessed curriculum. In addition to reviewing the PLDs for each specific subject and course, the committees reviewed the PLDs across grades and courses within a content area for reasonableness. Table 4.1 summarizes the committee composition for the entire committee that reviewed specific PLDs. The demographic characteristics of the larger standard-setting committee appears in Appendix 1 and is broken out by grade and subject-specific committees.

Table 4.1 Specific PLD Review Committee Composition

Gender		**Ethnicity**		**Position**	
Male	12	Native American	0	Teacher, General	13
Female	52	Asian/Pacific Islander	0	Teacher, Special Education	41
		African American	6	Administrator	5
		Hispanic	14	Other	4
		White	41		

Development of Borderline Descriptors

Development of BDs is one of the most challenging and time-consuming tasks that occurs during a standard-setting meeting. In an effort to devote as much time to this task as possible, the drafting of BDs was completed during the specific PLD committee meeting. Nine committees reviewed specific PLDs and drafted BDs. The nine committees were organized to match the structure of the subsequent standard-setting meetings. Specifically, the following structure of grades and subjects was used during the specific PLD meetings and the standard-setting meetings.

- English I and English II
- Reading grades 3–5
- Reading grades 6–8
- Writing grades 4 and 7
- Mathematics grades 6 and 7
- Mathematics grade 8 and Algebra I
- Mathematics grades 3–5
- Social Studies grade 8 and U.S. History
- Science grades 5, 8, and Biology

The review of specific PLDs and drafting of BDs followed a typical standard-setting meeting in that the committee was provided context for the use of PLDs and BDs. Table groups reviewed and provided feedback on the specific PLDs and then shared their feedback with the entire committee, and the entire committee agreed on the final terminology of the PLDs. The committee was then provided context on how the specific PLDs would be used during the standard-setting meeting and provided training on how to develop BDs. After completing the training on how BDs are generated, the committees broke into their small table groups and focused on drafting a BD for each performance level, starting with the Level II: Satisfactory Academic Performance level and then proceeding to Level III: Accomplished Academic Performance. When table groups had completed their drafts, table groups reconvened as an entire committee and reviewed and discussed their drafts. Edits and revisions were made to the BDs when the committee reached consensus on the concepts and terminology contained within a BD. As mentioned previously, these BDs were later used during the standard-setting committee meetings.

Review and Approval Process for the PLDs

Following the standard setting, TEA staff provided a final review of the specific PLDs verifying that any changes that might have occurred during the standard-setting meeting matched the performance expectations evident in the policy definitions. Specific PLD feedback from the standard-setting committees primarily reflected clarifications that the committees found useful in their discussions about student performance during the standard-setting process. The final PLDs can be found under STAAR Alternate 2 Performance Level Descriptors (PLDs) on the TEA website: http://tea.texas.gov/Student_Testing_and_Accountability/Testing/STAAR_Alternate/STAAR_Alternate_2_Performance_Level_Descriptors/.

Chapter 5: Standard Setting for STAAR Alternate 2

This chapter provides specific details about the standard-setting process and results of the standard-setting meetings. The chapter includes the following topics:

- Purpose of Standard-Setting Meetings
- Committee Composition and Attendees
- Implementing the Standard-Setting Process
- Standard-Setting Results
- Standard-Setting Process Evaluation Results

Purpose of Standard–Setting Meetings

All standard setting is based to a large degree on educator judgment. Panelists use their experience and knowledge to make expert recommendations. These judgments help establish the criteria for interpreting test scores using a specific standard-setting method. The purpose of holding STAAR Alternate 2 standard-setting meetings was to gather expert recommendations for the performance standards on each assessment.

Each committee was asked to recommend cut scores for Level II: Satisfactory Academic Performance and Level III: Accomplished Academic Performance using the following types of information:

- Content of the STAAR Alternate 2 assessments,
- Performance labels and policy definitions,
- Specific PLDs for each assessment,
- BDs
- Reasonable ranges (or neighborhoods) within which the cut scores should fall, and
- Impact data.

Committee Composition and Attendees

Nine standard-setting committees were convened. Table 5.1 shows the breakdown of assessments that were included in each meeting. Each committee recommended Level II: Satisfactory Academic Performance and Level III: Accomplished Academic Performance cut scores for each assessment covered in the meeting. For example, the grades 6–8 reading committee recommended six cut scores, two for each of the three assessments covered during the meeting.

Table 5.1: STAAR Alternate 2 Standard-Setting Meetings Schedule

Meeting	Dates	Number of Days
English I, English II	04/7/15 – 04/8/15	2
Writing Grades 4 & 7		
Biology Science Grades 5 & 8	04/7/15 – 04/9/15	3
Reading Grades 3–5		
Reading Grades 6-8		
Mathematics Grades 3–5		
Mathematics Grades 6–7	04/9/15 – 04/10/15	2
Algebra I Mathematics Grade 8		
U. S. History Social Studies Grade 8		

When selecting standard-setting panelists TEA placed an emphasis on experience with the population of students for which STAAR Alternate 2 is appropriate, as well as content knowledge and classroom experience. TEA also attempted to construct committees whose demographic representation reflected that of the alternate assessment's student population. As mentioned previously, the subset of educators who were invited to review and revise specific PLDs and draft BDs were also invited to attend the standard-setting meetings. Appendix 1 provides information regarding the characteristics and experience of the panelists on each standard-setting committee. These tables provide demographic information about the committee members, as well as information about the members' current positions in education, the number of years they have been in their positions, their experience working with the various types of student populations, and the types of districts they represent.

Implementing the Standard-Setting Process

The approach used for setting standards on the STAAR Alternate 2 assessments was an extended Angoff procedure (Angoff, 1971, Hambleton & Plake, 1995). During the standard-setting meeting, panelists engaged in various activities to familiarize themselves with the purpose of the meeting and the assessment itself. Meeting facilitators discussed the purpose of setting academic performance standards. Discussions focused on the idea that standard setting provides a systematic process for eliciting expert judgments from educators who understand both the content standards and the students who take the assessment. Panelists completed three rounds of judgments. During each round panelists were provided with data to inform their recommendations and feedback after the round to help inform their next set of recommendations. Specific instruction was provided with each round regarding the task the panelists were engaged in and the type of information that was shared at each round was described in detail. Panelists were told that their judgments ultimately related to the number of score points a borderline student would be required to obtain on STAAR Alternate 2 to achieve Level II or Level III performance standards. During the standard-setting process, panelists reviewed a variety of

resources related to academic performance, including policy definitions, specific PLDs, assessment items, and BDs. The remaining sections describe each specific step panelists followed during the standard-setting meetings.

Review Assessment Booklets

Panelists reviewed both the teacher and student test booklets for their assigned subject and grade/course. The process of reviewing the assessment items allowed panelists to become familiar with what students are asked to do during the assessment.

Each content-area specific assessment is comprised of six clusters of items. Each cluster contains four items progressing in difficulty from easiest to hardest. All items within a cluster are aligned to the same essence statement and utilize the same context in order to aid in student engagement with the items. Items are scored polytomously with a maximum of two points possible. A student receives maximum credit when answering the item correctly and independently. The student receives partial credit if the teacher has to provide the scripted assistance in order for the student to respond correctly. Panelists focused on the differences between the items and the type of teacher assistance built into the script for each item.

Review Specific PLDs and BDs

After reviewing the assessment booklets, panelists were given an opportunity to review and discuss the specific PLDs and BDs. Panelists reviewed the PLDs and BDs specific to each assessment after reviewing the assessment booklets. Panelists were asked to focus on what students should know and be able to do for a specific content area while taking into account that students with significant cognitive disabilities are a uniquely varied population of students. Panelists engaged in group discussions about the types of academic skills and levels of proficiencies that students meeting those definitions should demonstrate. Panelists were asked to form a conceptualization of what academic performance for a minimally capable or borderline student looks like. These conceptualizations were used throughout the standard-setting process to help panelists make judgments about the type of student described in each performance level and what academic behaviors differentiate students in different performance levels.

Round 1 Judgments

After reviewing the assessment booklets, specific PLDs, and BDs, panelists received training on the standard-setting process. Panelists were then guided to review every item while thinking about a borderline Level II: Satisfactory student and asked to indicate what score that borderline student should obtain on each item. Then they were asked to repeat that task for the borderline Level III: Accomplished student. Although reasonable ranges were created within which cut scores might fall, the first round of judgments was a pure content judgment using only specific PLDs and borderline descriptors as the foundation. Reasonable ranges were introduced in Round 2.

Round 1 Feedback

After Round 1, panelists were provided with two pieces of information. First, panelists were shown information about how their own judgments of the number of score points a borderline student is likely to obtain on a task compares with other panelists' judgments. Panelists were provided individual feedback data showing their recommended cut score, table-level feedback data, and group-level feedback data. Both table-level and group-level feedback data included the average and median score assigned by the group as well as the standard deviation, minimum, and maximum. Second, in preparation for Round 2 discussions, feedback also included a

total score comparison of panelists' judgments. Panelists were also provided information about the difficulty of items, represented by the average score point obtained on each item (ranging from 0 to 2) during the spring 2015 administration. Panelists were encouraged to discuss their ratings with their table and then to share their conversations with the committee as a whole.

Round 2 Judgments

To prepare for Round 2 judgments, the feedback from Round 1 was referenced along with additional discussion of the reasonable ranges within which a cut score might fall and the guiding principles for how those reasonable ranges were created. Panelists were asked to reflect on how their recommended cut scores did or did not align with the reasonable ranges. Round 2 judgments were made after panelists were comfortable with their understanding of round 1 feedback and the new reasonable range information.

Round 2 Feedback

Feedback for Round 2 included individual cut score results, table-level, and group-level summary statistics, consistent with Round 1 feedback. In addition, panelists were shown impact data from their Round 2 judgments. Impact data were calculated using data from the spring 2015 STAAR Alternate 2 administration. The percentage of students within each performance level was calculated for all students tested. Impact data were presented for the overall group of test-takers, as well as by gender, by ethnicity, and by economically disadvantaged status. Impact data were presented for each performance level. Panelists were asked to review the impact data for reasonableness in terms of the recommended cuts, policy definitions, and specific PLDs.

Round 3 Judgments

Once panelists reviewed the impact data, they were asked to make their final cut score recommendations. Appendix 2 shows the results of the standard-setting meeting by round of judgments for each committee.

Cross-Course Articulation

The final component of the standard-setting meeting involved judgments about the reasonableness of cut scores across subjects and grades or courses within a subject area. Committee members from each individual meeting were asked to contribute to the discussion of cross-course articulation across elementary, middle, and high school assessments. Both the cut score ranges that were provided to panelists as well as their final median cut score recommendations from Round 3 were presented during cross-course articulation. In addition, panelists were shown the impact data associated with the Round 3 cut score recommendations for each assessment.

During cross-course articulation, panelists were asked to review the recommended cut scores across subjects and grades or courses within a subject area with respect to the policy definitions, specific PLDs, and impact data. In light of their review of this information and group discussions across the subjects and grades or courses, panelists were asked to make recommendations about potential shifts in cut scores. Panelists set 44 cut scores throughout the course of the meetings. During articulation, only 16 of those cut scores were adjusted. The changes were typically made with the rationale that impact data across grades within a subject area should be more similar. Panelists also justified their changes based on the ranges of cut score recommendations throughout the standard-setting meeting. This final judgment by panelists during the cross-course articulation served as the final recommendation of cut scores from the standard-setting meeting to TEA.

Evaluation Survey

At the conclusion of the standard-setting meetings, each panelist completed an evaluation form related to the usefulness and success of each of the standard-setting activities. Panelists had an opportunity to provide feedback about their confidence in the final recommended cut score judgments, the training, time allowed for making judgments, and the types of information shared during the meetings. The survey information is used to support the quality of the design and implementation of the standard-setting procedure (Kane, 2001). Procedural validity evidence comes from indications that the steps proposed and supported by best practice were followed based on survey responses by the panelists.

Standard-Setting Process Evaluation Results

Full results from the completed process evaluations are provided in Appendix 4. In general, panelists across all meetings gave positive feedback on the standard-setting sessions and activities. The majority of panelists felt that their professional opinions were valued and that the three standard-setting rounds were sufficient for them to make judgments. Panelists' responses were more varied with regard to their confidence in the specific PLDs and their confidence in their Round 3 judgments. Even with this increased variation, the majority of responses indicated that panelists were either "confident" or "very confident" with the specific PLDs and their Round 3 judgments.

Meeting Success

The standard-setting panelists were asked to give their opinion on the success of the components of the standard-setting sessions. Across all meetings, the sessions were rated as "successful" or "very successful" by between 94% and 100% of the panelists. The science and social studies committees had 1 respondent, out of 18, who indicated the overview of STAAR Alternate 2 portion of the meeting was only partially successful. The English language arts committees had 1 respondent, out of 41, who indicated the overview of the standard-setting process was only partially successful.

Usefulness of Activities and Information

The standard-setting panelists were asked to judge how useful the standard-setting activities or information were in helping them make their recommendations for the STAAR Alternate 2 cut scores. Panelists had positive ratings across the meetings where 89% to 100% of panelists indicated that the activities and information used during the meeting were "useful" or "very useful."

Adequacy of the Meeting Elements

The standard-setting panelists were asked to rate the adequacy of various elements of the meeting including training provided, amount of time spent training, feedback provided between rounds, facilities used for the session, total amount of time for the ratings, and number of rounds for judgments. For all the subject area meetings, the elements were rated as "more than adequate" or "adequate" by between 98% and 100% of the panelists.

Confidence in Specific PLDs.

The standard-setting panelists were asked to give their opinion about how confident they felt in regard to how reasonable each of the specific PLDs was for each student performance level of the STAAR Alternate 2 assessment. Panelists' confidence in the specific PLDs showed lower ratings than other categories within the

survey. The majority of ratings indicate panelists were either confident or very confident in the specific PLDs, between 91% to 100% of panelists, with an even split between "confident" and "very confident" ratings. The lowest rating occurred in the mathematics committees with 9% of panelists indicating they were only "somewhat confident" with the Level I: Developing Academic Performance descriptor.

Confidence in Round 3 Ratings.

The panelists were asked to give their opinion about how confident they felt in regard to how well their Round 3 ratings represented the appropriate levels of student performance for the Satisfactory Academic Performance cut score and the Accomplished Academic Performance cut score. For the Round 3 recommended cut scores, between 94% and 100% of the panelists were "confident" or "very confident" across Level II and Level III performance categories.

Opportunities to Collaborate.

The standard-setting panelists were asked whether they believed they had adequate opportunities during the session to collaborate and express their professional opinions about certain aspects of the standard-setting process. Across all the subject area meetings, 100% of panelists indicated they had "adequate" or "more than adequate" opportunities during the sessions to express their opinions about various aspects of the process.

Respect.

The panelists were asked if they believed their opinions and judgments were treated with respect by their fellow judges and the facilitators. Committee responses indicated that between 97 and 100% of panelists felt as though they were treated with respect by their fellow panelists as well as by their facilitator. Responses less than 100% were for two committees where 1 panelist indicated they only sometimes felt they were treated with respect by their fellow panelists.

Chapter 6: Reasonableness Review, Approval, and Implementation of Performance Standards

This chapter provides details about the reasonableness review that took place following the STAAR Alternate 2 standard-setting meetings as well as the implementation of the standards. The sections in this chapter include:

- Purpose of Reasonableness Review
- Rationale for Adjustments Made During Reasonableness Review
- Implementation of Standards

Purpose of Reasonableness Review

After educator committees recommended Level II and Level III performance standards for the STAAR Alternate 2 assessments, TEA conducted reasonableness reviews of the cut score recommendations within and across content areas. The reasonableness review process following standard setting was intended to confirm that performance standards contribute to a well-articulated and coherent assessment system.

Rationale for Adjustments Made During Reasonableness Review

During the reasonableness review, TEA evaluated the results from all the standard-setting committees (see Appendix 2 and Appendix 3). The following pieces of information were considered:

- The content discussions that occurred during the standard-setting meetings;
- The alignment of content expectations to the performance level descriptors and policy definitions, both within and across content areas;
- Round 3 judgments of the standard-setting committees, including;
    - the impact data associated with the Round 3 judgments, and
    - the relationship of round 3 judgments to the guiding principles that established the reasonable ranges; and
- Changes recommended during the cross-course articulation, including:
    - the impact data associated with the articulation judgments
    - the relationship of articulation judgments to the guiding principles that established the reasonable ranges.

The review of this information was used to determine the final cut score recommendations for Level II: Satisfactory Academic Performance and Level III: Accomplished Academic Performance. Out of the 44 cut scores recommended by the standard-setting committee, adjustments were made to 7 of the cut scores. Half the adjustments made resulted in a final recommended cut score that was more similar to the panelists' recommendations after round 3 judgments but before vertical articulation. The other half of the adjustments were made to better align the impact data across grades or subject areas. Appendix 3 shows impact data using the final set of recommended cut scores that were presented by TEA to the Commissioner of Education.

Implementation of Standards

The standards set in April 2015 were approved by the Texas Commissioner of Education soon after the standard setting meetings were held. The standards apply to all students taking STAAR Alternate 2 in spring 2015 and

beyond. The final cut scores that were implemented in spring 2015 are shown in Table 6.1

Table 6.1: STAAR Alternate 2 Final Performance Standards

Assessment	Level II: Satisfactory Academic Performance	Level III: Accomplished Academic Performance
Grade 3 Reading	300	381
Grade 4 Reading	300	384
Grade 5 Reading	300	387
Grade 6 Reading	300	371
Grade 7 Reading	300	371
Grade 8 Reading	300	379
Grade 4 Writing	300	363
Grade 7 Writing	300	359
English I	300	367
English II	300	366
Grade 3 Mathematics	300	375
Grade 4 Mathematics	300	387
Grade 5 Mathematics	300	379
Grade 6 Mathematics	300	373
Grade 7 Mathematics	300	375
Grade 8 Mathematics	300	365
Algebra I	300	361
Grade 5 Science	300	387
Grade 8 Science	300	382
Biology	300	383
Grade 8 Social Studies	300	372
U.S. History	300	368

References

Angoff, W.H. (1971). Scales, norms, and equivalent scores. In R.J. Thorndike (Ed.),*Educational measurement* (pp. 508-600). Washington, DC: American Council on Education.

Hambleton, R.K. & Plake, B.S. (1995). Using an extended Angoff procedure to set standards on complex performance assessments. *Applied Measurement in Education*, *8*, 41-56.

Kane, M. T., (2001). So much remains the same: Conception and status of validation in setting standards. In G.J. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives* (pp. 53-88). Mahwah, NJ: Erlbaum.

O'Malley, K., Keng, L., & Miles, J. (2012). Using validity evidence to set performance standards. In G. J. Cizek (Ed.), *Setting Performance Standards* (2nd ed., pp.301–322). New York: Routledge.

Perie, M., Hess, K., & Gong, B. (March 27, 2008). Writing Performance Level Descriptors: Applying Lessons Learned from the General Assessment to Alternate Assessments based on Alternate and Modified Achievement Standards. Paper presented at the Annual Meeting of the National Council on Measurement in Education.

Appendix 1: STAAR Alternate 2 Standard-Setting Committee Composition

Mathematics Grades 3–5

Years of Professional Experience in Education

Years	N-Count
1-5 years	0
6-10 years	3
11-15 years	0
16-20 years	4
More than 20 years	2

Position

Years	N-Count
Administrator	1
Special Education	7
General Education	1
Other	0

District Size

Type	N-Count
Large	4
Medium	3
Small	1
Did Not Respond	1

Ethnicity Distribution

Ethnicity	N-Count
African American	2
Asian or Pacific Islander	0
Hispanic	3
Multi-racial	0
Native American	0
White	7

Gender Distribution

Gender	N-Count
Female	7
Male	2

District Socioeconomic Status

Type	N-Count
High	1
Moderate	3
Low	3
Did Not Respond	2

Experience with Student Population

Student Population	N-Count
General Education	8
Special Education	9
English Language Learners	7
Low Socioeconomic Status	8

District Type

Type	N-Count
Metro	3
Suburban	3
Rural	2
Did Not Respond	1

Mathematics Grades 6 and 7

Years of Professional Experience in Education

Years	N-Count
1-5 years	3
6-10 years	1
11-15 years	5
16-20 years	3
More than 20 years	0

Position

Years	N-Count
Administrator	1
Special Education	9
General Education	2
Other	0

District Size

Type	N-Count
Large	5
Medium	3
Small	4
Did Not Respond	0

Ethnicity Distribution

Ethnicity	N-Count
African American	3
Asian or Pacific Islander	0
Hispanic	2
Multi-racial	0
Native American	0
White	9

Gender Distribution

Gender	N-Count
Female	10
Male	2

District Socioeconomic Status

Type	N-Count
High	1
Moderate	3
Low	8
Did Not Respond	0

Experience with Student Population

Student Population	N-Count
General Education	7
Special Education	11
English Language Learners	5
Low Socioeconomic Status	8

District Type

Type	N-Count
Metro	1
Suburban	6
Rural	5
Did Not Respond	0

Grade 8 Mathematics and Algebra I

Years of Professional Experience in Education

Years	N-Count
1-5 years	0
6-10 years	7
11-15 years	3
16-20 years	0
More than 20 years	4

Position

Years	N-Count
Administrator	2
Special Education	8
General Education	4
Other	0

District Size

Type	N-Count
Large	6
Medium	4
Small	4
Did Not Respond	0

Ethnicity Distribution

Ethnicity	N-Count
African American	2
Asian or Pacific Islander	0
Hispanic	3
Multi-racial	0
Native American	0
White	12

Gender Distribution

Gender	N-Count
Female	13
Male	1

District Socioeconomic Status

Type	N-Count
High	2
Moderate	6
Low	5
Did Not Respond	1

Experience with Student Population

Student Population	N-Count
General Education	10
Special Education	13
English Language	10
Low Socioeconomic	12

District Type

Type	N-Count
Metro	2
Suburban	6
Rural	6
Did Not Respond	0

Reading Grades 3–5

Years of Professional Experience in Education

Years	N-Count
1-5 years	0
6-10 years	2
11-15 years	1
16-20 years	2
More than 20 years	4

Position

Years	N-Count
Administrator	0
Special Education	6
General Education	2
Other	1

District Size

Type	N-Count
Large	4
Medium	3
Small	2
Did Not Respond	0

Ethnicity Distribution

Ethnicity	N-Count
African American	1
Asian or Pacific Islander	0
Hispanic	2
Multi-racial	0
Native American	0
White	8

Gender Distribution

Gender	N-Count
Female	7
Male	2

District Socioeconomic Status

Type	N-Count
High	0
Moderate	4
Low	5
Did Not Respond	0

Experience with Student Population

Student Population	N-Count
General Education	6
Special Education	9
English Language	6
Low Socioeconomic	8

District Type

Type	N-Count
Metro	1
Suburban	3
Rural	3
Did Not Respond	2

Reading Grades 6–8

Years of Professional Experience in Education

Years	N-Count
1-5 years	0
6-10 years	2
11-15 years	3
16-20 years	0
More than 20 years	2

Position

Years	N-Count
Administrator	1
Special Education	5
General Education	0
Other	1

District Size

Type	N-Count
Large	1
Medium	4
Small	2
Did Not Respond	0

Ethnicity Distribution

Ethnicity	N-Count
African American	1
Asian or Pacific Islander	0
Hispanic	2
Multi-racial	0
Native American	0
White	4

Gender Distribution

Gender	N-Count
Female	6
Male	1

District Socioeconomic Status

Type	N-Count
High	0
Moderate	2
Low	5
Did Not Respond	0

Experience with Student Population

Student Population	N-Count
General Education	5
Special Education	7
English Language	4
Low Socioeconomic	5

District Type

Type	N-Count
Metro	2
Suburban	3
Rural	2
Did Not Respond	0

Writing Grades 4 and 7

Years of Professional Experience in Education

Years	N-Count
1-5 years	1
6-10 years	7
11-15 years	3
16-20 years	1
More than 20 years	2

Position

Years	N-Count
Administrator	0
Special Education	7
General Education	6
Other	0

District Size

Type	N-Count
Large	1
Medium	8
Small	4
Did Not Respond	1

Ethnicity Distribution

Ethnicity	N-Count
African American	2
Asian or Pacific Islander	0
Hispanic	4
Multi-racial	0
Native American	0
White	11

Gender Distribution

Gender	N-Count
Female	11
Male	3

District Socioeconomic Status

Type	N-Count
High	1
Moderate	4
Low	8
Did Not Respond	1

Experience with Student Population

Student Population	N-Count
General Education	9
Special Education	14
English Language	11
Low Socioeconomic	13

District Type

Type	N-Count
Metro	0
Suburban	5
Rural	8
Did Not Respond	1

English I and II

Years of Professional Experience in Education

Years	N-Count
1-5 years	1
6-10 years	4
11-15 years	2
16-20 years	1
More than 20 years	3

Position

Years	N-Count
Administrator	1
Special Education	4
General Education	6
Other	0

District Size

Type	N-Count
Large	4
Medium	3
Small	4
Did Not Respond	0

Ethnicity Distribution

Ethnicity	N-Count
African American	3
Asian or Pacific Islander	0
Hispanic	2
Multi-racial	0
Native American	0
White	8

Gender Distribution

Gender	N-Count
Female	9
Male	2

District Socioeconomic Status

Type	N-Count
High	1
Moderate	2
Low	8
Did Not Respond	0

Experience with Student Population

Student Population	N-Count
General Education	11
Special Education	11
English Language	10
Low Socioeconomic	11

District Type

Type	N-Count
Metro	3
Suburban	2
Rural	6
Did Not Respond	0

Science Grades 5 and 8 and Biology

Years of Professional Experience in Education

Years	N-Count
1-5 years	1
6-10 years	2
11-15 years	2
16-20 years	1
More than 20 years	3

Position

Years	N-Count
Administrator	0
Special Education	6
General Education	3
Other	0

District Size

Type	N-Count
Large	4
Medium	2
Small	3
Did Not Respond	0

Ethnicity Distribution

Ethnicity	N-Count
African American	0
Asian or Pacific Islander	0
Hispanic	2
Multi-racial	0
Native American	0
White	9

Gender Distribution

Gender	N-Count
Female	7
Male	2

District Socioeconomic Status

Type	N-Count
High	1
Moderate	5
Low	3
Did Not Respond	0

Experience with Student Population

Student Population	N-Count
General Education	9
Special Education	8
English Language	7
Low Socioeconomic	8

District Type

Type	N-Count
Metro	1
Suburban	5
Rural	2
Did Not Respond	0

Social Studies Grade 8 and U.S. History

Years of Professional Experience in Education

Years	N-Count
1-5 years	0
6-10 years	3
11-15 years	1
16-20 years	4
More than 20 years	1

Position

Years	N-Count
Administrator	1
Special Education	6
General Education	3
Other	0

District Size

Type	N-Count
Large	5
Medium	1
Small	3
Did Not Respond	0

Ethnicity Distribution

Ethnicity	N-Count
African American	1
Asian or Pacific Islander	0
Hispanic	3
Multi-racial	0
Native American	0
White	9

Gender Distribution

Gender	N-Count
Female	8
Male	2

District Socioeconomic Status

Type	N-Count
High	2
Moderate	2
Low	5
Did Not Respond	0

Experience with Student Population

Student Population	N-Count
General Education	6
Special Education	9
English Language	7
Low Socioeconomic	9

District Type

Type	N-Count
Metro	1
Suburban	4
Rural	4
Did Not Respond	0

Appendix 2: STAAR Alternate 2 Standard-Setting Results by Round

Mathematics Grades 3-8 and Algebra I

Level II: Satisfactory Academic Performance							
	Grade/Subject						
	3	**4**	**5**	**6**	**7**	**8**	**Algebra I**
Number of Panelists	9	9	9	12	12	14	14
Round 1							
Mean	20.2	19.2	22.1	21.3	21.5	21.6	23.8
Median	16	19	22	22	22	22	24
Standard Deviation	7	2.11	1.27	3.02	2.15	4.09	4.73
Minimum	14	17	20	18	18	13	15
Maximum	31	22	24	26	25	28	33
Round 2							
Mean	21.4	21.1	22.6	20.2	21.7	21.8	24.7
Median	21	21	22	20	21	21	23
Standard	2.4	1.17	1.01	2.41	1.87	3.26	4.71
Minimum	19	20	21	15	20	16	19
Maximum	26	23	24	23	26	27	38
Round 3							
Mean	21.3	21.4	22.7	20.8	21.4	23.6	23.9
Median	21	21	22	21	22	24	23
Standard Deviation	1.73	1.67	1.41	1.64	1.24	4.96	4.38
Minimum	19	18	21	18	20	16	20
Maximum	24	24	25	23	23	35	37

Level III: Accomplished Academic Performance							
Round 1	**3**	**4**	**5**	**6**	**7**	**8**	**Algebra I**
Mean	33.2	31.8	34.7	29.6	33.3	30.4	33.5
Median	33	32	35	30	34	31	34
Standard Deviation	2.11	2.64	1.41	2.35	2.56	5.24	3.23
Minimum	31	28	32	26	29	21	26
Maximum	37	35	36	34	37	38	39
Round 2							
Mean	34.7	33.2	35	28.7	33	30.6	33.9
Median	35	34	35	29	34	31	34
Standard Deviation	0.71	1.86	0.71	2.5	2.8	3.69	2.54
Minimum	34	30	34	25	27	24	30
Maximum	36	35	36	32	37	36	38
Round 3							
Mean	34.9	35.2	35.9	32.8	33.8	34.2	34
Median	35	35	36	33	35	35	34
Standard Deviation	1.36	1.09	1.27	2.63	2.29	2.89	2.39
Minimum	33	33	34	30	27	29	30
Maximum	38	37	37	37	35	38	38

Reading Grades 3-8, English I, and English II

Level II: Satisfactory Academic Performance								
	Grade/Subject							
	3	**4**	**5**	**6**	**7**	**8**	**Eng I**	**Eng II**
Number of Panelists	9	9	9	7	7	7	11	11
Round 1								
Mean	24	25.8	24.3	23.6	26	29	23.6	22.5
Median	24	26	25	24	25	28	26	23
Standard Deviation	4.9	3.93	2.74	3.41	4.55	2.45	6.59	2.77
Minimum	15	21	19	19	21	26	15	18
Maximum	33	35	28	28	32	33	33	28
Round 2								
Mean	25	25.6	24.9	23.6	27.8	28.4	23.1	23.2
Median	24	26	25	24	28	28	24	23
Standard Deviation	2.78	2.19	2.03	2.23	1.94	1.81	3.62	3.54
Minimum	21	23	22	20	26	27	16	19
Maximum	29	30	28	27	31	32	28	32
Round 3								
Mean	25.1	25.6	25	25.6	27.3	27.4	22.8	22.5
Median	25	25	25	25	27	28	23	22
Standard Deviation	2.47	3.05	2.24	2.76	1.86	2.23	3.03	3.01
Minimum	22	22	22	22	25	25	19	20
Maximum	30	33	29	30	30	31	28	31

Level III: Accomplished Academic Performance								
Round 1	**3**	**4**	**5**	**6**	**7**	**8**	**Eng I**	**Eng II**
Mean	33	34.9	34.6	32.4	33	35.4	32.3	32.9
Median	33	34	35	33	33	35	32	33
Standard Deviation	2.78	2.42	2.3	3.87	2.94	0.98	4.29	2.02
Minimum	28	31	30	28	29	34	25	29
Maximum	37	39	38	37	37	37	38	36
Round 2								
Mean	33.6	36	35.3	32.7	35.8	34.1	32.7	33.7
Median	33	36	35	32	36	35	33	34
Standard Deviation	1.67	1.5	1.73	1.98	1.17	1.77	2.76	1.74
Minimum	31	34	33	30	35	32	26	32
Maximum	36	38	38	35	38	36	36	38
Round 3								
Mean	35.8	35.8	35.3	33.1	35.3	34.9	33.5	33.5
Median	36	36	35	34	36	35	34	34
Standard Deviation	1.2	2.05	1.66	1.46	0.82	1.21	2.11	1.44
Minimum	34	32	33	31	34	33	30	31
Maximum	37	38	38	35	36	36	36	36

Grades 5 and 8 Science and Biology

Level II: Satisfactory Academic Performance			
	Grade/Subject		
	5	**8**	**Biology**
Number of Panelists	9	9	9
Round 1			
Mean	20.2	23.9	21.7
Median	19	23	22
Standard Deviation	3.7	4.17	3.39
Minimum	16	17	16
Maximum	26	31	27
Round 2			
Mean	21.3	21.7	21.4
Median	21	21	22
Standard Deviation	2.5	3.54	1.94
Minimum	18	18	17
Maximum	25	29	24
Round 3			
Mean	22.1	20.6	21.2
Median	22	20	22
Standard Deviation	1.96	1.74	1.48
Minimum	19	18	18
Maximum	25	23	23

Level III: Accomplished Academic Performance			
	5	**8**	**Biology**
Round 1			
Mean	31.6	33.8	32.9
Median	32	33	33
Standard Deviation	2.01	2.05	2.98
Minimum	29	30	27
Maximum	35	37	38
Round 2			
Mean	32.3	34.6	34
Median	33	35	34
Standard Deviation	2.06	1.81	2.24
Minimum	28	32	31
Maximum	34	37	37
Round 3			
Mean	34	34.7	35.2
Median	34	35	36
Standard Deviation	1	1.87	1.64
Minimum	32	32	32
Maximum	35	37	37

Grade 8 Social Studies and U.S. History

Level II: Satisfactory Academic Performance		
	Grade/Subject	
	8	**US History**
Number of Panelists	10	10
Round 1		
Mean	25.3	23.9
Median	27	23
Standard Deviation	5.96	3.98
Minimum	16	19
Maximum	31	31
Round 2		
Mean	25.6	24.4
Median	24	25
Standard Deviation	4.58	2.19
Minimum	20	20
Maximum	35	27
Round 3		
Mean	24.3	23.9
Median	24	24
Standard Deviation	4	2.03
Minimum	18	20
Maximum	33	27

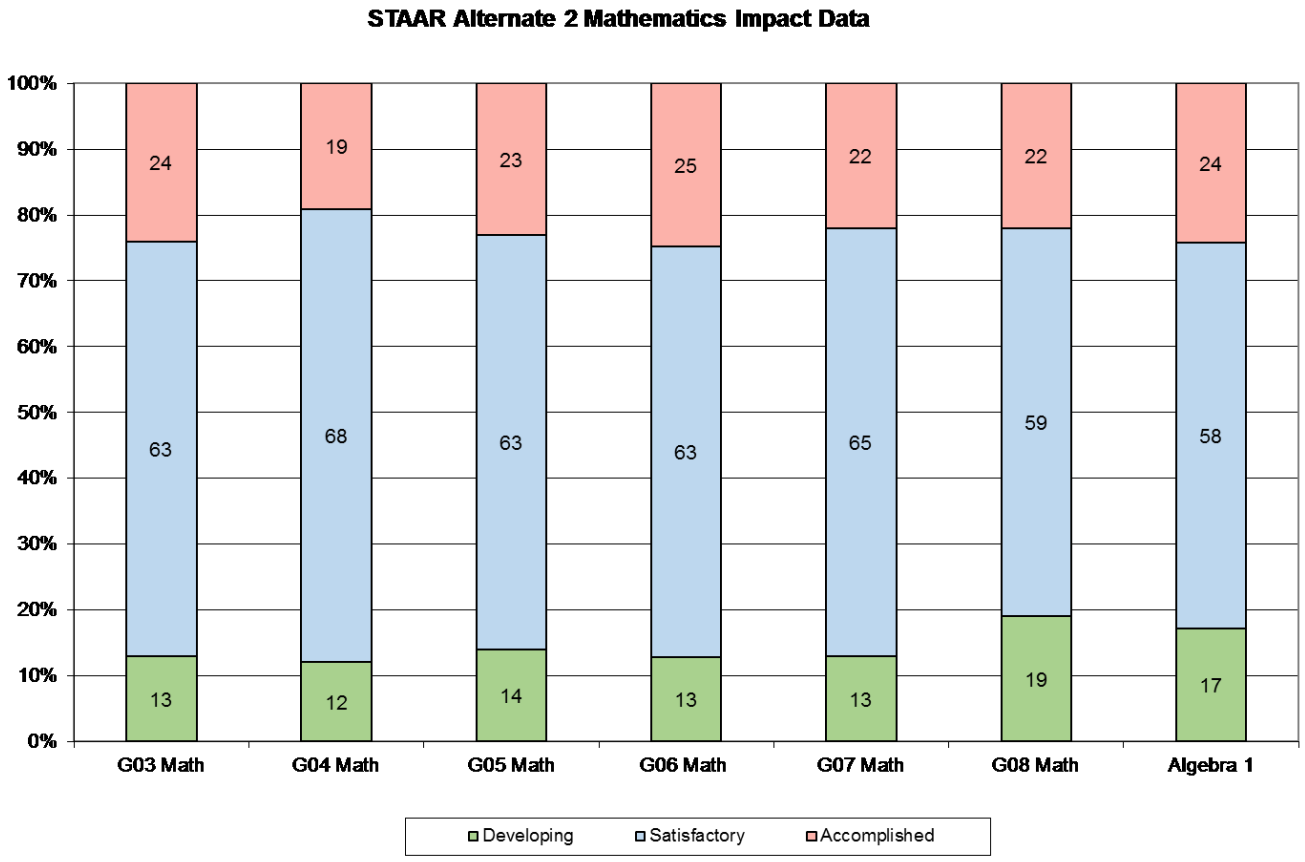
Level III: Accomplished Academic Performance		
Round 1	**8**	**US History**
Mean	33.2	33.2
Median	36	33
Standard Deviation	5.25	2.44
Minimum	21	29
Maximum	39	37
Round 2		
Mean	33.4	34.3
Median	34	34
Standard Deviation	1.51	1.94
Minimum	31	31
Maximum	36	38
Round 3		
Mean	33.2	36
Median	35	36
Standard Deviation	5.27	1.12
Minimum	20	34
Maximum	38	38

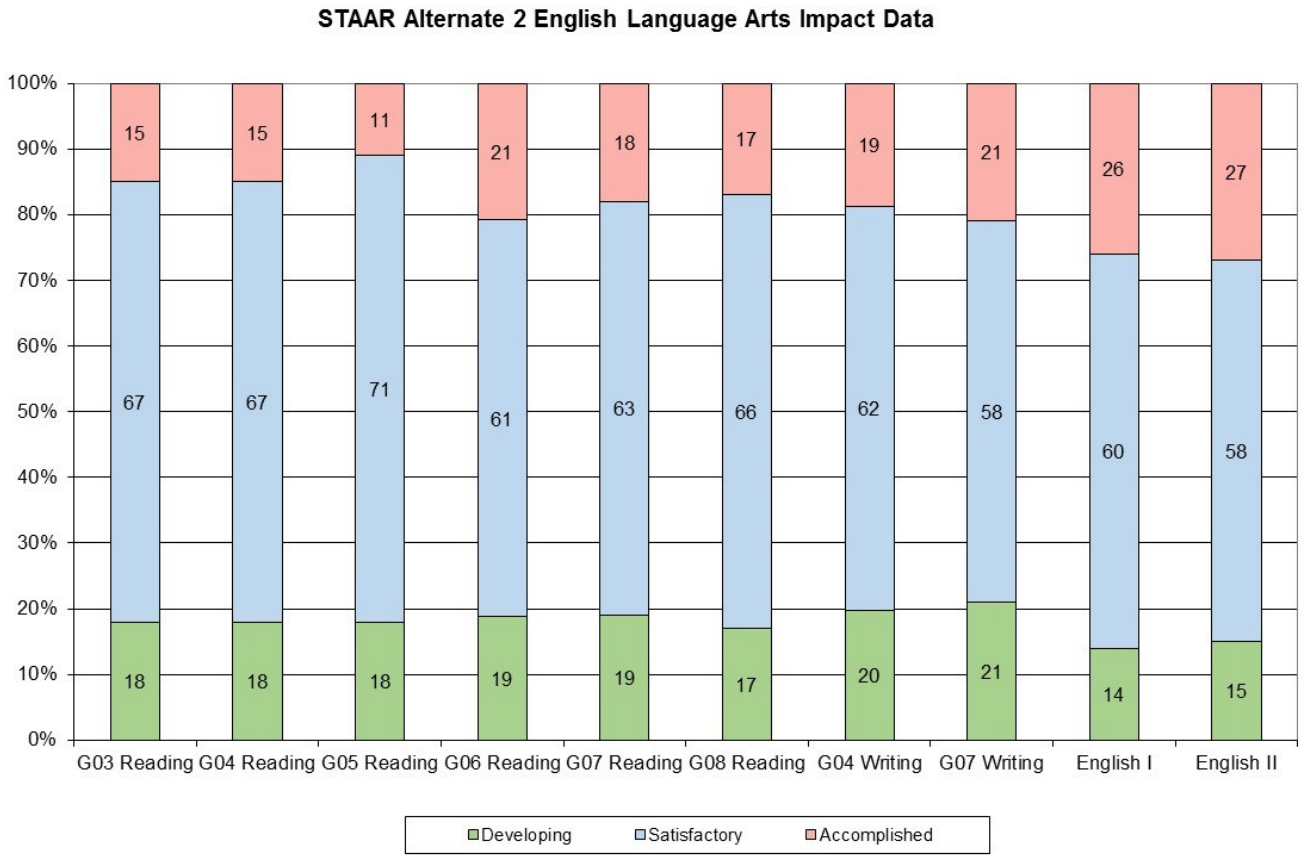
Grades 4 and 7 Writing

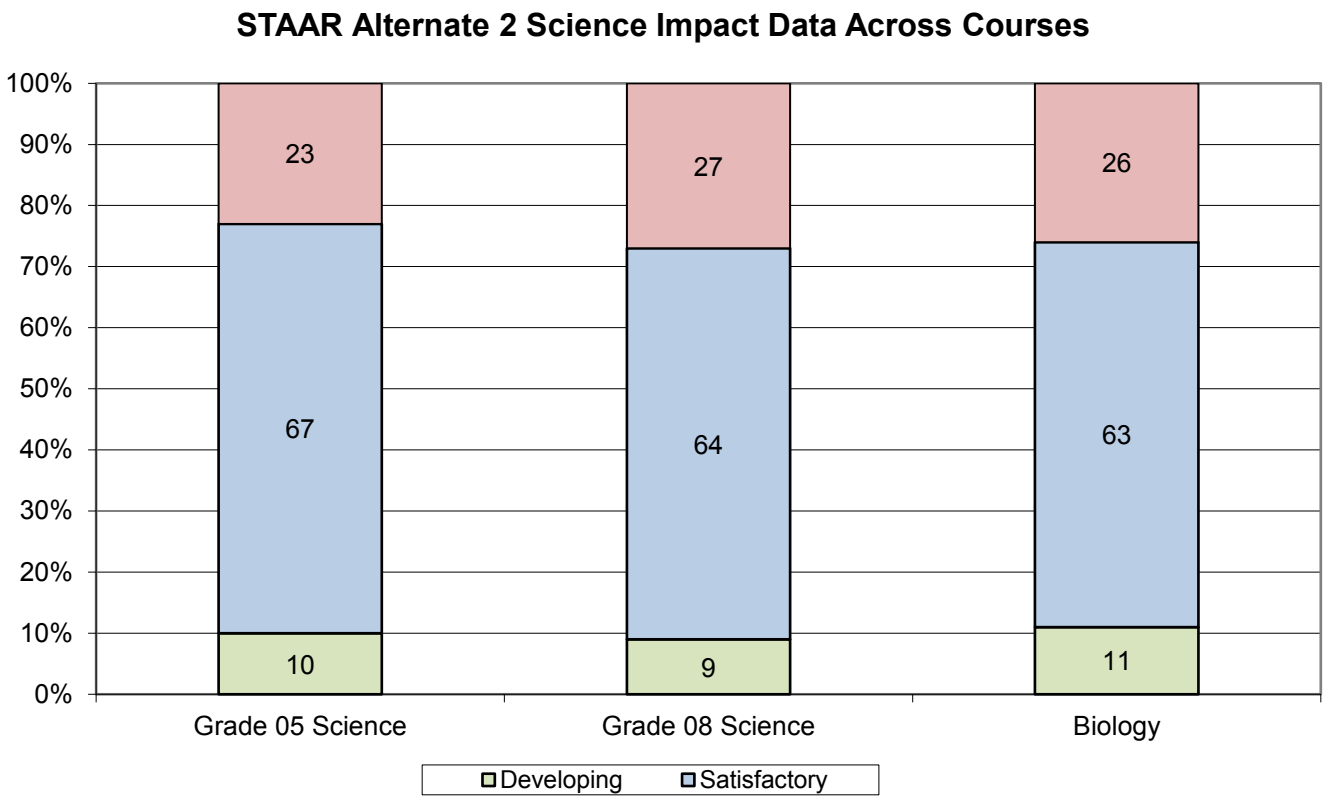
Level II: Satisfactory Academic Performance		
	Grade	
	4	**7**
Number of Panelists	14	14
Round 1		
Mean	24.7	23.8
Median	25	24
Standard Deviation	5.08	4.49
Minimum	18	17
Maximum	35	34
Round 2		
Mean	24.4	23.6
Median	24	23
Standard Deviation	3.8	3.18
Minimum	20	20
Maximum	31	31
Round 3		
Mean	23.7	33.6
Median	24	33
Standard Deviation	2.58	1.39
Minimum	20	32
Maximum	29	37

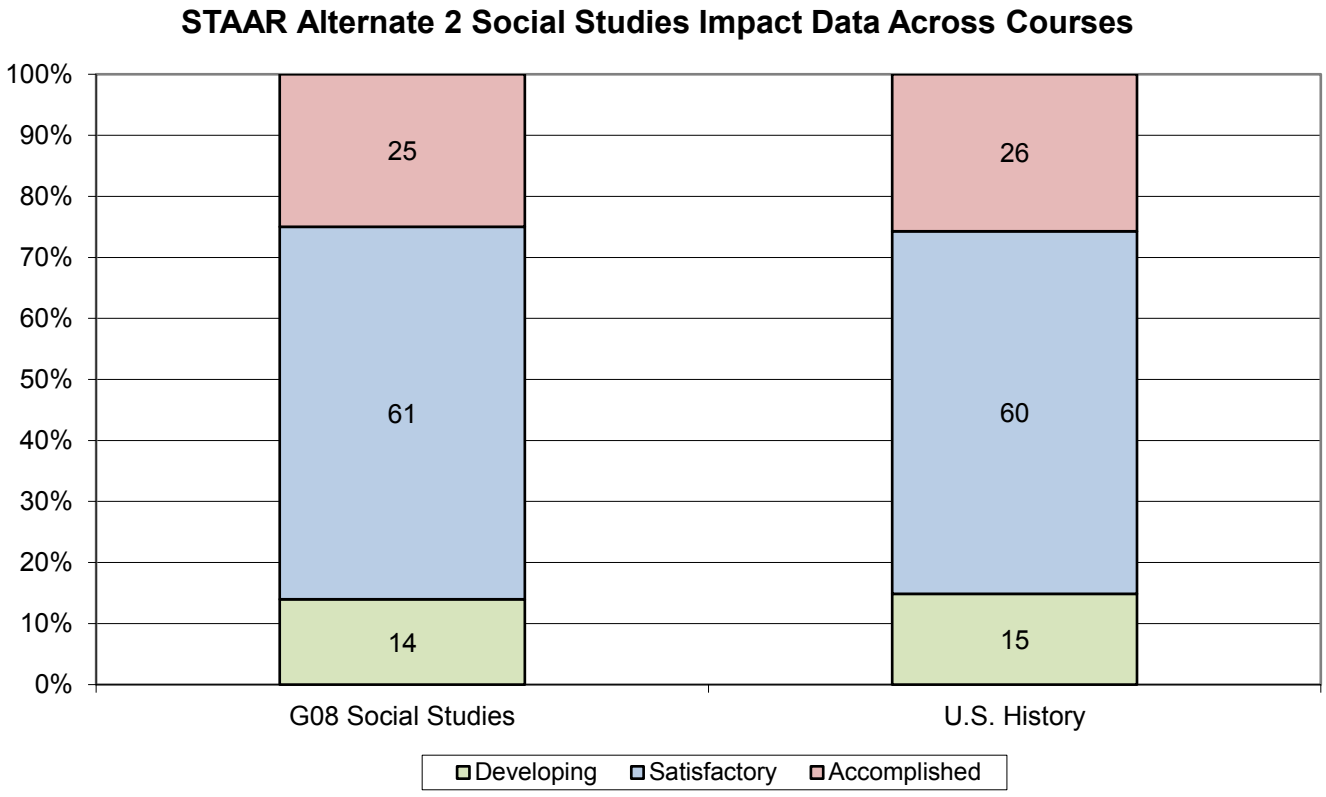
Level III: Accomplished Academic Performance		
Round 1	**4**	**7**
Mean	32.6	31.9
Median	33	32
Standard Deviation	3.48	2.57
Minimum	26	29
Maximum	38	37
Round 2		
Mean	33.3	32.3
Median	34	32
Standard Deviation	1.73	2.67
Minimum	30	29
Maximum	37	37
Round 3		
Mean	32.8	33.6
Median	33	33
Standard Deviation	1.67	1.39
Minimum	30	32
Maximum	35	37

Appendix 3: STAAR Alternate 2 Impact Data Using Final Cut Scores




STAAR Alternate 2 English Language Arts Impact Data
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
15
67
18
15
67
18
11
71
18
21
61
19
18
63
19
17
66
17
19
62
20
21
58
21
26
60
14
27
58
15
G03 Reading
G04 Reading
G05 Reading
G06 Reading
G07 Reading
G08 Reading
G04 Writing
G07 Writing
English I
English II
Developing
Satisfactory
Accomplished


STAAR Alternate 2 Science Impact Data Across Courses
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
23
67
10
27
64
9
26
63
11
Grade 05 Science
Grade 08 Science
Biology
Developing
Satisfactory


STAAR Alternate 2 Social Studies Impact Data Across Courses
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
25
61
14
26
60
15
G08 Social Studies
U.S. History
Developing
Satisfactory
Accomplished

Appendix 4: STAAR Alternate Process Evaluation Survey and Results

The purpose of this evaluation form is to collect information about your experience in recommending cut scores for the STAAR Alternate 2 assessment. Your opinions provide an important part of our evaluation of this meeting. **Please do not write your name on this evaluation form as we want your comments to be anonymous.** Thank you for your willingness to participate in this survey.

Current Position:

In which meeting did you participate? (Check only one.)

- ☐ Grades 3–5 Reading
- ☐ Grades 6–8 Reading
- ☐ Grades 4 & 7 Writing
- ☐ English I/English II
- ☐ Grades 3–5 Mathematics
- ☐ Grades 6 & 7 Mathematics
- ☐ Grade 8 Mathematics/Algebra I
- ☐ Grades 5 & 8 Science/Biology
- ☐ Grade 8 Social Studies/U.S. History

Years of Professional Experience in Education

- ☐ None
- ☐ 1–5 years
- ☐ 6–10 years
- ☐ 11–15 years
- ☐ 16–20 years
- ☐ More than 20 years

Do you have knowledge regarding policies and/or educational experience with any of the following populations? (Select all that apply.)

- ☐ Students receiving special education services
- ☐ Students of low socioeconomic status
- ☐ Students who are English language learners
- ☐ Students who are receiving general education instruction

If you currently work in a school district, please answer the questions below.		
Which word best describes the *size* of the district where you work? ☐ Small ☐ Medium ☐ Large	Which word best describes the *type* of district where you work? ☐ Rural ☐ Metropolitan ☐ Suburban	Which word best describes the *socioeconomic status* of the district where you work? ☐ Low ☐ Moderate ☐ High

1. Check the column below that best reflects your opinion about the level of success of the various components of the meeting in which you have just participated. The activities were designed to help you both understand the process and be supportive of the recommendations made by the committee.

a. Introduction to the process of setting performance standards				
b. Overview of STAAR Alternate 2				
d. Experiencing the assessment(s)				
f. Practice exercise for the standard-setting procedure				
h. Discussion after each round				

2. How useful do you feel the following activities or information were in assisting you to make your recommendations?

	Not Useful	Somewhat Useful	Useful	Very Useful
a. Performance Level Descriptors and Borderline Descriptors				
b. Training in the standard setting method				
c. Feedback data provided after Round 1				
d. Feedback data provided after Round 2				

3. How adequate were the following elements of the session?

	Not Adequate	Somewhat Adequate	Adequate	More Than Adequate
a. **Training** provided				
b. Amount of **time** spent training				
c. **Feedback** provided between rounds				
d. **Facilities** used for the session				
e. **Total amount of time** in breakout groups to make judgments				
f. **Number of rounds** for the judgments				

4. In applying the standard-setting method, you were asked to recommend cut scores (separating three proficiency levels) for student performance on STAAR Alternate 2 assessments. How confident do you feel that the Proficiency Level Descriptors (PLDs) are reasonable for each student performance level?

	Not Confident	Somewhat Confident	Confident	Very Confident
a. Developing				
b. Satisfactory				
c. Accomplished				

5. How confident do you feel that the final cut score recommendations represent appropriate levels of student performance?

	Not Confident	Somewhat Confident	Confident	Very Confident
a. Satisfactory				
b. Accomplished				

6. Did you have adequate opportunities during the session to:

	Not Adequate	Somewhat Adequate	Adequate	More Than Adequate
a. Express your opinions about student performance levels				
b. Ask questions about the standards and how they will be used				
c. Ask questions about the process of making cut score recommendations				
d. Interact with your fellow panelists				

7. Do you believe your opinions and judgments were treated with respect by:

	No	Sometimes	Yes
a. Fellow panelists			
b. Facilitators			

Please provide any additional comments you have regarding the standard-setting process, facilitator, etc.

Appendix 5: STAAR Alternate 2 Standard-Setting Process Evaluation Results

Mathematics Grades 3–8 and Algebra I

A total of 35 panelists responded to the process evaluation survey.

Section 1: Meeting Success

Meeting Component	Not Successful	Partially Successful	Successful	Very Successful	Omit
Introduction to the process of setting performance standards	0%	0%	29%	71%	0%
Overview of STAAR Alternate 2	0%	0%	23%	77%	0%
Discussion of performance level descriptors and borderline descriptors	0%	0%	20%	80%	0%
Experiencing the assessment(s)	0%	0%	20%	80%	0%
Overview of the standard-setting procedure	0%	0%	26%	74%	0%
Practice exercise for the standard-setting procedure	0%	0%	23%	77%	0%
Feedback data provided in each round	0%	0%	20%	80%	0%
Discussion after each round	0%	0%	20%	77%	3%

Section 2: Usefulness of Activities and Information

Activity or Information	Not Useful	Somewhat Useful	Useful	Very Useful	Omit
Performance Level Descriptors and Borderline Descriptors	0%	0%	23%	74%	3%
Training in the standard setting method	0%	0%	20%	80%	0%
Feedback data provided after Round 1	0%	0%	31%	69%	0%
Feedback data provided after Round 2	0%	3%	20%	77%	0%

Section 3: Adequacy of Meeting Elements

Meeting Element	Not Adequate	Somewhat Adequate	Adequate	More Than Adequate	Omit
Training provided	0%	0%	17%	83%	0%
Amount of time spent training	0%	0%	17%	83%	0%
Feedback provided between rounds	0%	0%	29%	71%	0%
Facilities used for the session	0%	0%	11%	89%	0%
Total amount of time in breakout groups to make judgments	0%	0%	20%	80%	0%
Number of rounds for the judgments	0%	0%	26%	71%	3%

Section 4: Specific PLDs

Performance Category	Not Confident	Somewhat Confident	Confident	Very Confident	Omit
Developing	0%	9%	40%	51%	0%
Satisfactory	0%	6%	46%	49%	0%
Accomplished	0%	6%	37%	57%	0%

Section 5: Cut Score Recommendations

Cut Score	Not Confident	Somewhat Confident	Confident	Very Confident	Omit
Satisfactory	0%	3%	40%	57%	0%
Accomplished	0%	3%	37%	60%	0%

Section 6: Opportunities to Collaborate

Category	Not Adequate	Somewhat Adequate	Adequate	More Than Adequate	Omit
Express your opinions about student performance levels	0%	0%	17%	83%	0%
Ask questions about the standards and how they will be used	0%	0%	17%	83%	0%
Ask questions about the process of making cut score recommendations	0%	0%	17%	83%	0%
Interact with your fellow panelists	0%	0%	17%	83%	0%

Section 7: Respect

Party	No	Sometimes	Yes	Omit
Fellow panelists	0%	3%	97%	0%
Facilitators	0%	0%	100%	0%

Reading Grades 3–8, Writing Grades 4 and 7, English I and English II

A total of 41 panelists responded to the process evaluation survey.

Section 1: Meeting Success

Meeting Component	Not Successful	Partially Successful	Successful	Very Successful	Omit
Introduction to the process of setting performance standards	0%	2%	41%	54%	2%
Overview of STAAR Alternate 2	0%	0%	34%	66%	0%
Discussion of performance level descriptors and borderline descriptors	0%	0%	37%	63%	0%
Experiencing the assessment(s)	0%	0%	24%	76%	0%
Overview of the standard-setting procedure	0%	0%	27%	73%	0%
Practice exercise for the standard-setting procedure	0%	0%	24%	76%	0%
Feedback data provided in each round	0%	0%	20%	80%	0%
Discussion after each round	0%	0%	27%	73%	0%

Section 2: Usefulness of Activities and Information

Activity or Information	Not Useful	Somewhat Useful	Useful	Very Useful	Omit
Performance Level Descriptors and Borderline Descriptors	0%	10%	27%	63%	0%
Training in the standard setting method	0%	0%	27%	73%	0%
Feedback data provided after Round 1	0%	0%	22%	78%	0%
Feedback data provided after Round 2	0%	0%	17%	83%	0%

Section 3: Adequacy of Meeting Elements

Meeting Element	Not Adequate	Somewhat Adequate	Adequate	More Than Adequate	Omit
Training provided	0%	0%	27%	71%	2%
Amount of time spent training	0%	0%	39%	61%	0%
Feedback provided between rounds	0%	2%	32%	63%	2%
Facilities used for the session	0%	0%	15%	85%	0%
Total amount of time in breakout groups to make judgments	0%	2%	24%	73%	0%
Number of rounds for the judgments	0%	0%	32%	68%	0%

Section 4: Specific PLDs

Performance Category	Not Confident	Somewhat Confident	Confident	Very Confident	Omit
Developing	2%	2%	56%	37%	2%
Satisfactory	2%	2%	54%	39%	2%
Accomplished	2%	7%	44%	44%	2%

Section 5: Cut Score Recommendations

Cut Score	Not Confident	Somewhat Confident	Confident	Very Confident	Omit
Satisfactory	0%	0%	56%	29%	15%
Accomplished	0%	5%	49%	32%	15%

Section 6: Opportunities to Collaborate

Category	Not Adequate	Somewhat Adequate	Adequate	More Than Adequate	Omit
Express your opinions about student performance levels	0%	0%	22%	78%	0%
Ask questions about the standards and how they will be used	0%	0%	24%	76%	0%
Ask questions about the process of making cut score recommendations	0%	0%	22%	78%	0%
Interact with your fellow panelists	0%	0%	12%	88%	0%

Section 7: Respect

Party	No	Sometimes	Yes	Omit
Fellow panelists	0%	2%	98%	0%
Facilitators	0%	0%	100%	0%

Science Grades 5 and 8, Biology and Social Studies Grade 8 and U.S. History

A total of 18 panelists responded to the process evaluation survey.

Section 1: Meeting Success

Meeting Component	Not Successful	Partially Successful	Successful	Very Successful	Omit
Introduction to the process of setting performance standards	0%	0%	44%	56%	0%
Overview of STAAR Alternate 2	0%	6%	28%	67%	0%
Discussion of performance level descriptors and borderline descriptors	0%	0%	39%	61%	0%
Experiencing the assessment(s)	0%	0%	33%	67%	0%
Overview of the standard-setting procedure	0%	6%	22%	72%	0%
Practice exercise for the standard-setting procedure	0%	0%	22%	78%	0%
Feedback data provided in each round	0%	0%	22%	78%	0%
Discussion after each round	0%	0%	39%	61%	0%

Section 2: Usefulness of Activities and Information

Activity or Information	Not Useful	Somewhat	Useful	Very Useful	Omit
Performance Level Descriptors and Borderline Descriptors	0%	11%	22%	67%	0%
Training in the standard setting method	0%	0%	22%	78%	0%
Feedback data provided after Round 1	0%	0%	11%	89%	0%
Feedback data provided after Round 2	0%	0%	6%	94%	0%

Section 3: Adequacy of Meeting Elements

Meeting Element	Not Adequate	Somewhat Adequate	Adequate	More Than Adequate	Omit
Training provided	0%	0%	33%	67%	0%
Amount of time spent training	0%	0%	33%	67%	0%
Feedback provided between rounds	0%	0%	22%	78%	0%
Facilities used for the session	0%	0%	17%	83%	0%
Total amount of time in breakout groups to make judgments	0%	0%	11%	89%	0%
Number of rounds for the judgments	0%	0%	33%	67%	0%

Section 4: Specific PLDs

Performance Category	Not Confident	Somewhat Confident	Confident	Very Confident	Omit
Developing	0%	0%	39%	61%	0%
Satisfactory	0%	0%	33%	67%	0%
Accomplished	0%	0%	50%	50%	0%

Section 5: Cut Score Recommendations

Cut Score	Not Confident	Somewhat Confident	Confident	Very Confident	Omit
Satisfactory	0%	0%	39%	56%	6%
Accomplished	0%	6%	44%	44%	6%

Section 6: Opportunities to Collaborate

Category	Not Adequate	Somewhat Adequate	Adequate	More Than Adequate	Omit
Express your opinions about student performance levels	0%	0%	22%	78%	0%
Ask questions about the standards and how they will be used	0%	0%	11%	89%	0%
Ask questions about the process of making cut score recommendations	0%	0%	17%	83%	0%
Interact with your fellow panelists	0%	0%	17%	83%	0%

Section 7: Respect

Party	No	Sometimes	Yes	Omit
Fellow panelists	0%	0%	100%	0%
Facilitators	0%	0%	100%	0%