

Chapter 5 STAAR Modified



Overview

Participation Requirements

Testing Requirements for Graduation

Test Development

Training

Test Administrations

Testing Accommodations

Student Success Initiative

Scores and Reports

Parent Brochures

Performance Standards

Scaling

Equating

Reliability

Validity

Measures of Student Progress

Sampling

Test Results

Overview

State of Texas Assessments of Academic Readiness Modified (STAAR Modified) is an alternate assessment based on modified academic achievement standards for students receiving special education services who meet participation requirements. STAAR Modified has been designed to meet federal requirements mandated under the No Child Left Behind (NCLB) Act. According to federal regulations, all students, including those receiving special education services, must be assessed on grade-level curriculum. STAAR Modified covers the same grade-level/course content as STAAR, but STAAR Modified assessments have been changed in format (larger font size, fewer items per page, etc.) and test design (shorter test blueprint, fewer answer choices, simpler vocabulary and sentence structure, etc.).



In 2013, STAAR Modified included assessments at grades 3–8 and nine end-of-course (EOC) assessments at the high school level. The initial administration of STAAR Modified occurred in spring 2012. However, not all of the STAAR Modified assessments were available at that time. House Bill (HB) 3 legislation called for the STAAR assessments to replace the Texas Assessment of Knowledge and Skills (TAKS) assessments for students in grades 9 and below starting with the 2011–2012 school year. As a result, the STAAR Modified assessments were available for students in grades 3–10 in 2013. Students in grade 11 continued to take the TAKS–M assessments if appropriate. Consequently, the STAAR Modified EOC assessments for courses typically taken in grade 11 will be administered for the first time in spring 2014. Refer to [chapter 8, “Texas Assessment of Knowledge and Skills \(TAKS\), TAKS \(Accommodated\), TAKS–Modified \(TAKS–M\) and Linguistically Accommodated Testing \(LAT\),”](#) for detailed information about these programs.

Due to HB 5 legislation, the English I and II assessments will combine the reading and writing assessments beginning in spring 2014, while other EOC assessments will no longer be administered. As a result, the implementation plan for STAAR Modified has been adjusted. The plan for implementing STAAR Modified is summarized below. Refer to the [Test Administrations](#) section for additional information about the various types of the STAAR Modified administrations.

Spring 2013

- STAAR Modified operational administrations with embedded field-test items: STAAR 3–8 assessments, Algebra I, geometry, biology, English I reading, English I writing, English II reading, English II writing, world geography
- STAAR Modified special operational administration: world history

Spring 2014

- STAAR Modified operational administrations with embedded field-test items: STAAR 3–8 assessments, Algebra I, biology, English I, English II
- STAAR Modified special operational administration: U.S. history

The STAAR Modified assessments administered in 2012–2013 are shown by grade and course in Table 5.1.

Table 5.1. 2012–2013 STAAR Modified Assessments

2012–2013 STAAR Modified Assessments	
Grade/Course	Assessed Content Area/Course
Grade 3	mathematics and reading
Grade 4	writing, mathematics, and reading
Grade 5	mathematics, reading, and science
Grade 6	mathematics and reading
Grade 7	writing, mathematics, and reading
Grade 8	mathematics, reading, science, and social studies
End-of-Course Assessments	Algebra I, geometry, biology, English I reading, English I writing, English II reading, English II writing, world geography, world history

Participation Requirements

Federal and state laws require the admission, review, and dismissal (ARD) committee for each student to make decisions about a student’s placement in the Texas assessment program. The first consideration is the general assessment, STAAR. If the general assessment is not appropriate, the ARD committee then considers the participation guidelines for STAAR Modified. The decision to administer STAAR Modified is neither based solely on disability category or placement setting, nor is it determined administratively; rather, the decision lies with the ARD committee based on the participation requirements.

Students receiving special education services are eligible to take STAAR Modified in one or more content areas if the answer to all three eligibility questions below is “Yes.”

1. Do the student’s present level of academic achievement and functional performance (PLAAFP) statements in the Individualized Education Program (IEP) lead the ARD committee to conclude that the student is multiple years behind grade-level or course expectations and will not progress at the same rate and level of rigor as his or her non-disabled peers?
2. Does the student’s IEP contain standards-based (TEKS-based) goals indicating that the student requires modified content in order to access the grade-level or course curriculum?
3. Does the student require direct and intensive instruction in order to acquire, maintain, and transfer skills to other contexts?

Refer to the [STAAR Modified Participation Requirements](#) page on TEA’s Student Assessment Division website for more information. Any student who meets participation requirements for STAAR Modified may take the STAAR Modified assessments. However, only two percent of the tested population can count as proficient for the purpose of Adequate Yearly Progress (AYP) performance calculations.



Testing Requirements for Graduation



Students taking the STAAR Modified EOC assessments are in the Minimum High School Program (MHSP) because they are receiving modified instruction. For those students eligible to take STAAR Modified EOC, passing the assessments is not a requirement for graduation. A student's ARD committee determines individual requirements for graduation.

Test Development

The test development process for STAAR Modified follows, as closely as possible, the procedures used for development of other assessment programs in Texas, coupled with the additional requirements specific to STAAR Modified.

The STAAR Modified blueprints mirror the STAAR blueprints and reflect that the students taking STAAR Modified are assessed on the same grade-level curriculum as general education students. The number of items on the STAAR Modified blueprints is reduced from the number of items on the STAAR blueprints by 20 percent. The STAAR Modified blueprints contain similar percentages of readiness and supporting items as found in the STAAR blueprints.

Issues of validity, reliability, fairness, accessibility, and consistency in meaning are carefully considered as part of the item modification and review processes. As STAAR Modified items are developed and reviewed, attention is given to the standards of fairness and the principles of alignment and universal design. According to the principles of universal design, every item must have precisely defined constructs and demonstrate maximum legibility, maximum readability, and maximum comprehensibility. Similarly, each item must be adaptable for purposes of accommodations, be accessible and non-biased, and take special populations into consideration.

Using results from a literature review of modifications that would be appropriate for students with disabilities, TEA modified existing STAAR items. Modification guidelines were developed for each subject/content area to ensure that the modifications did not affect the construct of the items, and the item modifications would be consistent across development years.

In addition to strictly adhering to these modification guidelines, TEA convenes internal and external educator item review meetings to further ensure that the modified items meet the criteria listed above. Texas educators—general and special education classroom teachers, curriculum specialists, administrators, and regional education service center (ESC) staff—play a vital role in the STAAR Modified development process. The participation of these education professionals enables TEA to develop high-quality alternate assessment instruments that accurately reflect the Texas Essential Knowledge and Skills (TEKS) curriculum.



Training

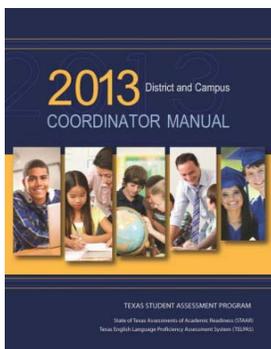
The test administration training for STAAR Modified is the same as the training available for STAAR, and can be found in the Training section of [chapter 4, “State of Texas Assessments of Academic Readiness \(STAAR\).”](#) For each test administration in the 2012–2013 school year, ESC personnel and district coordinators were given a district testing coordinator packet that contained all the information and materials necessary for overseeing test administrations and copies of the coordinator and test administrator manuals.

Training opportunities were conducted via the Texas Education Telecommunication Network (TETN), 2013 ESC Training, and the 2012 Texas Assessment Conference.

The presentations addressed these topics:

- accommodation policies and procedures
- required documentation for assessment and accommodation decisions
- modification guidelines and processes
- information regarding making assessment decisions in ARD committee meetings

District and Campus Coordinator Manual



The [2013 District and Campus Coordinator Manual \(DCCM\)](#) explains the responsibilities of district and campus testing coordinators for the STAAR program. This manual contains preparation and administration procedures for every program for the 2013 calendar year. Separate test administrator manuals are available for districts prior to the first assessment administration for each grade.

Test Administrations

In spring 2013, the STAAR Modified assessments had two types of administrations: operational administrations and a special operational administration.

An operational administration is one in which all items included on the test have been previously field-tested and evaluated before the test administration to determine whether the items should appear on an operational assessment. Students receive a score based on their performance.

A special operational administration is one in which all items included on the test have not been previously field-tested. However, students do receive a test score based on



their performance on the assessment. To give students their scores on a special operational administration, items are analyzed immediately after the administration, and item statistics are evaluated to determine which items are appropriate to contribute to a student’s score. These items are also deemed suitable for use on future operational assessments. The STAAR Modified world history assessment was the only special operational administration in 2013.

More than 349,000 STAAR Modified assessments were administered in 2012–2013. The number of STAAR Modified assessments administered is shown in Table 5.2.

Table 5.2. STAAR Modified Assessments Administered in 2012–2013

STAAR Modified Assessments Administered 2012–2013	
STAAR Modified Assessment	Assessments Administered
Grade 3 reading	9,933
Grade 3 mathematics	11,544
Grade 4 reading	14,162
Grade 4 mathematics	12,796
Grade 4 writing	14,070
Grade 5 reading	20,282
Grade 5 mathematics	21,466
Grade 5 science	12,611
Grade 6 reading	15,000
Grade 6 mathematics	14,622
Grade 7 reading	13,330
Grade 7 mathematics	13,832
Grade 7 writing	13,461
Grade 8 reading	17,504
Grade 8 mathematics	19,440
Grade 8 social studies	11,214
Grade 8 science	11,559
Algebra I	13,326
Geometry	10,284
Biology	11,447
English I reading	12,633
English I writing	12,571
English II reading	10,940
English II writing	10,958
World geography	11,020
World history	9,423



Educational Materials Required for Testing

DICTIONARIES AND THESAURUSES

English-language dictionaries and thesauruses must be provided to students for the writing assessment at grade 7, the reading assessments at grades 6–8, and the English I and II reading and writing assessments.

There must be at least one dictionary for every five students; it is also recommended that there be one thesaurus for every five students, if possible. Students may also use a combination dictionary/thesaurus. An English as a second language (ESL) dictionary that uses simple English and pictures to define words may be provided for ELLs. Both paper and electronic dictionaries are permitted, though electronic dictionaries must not allow access to the Internet.

Specific information regarding dictionary policies for each STAAR assessment, including STAAR Modified, can be found on the [STAAR Resources](#) page on TEA's Student Assessment Division website.

CALCULATORS

Calculators must be provided to students for the STAAR Modified Algebra I, geometry, and biology assessments. Students may use their own calculators instead of those provided by the district. Districts may provide students with more than one calculator during the assessment but must provide, at a minimum, the following:

- a graphing calculator for each student taking Algebra I and geometry
- a calculator for every five students taking biology

Any calculator may be used to fulfill these minimum requirements except for those that include a computer algebra system (CAS) or that allow access to the Internet. In addition, an electronic device that has a calculator as an application cannot be used (e.g., a cell phone or smartphone).

All calculator memory must be cleared to factory default both before and after testing. Any programs or applications must be removed or disabled prior to testing.

Calculators may not be provided to a student taking the grades 3–8 mathematics assessments or the grades 5 and 8 science assessments unless the student meets the criteria for such an accommodation.

Specific information regarding calculator policies for each STAAR assessment can be found on the [STAAR Resources](#) page on TEA's Student Assessment Division website.

Testing Accommodations

Accommodations are practices and procedures that provide equitable access to grade-level or course curriculum during instruction and assessment. The decision to use a particular accommodation with a student is made on an individual basis and takes into



consideration both the needs of the student and whether the student routinely receives the accommodation in classroom instruction and testing. Further information on testing accommodations can be found in [chapter 4, “State of Texas Assessments of Academic Readiness \(STAAR\).”](#) Specific information about each accommodation can be found on the [Accommodations Resources](#) page on TEA’s Student Assessment Division website.

Accommodations for Students with Disabilities

For a student who receives special education or Section 504 services, the decision to allow the student to use accommodations during the statewide assessments is made by the student’s ARD committee or Section 504 placement committee. In those rare instances where a student does not receive special education services but does meet the eligibility criteria due to a disabling condition, the decision to allow accommodations on the statewide assessments is made by the appropriate team of people at the campus level, such as the Response to Intervention (RTI) team or student assistance team. Further information on the types of accommodations available for students with disabilities can be found in [chapter 4, “State of Texas Assessments of Academic Readiness \(STAAR\).”](#)

After determining the instructional accommodation(s) that are effective for a student with disabilities, the educator should investigate whether each accommodation is allowed on a statewide assessment.

Dyslexia Accommodations

Accommodations are available for students who have dyslexia and other reading disabilities on the STAAR Modified reading assessments in grades 3 through high school. These accommodations include:

- having all items and answer choices read aloud to a student, and
- extending the four-hour testing time over an entire school day.

The needs of the student should be carefully considered when determining the most appropriate accommodations for the STAAR Modified assessments.

Oral Administration

Oral administration is an accommodation that allows for test questions and answer choices for reading, mathematics, science, and/or social studies tests to be read aloud or signed to eligible students taking the STAAR assessments. Two levels of reading support are available: read questions and answer choices at the student’s request, and read all questions and answer choices. Required reference materials (e.g., a dictionary) and allowable supplemental aids (e.g., a list of grammar rules or a math chart) may be read to students, but the test administrator cannot interpret or help apply the information contained within. STAAR may be administered orally to individual students or to a group of students, depending on student needs. All references to

reading support during an oral administration also apply to signing during a signed administration.

A student is eligible for an oral administration if he or she routinely and effectively uses this accommodation during classroom instruction and testing and meets at least one of the following criteria:

- the student receives special education services and is identified with dyslexia or has evidence of reading difficulties,
- the student receives Section 504 services and is identified with dyslexia or has evidence of reading difficulties, or
- the student does not receive special education or Section 504 services but is identified with dyslexia.

The ARD committee or Section 504 committee decides whether a student exhibits evidence of a reading difficulty, which is defined as a problem with reading that may be caused by a learning disability or other conditions such as ADHD, an emotional or behavioral disability, or processing or memory issues.

Students with Visual Impairments

Test administrators receive specific instructions for testing visually impaired students using braille test booklets. Districts are required to indicate on the answer document whether the student used a large-print or braille version of a test. Large-print and braille test booklets are available for all STAAR Modified operational administrations.

Linguistic Accommodations

Linguistic accommodations are language supports that make grade-level academic assessments in English more accessible to ELLs. The Texas English Language Proficiency Standards (ELPS) require all teachers to linguistically accommodate the instruction of ELLs in their classes in a manner that is commensurate with the students' English language proficiency levels. The policies for the STAAR linguistic accommodations support these ELPS requirements.

More information about the STAAR linguistic accommodations can be found in [chapter 4, “State of Texas Assessments of Academic Readiness \(STAAR\)”](#) and is available in the document [Linguistic Accommodations for ELLs Participating in the STAAR Program](#) on the [Accommodations Resources](#) page on TEA's Student Assessment Division website.

Additional information regarding accommodations, including eligibility requirements, is available on the [STAAR Resources](#) page on TEA's Student Assessment Division website.



Student Success Initiative



The Student Success Initiative (SSI) provides a system of academic support to help students achieve on grade level in reading and mathematics. The SSI incorporates a grade-advancement component adopted by the Texas Legislature in 1999 to ensure that all students receive the instruction and support they need to be academically successful in reading and mathematics.

Under the SSI grade advancement requirements, students are required to pass the STAAR Modified grade 5 reading and mathematics assessments to be promoted to sixth grade. Additionally, students are required to pass the STAAR Modified grade 8 reading and mathematics assessments to be promoted to ninth grade.

In 2012–2013, students in grades 5 and 8 had three opportunities (April, May, and June) to pass the STAAR Modified reading and mathematics assessments. For students who take the STAAR Modified assessments and do not achieve at least Level II: Satisfactory Academic Performance, the ARD committee determines if students must retest. The ARD committee also makes decisions regarding promotion, accelerated instruction, and retention.

More information about the [Student Success Initiative](#) is available on TEA's Student Assessment Division website.

Scores and Reports

There are a variety of reports that show a student's performance on the assessments in the STAAR Modified program. Refer to the information below for details about the types of scores given on reports and the types of reports available.

Description of Scores

Scores for the STAAR Modified assessments consist of the number of items answered correctly (raw scores), scale scores, and reporting category performance information.

RAW SCORE

The number of items that a student answers correctly on a STAAR Modified assessment is the student's raw score. The raw score can be interpreted only in terms of a specific set of test items on that test form. However, because the difficulty of items might vary among test forms over time, differences in student performance across tests or administrations cannot be compared using raw scores alone. To compare student scores across different test forms and different administrations, 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: Advanced Academic Performance. (Performance-level cut scores are discussed in the [Performance Standards](#) section of this chapter.) Scale scores were available for all STAAR Modified assessments administered in 2013, except for world history. Performance standards for world history were to be set in summer 2013, but due to HB 5 legislation, the assessment was discontinued after the spring 2013 administration. As a result, only raw scores for world history were available to districts after the spring 2013 administration.

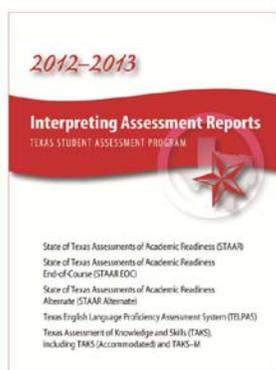
Scale scores are also used to compare the performance of an individual student with the performance of a demographic group, a program group, an entire campus, or a district at a particular grade. For example, the scores for a Hispanic student in a gifted and talented program can be compared with the average scores of other Hispanic students, other gifted and talented students, all the students on a campus, or any combination of these aggregations at that grade.

ADDITIONAL PERFORMANCE INFORMATION

Other scores can provide information about a student's relative strengths or weaknesses in core academic areas. For example, reporting category-level data can identify areas where a student might be having difficulty. This identification can help campuses plan the most effective instructional intervention. Finally, individual student test scores are also used in conjunction with other performance indicators to assist in making placement decisions. While scores can contribute to decisions regarding placement, educational planning for a student should take into account as much student information as possible.

Report Formats

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



For more information about scoring and reporting for STAAR Modified, refer to the TEA publication [Interpreting Assessment Reports](#).

Use of Test Results



Test results can be used to evaluate the performance of a group over time. Average scale scores and the percentage of students meeting the Level II and Level III performance standards can be analyzed by grade and content area across administrations to give insight into whether student performance is improving across years. For example, the average scale score for students who receive special education services and who took the STAAR Modified grade 4 writing test can be compared over time.

Test results can also be used to compare the performance of different demographic or program groups. The STAAR Modified scores can be analyzed within the same content area of any single administration to determine which demographic or program group had the highest average scale score, or which group had the lowest percentage meeting the Level II performance standard, or which group had the highest percentage achieving Level III performance standard, etc. Other scores can be used to help evaluate the academic performances of demographic or program groups in core academic areas. For example, reporting category data can help campuses and districts identify areas of potential academic weakness for a group of students. The same methodology can be applied to an entire campus or district. Test results for groups of students can be used when evaluating instruction or programs that require average-score or year-to-year comparisons. Because the tests are designed to measure content areas within the required state curriculum, the consideration of test results by content area and by reporting category might be helpful when evaluating curriculum and instruction. In addition, all test scores can be compared with regional and statewide performance within the same content area for any administration.

Generalizations from test results can be made to the specific content area being measured on the test. However, because each test measures a finite set of skills with a limited set of items, any generalizations about student achievement derived solely from a particular assessment should be made with great care and with full reference to the fact that the conclusions were based only on that assessment. Instruction and program evaluations should take into account as much information as possible to provide a more complete picture of performance.

Parent Brochures

TEA's Student Assessment Division produces a brochure titled "[Understanding Your Child's Confidential Student Report \(CSR\): A Guide for Parents](#)" to help parents understand their child's STAAR 3–8 test results. This brochure provides a brief summary of the STAAR program, including STAAR Modified, and explains information contained on a CSR so that parents can understand their child's test report. The brochure, available in both English and Spanish, was provided to districts for distribution with individual student STAAR Modified performance results. For STAAR Modified EOC, an explanation of the test results is printed on the CSR for each individual assessment.



Performance Standards

Performance standards relate levels of test performance directly to what students are expected to learn, as defined in the statewide curriculum. This is done by establishing cut scores that distinguish between performance levels or categories. Standard setting is the process of establishing these cut scores that define the performance levels for an assessment.

Performance Levels and Policy Definitions

For the STAAR Modified assessments, the performance levels are:

- Level I: Unsatisfactory Academic Performance
- Level II: Satisfactory Academic Performance
- Level III: Advanced Academic Performance

More detailed descriptions, known as policy definitions, of each performance level are as follows:

LEVEL I: UNSATISFACTORY ACADEMIC PERFORMANCE

Performance in this category indicates that students are inadequately prepared for the next grade or course even with instructional modifications such as simplified language and concepts. They demonstrate an insufficient understanding of the assessed knowledge and skills. Students in this category are unlikely to succeed in the next grade or course without significant and/or additional modifications and increased support.

LEVEL II: SATISFACTORY ACADEMIC PERFORMANCE

Performance in this category indicates that students are sufficiently prepared for the next grade or course with instructional modifications such as simplified language and concepts. They generally demonstrate the ability to understand and apply the assessed knowledge and skills in familiar contexts. Students in this category have a reasonable likelihood of success in the next grade or course with continued modifications and support.

LEVEL III: ADVANCED ACADEMIC PERFORMANCE*

Performance in this category indicates that students are well prepared for the next grade or course with instructional modifications such as simplified language and concepts. They demonstrate the ability to understand and apply the assessed knowledge and skills. Students in this category have a high likelihood of success in the next grade or course with continued modifications and support.

* For English III, this level of performance also indicates that students are prepared for postsecondary endeavors with support.

Standard-Setting Process for STAAR Modified

Standard setting for STAAR Modified took into consideration a variety of factors such as policy, 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.

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

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

Tables 5.3 and 5.4 provide high-level descriptions and timelines for these nine steps as implemented in the STAAR Modified EOC and STAAR Modified 3–8 standard-setting processes, respectively. Although HB 5 legislation removed the requirement for the review of performance standards at least once every three years, step nine is included in the process because this step was considered when setting performance standards for STAAR Modified.

Additional detail about each step in the STAAR standard-setting process is given in the [“STAAR Modified Standard-Setting Technical Report,”](#) available on the STAAR Resources page on TEA’s Student Assessment Division website.

Table 5.3. Overview of the STAAR Modified EOC Standard-Setting Process

Overview of the STAAR Modified EOC Standard-Setting Process		
Standard-Setting Step	Description	Timeline
1. Conduct empirical studies	Scores on each assessment are linked to performance on other assessments in the same content area (when available).	Studies started in spring 2012 and will continue throughout the program.
2. Develop performance labels and policy definitions	A committee was convened jointly by TEA and the THECB to recommend performance categories, performance category labels, and general policy definitions for each performance category for the STAAR program. This information was adapted to apply to the STAAR Modified program.	September 2010
3. Convene a policy committee	Committee considers policy implications of performance standards and empirical study results and makes recommendations to identify reasonable ranges (“neighborhoods”) for the STAAR EOC cut scores. This information is used to inform the reasonable ranges for the STAAR Modified cut scores.	February 1–2, 2012
4. Develop grade-/course-specific performance level descriptors (PLDs)	Committees consisting primarily of educators started with the general STAAR PLDs to develop STAAR Modified PLDs as an aligned system, describing a reasonable progression of skills within each content area (English, mathematics, science, and social studies).	June 2012
5. Convene standard-setting committees	Committees consisting of K–12 educators used the performance labels, policy definitions, PLDs, and reasonable ranges to recommend cut scores for each STAAR Modified EOC assessment.	August 8–14, 2012
6. Review performance standards for reasonableness	TEA reviews the recommendations across content areas.	September 2012
7. Approve performance standards	The commissioner of education approves the performance standards.	December 2012
8. Implement performance standards	Performance standards are reported to students for the spring 2012 administration with phase-in standards applied.	January 2013
9. Review performance standards	Performance standards are reviewed at least once every three years.*	If applicable

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



Table 5.4. Overview of the STAAR Modified 3–8 Standard Setting Process

Overview of the STAAR Modified 3–8 Standard Setting Process		
Standard-Setting Step	Description	Timeline
1. Conduct empirical studies	Scores on each assessment are linked to performance on other assessments in the same content area (when available).	Studies started in spring 2012 and will continue throughout the program.
2. Develop performance labels and policy definitions	Committee is convened jointly by TEA and the THECB to recommend performance categories, performance category labels, and general policy definitions for each performance category for the STAAR program. This information was adapted to apply to the STAAR Modified program.	September 2010
3. Develop performance standard ranges	STAAR Modified EOC performance standards and empirical study results are used to identify reasonable ranges (“neighborhoods”) for the cut scores for Levels II and III.	August 2012
4. Develop grade/subject specific performance level descriptors (PLDs)	Committees consisting primarily of educators started with the general STAAR PLDs to develop STAAR Modified PLDs as an aligned system, describing a reasonable progression of skills within each content area (English, mathematics, science, and social studies).	September 2012
5. Convene standard-setting committees	Committees consisting of K–12 educators used the performance labels, policy definitions, PLDs, and reasonable ranges to recommend cut scores for each STAAR Modified assessment.	October 29– November 9, 2012
6. Review performance standards for reasonableness	TEA reviews the cut-score recommendations across grades and subjects.	November 2012
7. Approve performance standards	The commissioner of education approves performance standards.	December 2012
8. Implement performance standards	Performance standards are reported to students for the spring 2012 administration with phase-in standards applied.	January 2013
9. Review performance standards	Performance standards are reviewed at least once every three years.*	If applicable

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





Standard-Setting Committees

The task of each standard-setting committee was to recommend the two cut scores that would define the three performance levels for each of the STAAR Modified assessments. The standard-setting committees were made up of K–12 educators. Each committee included general education teachers who were experts in both the assessed content and the curriculum, and special education teachers having expertise with students who take the STAAR Modified assessments.

Committee members were provided with reasonable ranges within which performance standards should be set. The ranges were determined after a careful consideration of the alignment of performance standards with the STAAR Modified EOC and 3–8 assessments in the same content area, the relevant information from the policy committee, and the results of various empirical studies. (The studies provided research-based anchors for setting performance standards that were meaningful and rigorous.)

In August 2012, TEA convened standard-setting committees that recommended performance standards for the STAAR Modified Algebra I, geometry, biology, English I reading, English II reading, English I writing, English II writing, and world geography assessments.

In October and November 2012, TEA convened standard-setting committees that recommended performance standards for STAAR Modified grades 3–8 mathematics, grades 3–8 reading, grades 5 and 8 science, grade 8 social studies, and grades 4 and 7 writing.

STAAR Modified world history was administered for the first time in May 2013. Performance standards for world history were to be set in summer 2013; but due to HB 5 legislation, the assessment was discontinued after the spring 2013 administration and performance standards were not set.

Phase-in of Performance Standards

A phase-in period has been implemented for the STAAR Modified performance standards in order to provide school districts with sufficient time to adjust instruction, provide new professional development, increase teacher effectiveness, and close knowledge gaps. The commissioner of education determines the appropriate timeline for phasing in the performance standards.

A two-step phase-in for Level II: Satisfactory Academic Performance is in place for all STAAR 3–8 and EOC assessments. Phase-in 1 performance standards for Level II will be in effect for the 2011–2012, 2012–2013, and 2013–2014 school years. Phase-in 2 performance standards are currently planned to take effect in the 2014–2015 school year unless the commissioner of education determines a different phase-in standard is warranted. The final recommended Level II standards will be in place following phase-in 2. There is no phase-in for the Level III performance standard for any of the STAAR Modified assessments.



The STAAR Modified EOC phase-in periods for performance standards are on a student-by-student basis by content area (mathematics, English, science, and social studies). The phase-in standard to which students are held depends on when the students begin testing in a content area, and it applies to all assessments in that content area. For example, for students who first took Algebra I in spring 2012, the first phase-in standard for Level II: Satisfactory Performance applies to both mathematics assessments (Algebra I and geometry).

Outcome of Standard Setting

The standard-setting process elicited recommended cut scores that reflect the level of performance a student must achieve for each performance category of the STAAR Modified assessments. Performance standards for the STAAR Modified grades 3–8 assessments and Algebra I, geometry, biology, English I reading, English I writing, English II reading, English II writing, and world geography were recommended by standard-setting committees and approved by the commissioner of education in December 2012. Tables 5.5 and 5.6 show the approved performance standards in scale score units for the STAAR Modified 3–8 and EOC assessments, respectively.

Table 5.5. STAAR Modified 3–8 Performance Standards

STAAR Modified 3–8 Performance Standards				
Assessment	Phase-in 1 Level II	Phase-in 2 Level II	Final Recommended Level II	Final Recommended Level III
Grade 3 mathematics	2800	2900	3000	3578
Grade 4 mathematics	2800	2900	3000	3526
Grade 5 mathematics	2800	2900	3000	3691
Grade 6 mathematics	2800	2900	3000	3462
Grade 7 mathematics	2800	2900	3000	3551
Grade 8 mathematics	2800	2900	3000	3577
Grade 3 reading	2800	2900	3000	3306
Grade 4 reading	2800	2900	3000	3238
Grade 5 reading	2800	2900	3000	3312
Grade 6 reading	2800	2900	3000	3316
Grade 7 reading	2800	2900	3000	3368
Grade 8 reading	2800	2900	3000	3436
Grade 4 writing	2800	2900	3000	3349
Grade 7 writing	2800	2900	3000	3422
Grade 5 science	2800	2900	3000	3234
Grade 8 science	2800	2900	3000	3509
Grade 8 social studies	2800	2900	3000	3348

Table 5.6. STAAR Modified EOC Performance Standards

STAAR Modified EOC Performance Standards				
Assessment	Phase-in 1 Level II	Phase-in 2 Level II	Final Recommended Level II	Final Recommended Level III
Algebra I	2800	2900	3000	3470
Geometry	2800	2900	3000	3743
Biology	2800	2900	3000	3500
English I reading	1430	1450	1500	1672
English II reading	1430	1450	1500	1652
English I writing	1430	1450	1500	1612
English II writing	1430	1450	1500	1604
World geography	2800	2900	3000	3354



Review of Performance Standards

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

Scaling

Scaling is a statistical procedure that places raw scores on a common scoring metric in order to make test scores easier to interpret and compare across test administrations. As with previous Texas assessment programs, the STAAR program uses the Rasch Partial-Credit Model (RPCM) to place test items on the same scale across administrations for a given STAAR Modified assessment. Once performance standards have been set for an assessment, its Rasch scale is then transformed to the more user-friendly metric of a reporting scale in order to further facilitate interpretation of the test scores. Details of the RPCM scaling method used in Texas are provided in [chapter 3, “Standard Technical Processes.”](#)

Reporting Scales

Scale scores for STAAR Modified assessments are reported on a horizontal scale. Horizontal scale scores allow for direct comparisons of student performance between specific sets of test items from different test administrations. For all STAAR Modified mathematics, science, and social studies assessments, and the STAAR Modified grades 3–8 reading assessments and grades 4 and 7 writing assessments; a scale score of 3000 represents the final recommended Level II performance standard. In



addition, the standard deviation for those scales was set at 200. For all STAAR Modified EOC English assessments, a scale score of 1500 represents the final recommended Level II performance standard. The standard deviation for the STAAR Modified EOC English scales was set at 100. It is important to note that although Level II scale score values are fixed across horizontally scaled assessments within content areas, Level III scale score values vary across all STAAR Modified assessments. However, these Level III scale score values will stay constant over time (not accounting for the phase-in of standards). The STAAR Modified scale scores represent linear transformations of Rasch-based performance estimates (θ). Specifically, the transformation is made by first multiplying any given θ by a slope (A) and then adding an intercept (B). This operation is described by the equation below:

$$SS_{\theta} = A \times \theta + B \quad (1)$$

A and B in Equation (1) are referred to as the horizontal scaling constants. These same transformations will be applied each year to the Rasch proficiency level estimates for that year's set of test items. Values for the horizontal scaling constants are provided in Tables 5.7 and 5.8 for the STAAR Modified 3–8 and EOC assessments, respectively.

Table 5.7. Horizontal Scaling Constants for STAAR Modified 3–8

Horizontal Scaling Constants for STAAR Modified 3–8		
STAAR Assessment	A	B
Grade 3 reading	244.2943	2877.1200
Grade 3 mathematics	317.5113	2864.1052
Grade 4 reading	232.5136	2797.9457
Grade 4 mathematics	304.9280	2914.9251
Grade 4 writing	386.5413	2919.2129
Grade 5 reading	275.1252	2862.7125
Grade 5 mathematics	389.5123	2988.7041
Grade 5 science	260.2186	2679.6709
Grade 6 reading	281.7367	2852.3700
Grade 6 mathematics	446.3956	3011.6063
Grade 7 reading	278.1452	2880.1194
Grade 7 mathematics	459.9354	3040.9343
Grade 7 writing	408.0328	2862.4930
Grade 8 reading	253.5742	2857.9984
Grade 8 mathematics	457.1662	3030.1730
Grade 8 science	328.2878	2817.4720
Grade 8 social studies	301.6569	2853.0931

Table 5.8. Horizontal Scaling Constants for STAAR Modified EOC

Horizontal Scaling Constants for STAAR Modified EOC		
STAAR Assessment	<i>A</i>	<i>B</i>
Algebra I	405.0300	2972.8630
Geometry	493.9173	3069.6423
Biology	341.6655	2847.2755
English I reading	156.6402	1485.9024
English II reading	151.3773	1478.0503
English I writing	126.1548	1411.6916
English II writing	150.1727	1450.1427
World geography	340.4011	2946.5570

Equating

Used in conjunction with the scaling process, equating is the statistical process that takes into account the slight differences in difficulty across test forms and administrations and allows the scores to be placed onto a common scale. By using statistical methods, TEA “equates” the results of different tests so that scale scores across test forms and testing administrations can be compared. In the 2012–2013 school year, the STAAR Modified equating activities included pre-equating and field-test equating. Refer to [chapter 3, “Standard Technical Processes,”](#) for detailed information about equating.

Pre-Equating

In general, pre-equating is conducted for STAAR Modified assessments. The pre-equating process is one in which a newly-developed test form is linked, before it is administered, to a set of items that appeared previously on one or more test forms. During the 2012–2013 school year, pre-equating was conducted for all STAAR Modified operational assessments, except world history. Equating was not conducted for STAAR Modified world history due to HB 5 legislation.

Field-Test Equating

To replenish the item bank as new tests are developed each year, items must be field-tested and equated to the scale of the item bank. Whenever possible, embedded designs are used to field-test new items so that test takers are unable to distinguish between the field-test items and the operational items on each test form. This results in student performance data that are more reliable. The embedded field-test design is used to collect data on field-test items for STAAR Modified. The equating procedure is described in the technical details and procedures in [chapter 3, “Standard Technical Processes,”](#) for STAAR Modified assessments.



Reliability

Reliability refers to the expectation that repeated administrations of the same test should generate consistent results. Reliability is a critical technical characteristic of any measurement instrument because unreliable instruments cannot be used to make valid interpretations.

During the 2012–2013 school year, the reliability of the STAAR Modified scores was estimated using statistical methods such as internal consistency, classical standard error of measurement, conditional standard error of measurement, and classification accuracy. Refer to [chapter 3, “Standard Technical Processes,”](#) for detailed information about reliability.

Internal Consistency

For the STAAR Modified assessments administered as operational tests in spring 2013, the internal consistency estimates ranged from 0.51 to 0.86. Internal consistency estimates across most grades and content areas were found to be of a similarly high level; however, there was a noticeable decrease in estimates for several grades and content areas (e.g., Algebra I and geometry). For the different student groups, estimates were found to be similar; for grade 8 reading, for example, the reliability for the total group was 0.83, for female only was 0.83, for male only was 0.84, for African American only was 0.83, for Hispanic only was 0.82, and for white only was 0.84.

Because internal consistency estimates typically decrease as the number of test items decrease, internal consistency estimates made at the reporting category level can be noticeably different from those made at the level of the full test. In spring 2013, the internal consistency estimates at the reporting category level were generally lower than at the total score level, indicating that, as expected, interpretations of student reporting category scores are not as reliable as those based on the complete test. For example, the STAAR Modified grade 5 mathematics test contains 40 items, and its overall estimated reliability was 0.82. The STAAR Modified grade 5 mathematics reporting category “Numbers, Operations, and Quantitative Reasoning” contains 14 items. The estimated reliability for the scores in this reporting category was 0.64. Therefore, the lower reliability at the reporting category level should be taken into account when making interpretations of the scores at this level.

Estimates of internal consistency at the overall level as well as by reporting categories and for student groups for all spring 2013 STAAR Modified assessments are provided in [Appendix C](#).

Classical Standard Error of Measurement (SEM)

For the STAAR Modified assessments administered operationally in spring 2013, SEM values range from approximately 2 to 3 raw score points across grades and content areas. The SEM values for the operational STAAR assessments administered in spring 2013 are provided in [Appendix C](#).



Conditional Standard Error of Measurement (CSEM)

Conditional standard error of measurement (CSEM) provides a reliability estimate at each score point on a test. It is an estimate of the average test score measurement error conditional on the proficiency estimate or scale score estimate.

The CSEM values for the primary spring 2013 STAAR Modified assessments are provided in [Appendix C](#).

Classification Accuracy

Classification accuracy provides an estimate of the accuracy of student classifications into performance categories based on current test results. Classification accuracy rates for STAAR Modified assessments ranged from 79.4 to 87.2 percent. The classification accuracy rates for the primary spring 2013 STAAR Modified assessments are provided in [Appendix C](#).

Validity

The results of the STAAR Modified assessments are used to guide educational planning related to the knowledge and skills that students are acquiring in each academic content area. Texas collects validity evidence annually to support the many uses of the STAAR Modified scores. Texas follows national standards of best practice to continue to build its body of validity evidence for the STAAR Modified assessments. TEA also receives ongoing input from the Texas Technical Advisory Committee with regard to plans for collecting validity evidence for the Texas assessment program. The sections that follow describe how different types of validity evidence were collected for the STAAR Modified assessments during the 2012–2013 school year. Refer to [chapter 3, “Standard Technical Processes,”](#) for more detailed information about validity.

Evidence Based on Test Content

The STAAR Modified assessments have been developed to align with content as defined by the TEKS. Content validity evidence is collected at all stages of the test-development process. Nationally established test-development processes for the Texas assessment program are followed while developing the STAAR Modified assessments in order to support the use of the STAAR Modified scores in making inferences about students’ knowledge and understanding of the TEKS.

Because STAAR Modified is a version of STAAR, the test-development processes for both assessments play an intricate role in building validity evidence. To achieve the highest level of content validity, the process of aligning both STAAR and STAAR Modified to the curriculum included review by numerous committees of Texas educators.

RELATIONSHIP TO THE STATEWIDE CURRICULUM

As part of the transition to a high-stakes assessment program in 2012, teachers, curriculum specialists, test-development specialists, college educators, and TEA staff worked together in advisory committees to identify appropriate assessment reporting categories for STAAR, including the STAAR Modified assessments. The input of the advisory committees is reflected in the assessed curricula and test blueprints.

Early in the development process, prototype items were developed for the STAAR Modified assessments. As part of the item-development process, advisory committees and TEA staff reviewed these prototypes to identify how well these items would measure the student expectations to which the items were aligned. These early reviews provided valuable suggestions for item development guidelines and item types. Item development guidelines continued to be refined throughout the test-development process, as various STAAR Modified item review committees shared their feedback in 2011–2012 about how the student expectations could be effectively assessed.

EDUCATOR INPUT

As part of the annual process of item development, committees of Texas educators meet to review the STAAR Modified items and confirm that each item appropriately measures the TEKS to which it is aligned. These item review committees also review the assessment items for content and bias. The committees are made up of Texas K-12 educators, and these committees revise and edit items, as appropriate, prior to test administrations. Item review committees are convened for all STAAR Modified assessments.

TEST DEVELOPER INPUT

Item writers and reviewers follow test development guidelines that explain how content, aligned to given TEKS statements, should be measured. At each stage of development, writers and reviewers verify the alignment of the items with the assessed student expectations. When STAAR Modified was designed as the alternate assessment based on modified achievement standards, special education content specialists developed detailed guidelines so that the modifications made to the STAAR items were consistent. After the items were modified, educator committees for each content area at each grade level reviewed the original STAAR item and the STAAR Modified version of the item to make sure that the modified item still measured the same underlying skill as the original item. In this way, the alignment between the TEKS curriculum and the STAAR items carries through to the STAAR Modified items.

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 support the expectation that the way students respond to items on the STAAR Modified assessments reflects the accurate measurement of the construct.





ITEM TYPE

Student response processes on the STAAR Modified assessments vary according to item types. Across STAAR Modified, three types of items are administered to students: multiple-choice items on all assessments; gridded-response items on mathematics assessments; and written compositions on grades 4 and 7 and English I and II writing assessments.

When item types are initially modified for STAAR Modified, the items are reviewed by educator committees to help ensure that the modifications made the items accessible to the STAAR Modified student population.

Texas also gathers evidence to show that response processes do not advantage or disadvantage one or more student groups. This evidence comes from several sources. When item types were initially considered for inclusion in the STAAR Modified assessments, the item types were pilot-tested to study the way students engage with the various item presentations. After item types were determined to be appropriate for STAAR Modified, evidence about student responses is gathered annually through educator and expert reviews and analyses of individual student responses to these items. Every year, during educator reviews, educators evaluate whether item content for a given item type 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 data are gathered about students' responses. Data such as item difficulty, item point-biserial correlations, and differential item functioning are all evaluated with regard to the item type. For additional information, refer to the Item Analyses section of [chapter 3, "Standard Technical Processes."](#)

SCORING PROCESS

For multiple-choice items, statistical keychecks are conducted for all STAAR Modified assessments during the equating process. Score reliability and validity indices are generated and evaluated for every STAAR Modified assessment (refer to the Item Analyses section of [chapter 3, "Standard Technical Processes"](#)).

Tables 5.9 and 5.10 summarize reader agreement rates by grade and the validity results, respectively, for the STAAR Modified assessments administered in spring 2013 that included written compositions. The reader agreement rate is expressed in terms of absolute agreement between the first reader's score and the second reader's score. Validity is expressed in terms of exact agreement between the score assigned by a given reader and the "true" score approved by TEA.

Table 5.9. Summary of Reader Agreement (Reliability) for 2013 STAAR Modified

STAAR Modified Assessment	Number of Responses Read	Agreement Rate (%) After 2 Readings	Number of Third Readings	Agreement Rate (%) After 3 Readings
Grade 4 writing	14,253	79%	235	98%
Grade 7 writing	13,649	78%	178	99%
English I writing	12,839	77%	65	99%
English II writing	11,390	79%	16	100%

Table 5.10. Summary of Validity Packet Results for 2013 STAAR Modified

STAAR Modified Assessment	Agreement Rate (%)
Grade 4 writing	87%
Grade 7 writing	88%
English I writing	83%
English II writing	82%

Evidence Based on Internal Structure

The internal consistency of the STAAR Modified assessments is evaluated every year using the KR20 for assessments that have only dichotomously scored items (i.e., multiple-choice and gridded-response items). For the STAAR Modified assessments that have a combination of multiple-choice items and written compositions (i.e., the writing assessments), internal consistency is evaluated using the stratified coefficient alpha. These internal consistency evaluations are made for all students and for reported student groups, such as female, male, African American, Hispanic, and white students. Estimates of internal consistency are made for the full assessment as well as for each reporting category within a content area and can be found in the [Reliability](#) section of this chapter.

Evidence Based on Relationships to Other Variables

Another method by which Texas provides validity evidence for the STAAR Modified assessments is by analyzing the relationship between performance on a given STAAR Modified assessment and performance on another STAAR Modified assessment, a process that supports what is referred to as criterion-related validity. By examining this relationship, evidence can be collected to show that the relationships are consistent with those expected based on empirical data.



Correlations between the STAAR Modified content area scale scores were calculated. The correlations across all grades and courses ranged from 0.14 to 0.73. Most correlations are considered moderate correlations, which suggests that scores across content areas are related and neither redundant nor irrelevant. This is expected because the constructs being measured are both academic content areas but assess different types of knowledge and skills. Correlations between writing and science and writing and social studies in grades 3–8 are not included because students do not take these assessments in the same grade. Correlations between the STAAR Modified content-area scale scores at the same grade level can be found in [Appendix C](#).

The correlations between the total score and the STAAR Modified reporting category scores were also calculated within grade and content area. Across all subjects and grades, the correlations between each reporting category and score ranged from 0.49 to 0.90. The magnitudes of these correlations were found to support theoretical relations between reporting categories and the overall test. More specifically, the range of correlations within reading across all grades was 0.66 to 0.90. For mathematics, the range of correlations was 0.49 to 0.88. Science had a correlation range of 0.64 to 0.84, while social studies had a correlation range of 0.68 to 0.87. Lastly, the correlations of the total test score to the reporting category scores for writing ranged from 0.74 to 0.87.

Additional validity evidence was collected in the form of discriminant validity, which demonstrates that the STAAR Modified scores are unrelated to demographic variables (e.g., gender and ethnicity). Theoretically, student characteristics such as ethnicity and gender should not relate to their performance on the assessment; therefore, the lack of meaningful empirical relationships between these measures is expected.

To investigate the relationship between the STAAR Modified scores and demographic variables, correlations were computed specifically for gender and ethnicity. The correlation between the STAAR Modified scores and gender was 0.033, and the correlation between the STAAR Modified scores and ethnicity was 0.060. Both the gender and ethnicity correlations are very small and do not indicate a meaningful relationship between the STAAR Modified scores and either demographic variable.

Evidence Based on Consequences of Testing

In 2011–2012, TEA formally captured and documented the intended and unintended consequences of the STAAR Modified assessment program by administering a consequential validity survey to educators. The consequential validity survey allowed educators to document the extent to which they believed the administration of STAAR Modified has led to changes in certain areas, such as student achievement, impact on teachers, and curricular and instructional reform.

Surveys asking about the intended and unintended consequences resulting from the STAAR Modified assessments were administered to 111 educators attending the fall 2012 STAAR Modified 3–8 standard-setting meetings. The educators who attended the meetings were selected to represent the diversity of Texas educators. Fifty-four percent of the educators taught primarily general education, while the remaining 46



percent taught primarily special education. The three most commonly taught subjects were mathematics (41 percent), ELA (20 percent), and reading (16 percent); and the most commonly taught grades across prekindergarten to grade 12 were grades 3–8. The three most commonly taught grades were grade 8 (16 percent), grades 7 and 6 (12 percent each), and grade 5 (11 percent). The average number of years taught was 16, with two years of teaching being the minimum number of years taught and 45 years being the maximum number of years taught.

Overall, the responses from educators indicate that they believe that the administration of STAAR Modified has led to positive changes in many areas. The majority of educators responded that the STAAR Modified assessment represents what students are expected to know; and because of the administration of STAAR Modified, more students are receiving instruction in higher-order thinking skills. Most educators reported that more students assessed with STAAR Modified are being included in the general education classrooms. The majority of educators also reported that STAAR Modified helps to clarify those skills that are most important for college and career readiness, helps to identify student strengths and weaknesses, and helps schools and teachers obtain additional resources for the low-performing students taking the assessment.

In addition, the majority of educators responded that the administration of STAAR Modified has led to an increase in instruction in the grade-level curriculum and in professional development opportunities, and it has fostered collaboration between general education and special education teachers. The majority of educators also responded that they believe the assessment will provide the feedback about the quality of instruction the students are receiving in the grade-level curriculum and teachers are receiving training and resources to help all students learn.

Educators reported many positive changes due to the administration of STAAR Modified; however, a majority of educators also responded that there are teachers who are feeling more burdened when preparing students to take STAAR Modified and that the public may not perceive STAAR Modified as being helpful for students. Despite these results, more than half of the educators responded that students are responding positively when taking STAAR Modified.

Measures of Student Progress

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

Because STAAR is the basis of the STAAR Modified assessments, the progress measure for STAAR Modified will mirror that of STAAR using a gain score approach. Student progress on STAAR Modified will be classified as *Did Not Meet*, *Met*, or *Exceeded* in relation to a progress target. For more information about the STAAR

progress measure refer to Measures of Student Progress section in [chapter 4](#), “STAAR”.

As the development of the STAAR Modified progress measure proceeds, additional information will be made available on the [STAAR Resources](#) page of TEA’s Student Assessment Division website.

Sampling

In 2012–2013, there were no research studies, audits, or field tests conducted for STAAR Modified. Therefore, sampling was not required.

Test Results

[Appendix C](#) provides scale score distributions and statistics and RSSS conversion tables, as well as mean p-values and reliability estimates by reporting category and content area, for all operational STAAR Modified assessments administered in spring 2013. Table 5.11 shows spring 2013 pass rates for STAAR Modified.



Table 5.11. STAAR Modified Spring 2013 Pass Rates (at the Phase-in 1 Standard)

Content Area	Grade/Course	Pass Rate
Mathematics	Grade 3	64%
	Grade 4	66%
	Grade 5	64%
	Grade 6	60%
	Grade 7	60%
	Grade 8	61%
	Algebra I	44%
	Geometry	42%
Reading	Grade 3	71%
	Grade 4	69%
	Grade 5	76%
	Grade 6	66%
	Grade 7	67%
	Grade 8	68%
	English I	71%
	English II	68%
Writing	Grade 4	58%
	Grade 7	68%
	English I	61%
	English II	75%
Science	Grade 5	56%
	Grade 8	69%
	Biology	50%
Social Studies	Grade 8	62%
	World geography	67%

