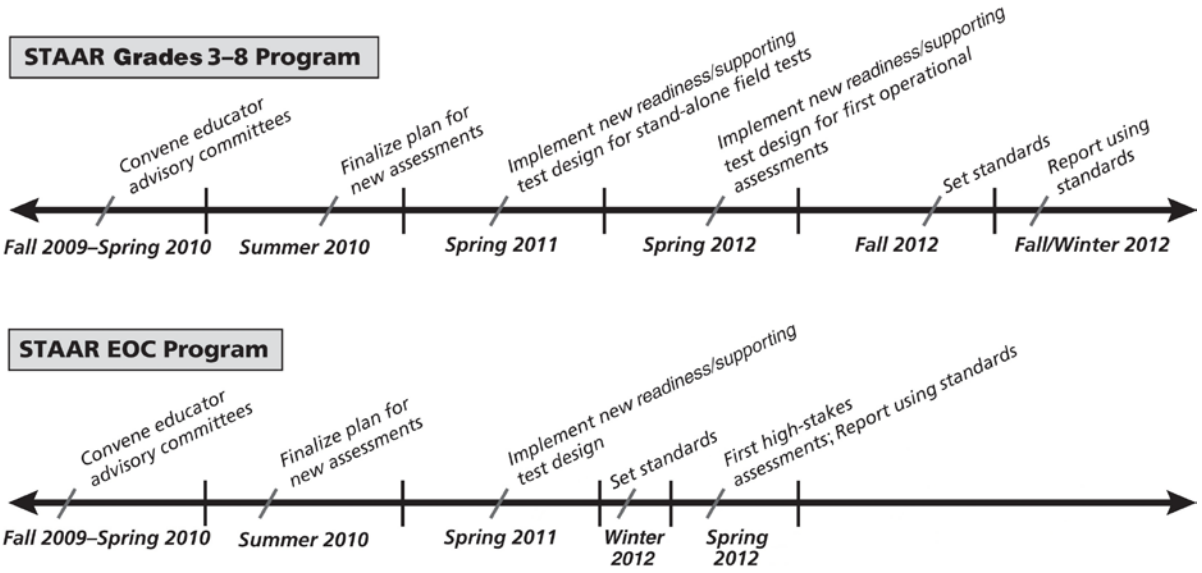
Chapter 1
Timeline for the Development and Implementation of the State of Texas Assessments of Academic Readiness (STAAR) Grades 3–8 and End-of-Course (EOC) General Assessments

Introduction

Beginning in the 2011–2012 school year, the Texas Education Agency (TEA) will implement the STAAR program, which will include new assessments in grades 3 through 8 and twelve EOC assessments. This section provides a timeline for the development and implementation of the STAAR grades 3–8 and STAAR EOC assessments for the general student population and includes a history of the Texas assessment program, comparisons of STAAR and Texas Assessment of Knowledge and Skills (TAKS), information on STAAR test development and field testing, and a proposed STAAR test administration schedule. Additional details regarding the timeline can be found in subsequent sections of this report.

A general timeline for the development and implementation of the STAAR assessment program is shown below.



History of the Texas Assessment Program

In 1979, the state of Texas implemented a statewide testing program that since its inception has grown in size, rigor, and scope following changes in policy and legislation. As required by state statute, Texas assessed minimum skills in reading, mathematics, and writing with the Texas Assessment of Basic Skills (TABS) tests (1981–1984) and then with the Texas Educational Assessment of Minimum Skills (TEAMS) tests (1985–1990).

The implementation of the Texas Assessment of Academic Skills (TAAS) testing program in 1990 shifted the focus of assessment from minimum skills to academic skills. TAAS mathematics, reading, and writing tests were administered to students in grades 3, 5, 7, 9, and 11. A Spanish-language version of the grade 3 test was available for eligible English language learners (ELLs).

During the 1993–1994 school year, the TAAS testing program was reconfigured. Between 1994 and 2002, TAAS was administered every spring to students in grades 3 through 8 and 10 in reading and mathematics; grades 4, 8, and 10 in writing; and grade 8 in science and social studies. The Spanish-language TAAS program was expanded in the 1996–1997 school year so that Spanish-language TAAS tests were available to eligible students in grades 3 through 6.

In 1995, the 74th Texas Legislature included EOC assessments as an option for meeting graduation requirements. Between 1995 and 2002, EOC assessments were administered in Algebra I, English II, biology, and U.S. history.

In 1999, the 76th Session of the Texas Legislature enacted Senate Bill (SB) 103, which required the development of a new statewide testing program. The new testing program, subsequently named TAKS, replaced TAAS as the primary statewide student assessment program in spring 2003. The TAKS tests were designed to measure the extent to which a student has learned and is able to apply the defined knowledge and skills at each grade level tested. Every TAKS test was directly aligned to the state content standards, the Texas Essential Knowledge and Skills (TEKS). When TAKS replaced TAAS in the 2002–2003 school year, EOC assessments were no longer administered with the exception of Algebra I, which remained a voluntary operational assessment.

In 2005, Governor Rick Perry issued Executive Order RP53, which called for an increase in college-readiness programs in Texas public schools and authorized "the development of a series of voluntary end-of-course assessments in science, mathematics, and other subjects currently assessed by the eleventh grade Texas Assessment of Knowledge and Skills, to measure student performance…." In response to the order, TEA began to develop new EOC assessments in geometry, biology, chemistry, physics, and U.S. history.

In 2007, the 80th Session of the Texas Legislature enacted SB 1031, which expanded the scope of the EOC program. This legislation required the phase-out of the current high school TAKS assessments and replaced them with EOC assessments to be administered beginning with students entering the ninth

grade during the 2011–2012 school year. SB 1031 also required the development of six additional EOC assessments: Algebra II, world geography, world history, and English I, II, and III.

In 2009, the 81st Session of the Texas Legislature enacted House Bill (HB) 3, which required that

- new assessments be developed at grades 3–8 and linked to EOC assessments and to college readiness;
- college readiness be defined as "the level of preparation a student must attain in English language arts and mathematics courses to enroll and succeed, without remediation, in an entry-level general education course for credit in that same content area" at a general academic teaching institution or an institution that offers associate degrees or certificates;
- performance on Algebra II and English III indicate college readiness based on studies correlating performance on the EOC assessments with college readiness;
- performance on Algebra I and English II correlate to performance on Algebra II and English III;
- performance on English I correlates to performance on English II;
- performance on grade 8 assessments correlate to performance on Algebra I and English I;
- performance on grades 3–7 assessments correlate to assessments in the same content area at the next grade;
- performance on college readiness standards be set on certain science and social studies EOC assessments if a link is established between performance on the assessment and college readiness; and
- the commissioner of education and commissioner of higher education establish college readiness performance standards for Algebra II and English III, and all other performance standards be established by the commissioner of education.

Changes from TAKS to STAAR

To meet legislative requirements, the new STAAR program will differ significantly from the current TAKS program in several ways. In the following section the key differences are outlined, specifically with regard to the rigor and test design of the STAAR program. A detailed document comparing the attributes of TAKS and STAAR is located at the end of this section.

Increased Rigor

Item development guidelines have remained generally consistent for TAKS since 2001, although modifications have been made to address changes in the content standards (TEKS). Performance standards were set by the State Board of Education (SBOE) in 2002, and were phased in over a three-year period. These performance standards remained the same until the vertical scale was established in reading and mathematics at grades 3–8. At that time, some changes in the performance standards were required to

implement the vertical scale. Over time, increasing numbers of students have begun to reach higher performance standards, making differentiation of "Commended Performance" difficult because too few test items are currently rigorous enough to reflect this performance category.

STAAR is designed to better assess students' academic achievement at all performance levels. STAAR will be more rigorous than TAKS in the following ways.

- STAAR assessments will assess content and skills from the TEKS at a greater depth and higher level of cognitive complexity.
- STAAR assessments will contain more items to facilitate the measurement of a student's knowledge and skills at all performance levels.
- Some items will assess more than one student expectation from the TEKS, allowing skills to be tested in more integrated and authentic ways.
- STAAR assessments will focus on the student expectations that are necessary both for success in the current grade or course and for success in the next grade or course. Algebra II and English III will emphasize the student expectations that are necessary both for success in those courses and for postsecondary readiness.
- In writing, students will be required to respond to two writing tasks (personal narrative, literary, expository, persuasive, or analytic), rather than one writing task, as was required on TAKS.
- In science and mathematics assessments, the number of open-ended (griddable) items on most tests will increase. Griddable items are more rigorous because they require students to derive answers independently rather than to select a correct response from a list of possible responses.
- In grades 5 and 8 science, there will be an increased focus on promoting preparedness for high school science through an emphasis on the content and skills in grades 3–5 and 6–8 that link directly to the high school content standards for biology, chemistry, and physics.
- Performance standards will be set using empirical data gathered from studies that link year-to-year performance from grades 3–8 to high school and from specific courses (Algebra II and English III) to postsecondary readiness.
- Expectations for student performance on STAAR will be raised to achieve the goal of graduating students who are college and career ready.
- Performance standards will be reviewed at least once every three years and, if necessary, adjusted to ensure that the assessments maintain a high level of rigor.
- Empirical studies will be used to compare performance on STAAR with national and international norm-referenced assessments.

Course Specificity for EOC Assessments

High school TAKS currently assesses students at grade 9 in reading and mathematics and students at grades 10 and 11 in English language arts (ELA), mathematics, science, and social studies. The content

assessed in any one high school TAKS assessment typically includes content from multiple courses. For example, the TAKS grade 11 mathematics assessment draws on content from Algebra I, geometry, and grade 8 mathematics.

The STAAR EOC assessment model differs from TAKS in that each EOC assessment will cover only the content from a particular course (for example, Algebra II will assess only Algebra II content). This model allows for a more focused assessment that is aligned to the course content on which the student has received instruction that year.

It should be noted that there are specific course requirements for students depending on their graduation programs, but there is not a state-mandated course sequence. However, the typical course sequence that most students follow is provided below.

Typical High School Course Sequence

	English	Mathematics	Science	Social Studies
Grade 9	English I	Algebra I	Biology	World Geography
Grade 10	English II	Geometry	Chemistry	World History
Grade 11	English III	Algebra II	Physics	U.S. History

Performance Standards

STAAR assessments will continue to measure student performance as well as academic growth, as required by HB 3. STAAR differs from TAKS, however, in that tests in mathematics and reading must be linked from grade to grade and to postsecondary readiness performance standards for the Algebra II and English III EOC assessments. Because the STAAR performance standards will be set as an aligned system across grades and courses within a content area (from grades 3–8 through high school), performance on STAAR assessments can provide early indications of each student's preparedness for secondary and postsecondary education.

STAAR Test Development

College and Career Readiness Standards

In 2008, the Texas Higher Education Coordinating Board (THECB) and the commissioner of education adopted the Texas College and Career Readiness Standards (CCRS). Since then, the SBOE has incorporated these standards into the TEKS for the four foundational content areas: English language arts, mathematics, science, and social studies. The following chart gives more information regarding this timeline.

Incorporation of the CCRS into the TEKS

Foundational Content Areas	TEKS Revisions Process Incorporated CCRS
English language arts	May 2008
Mathematics	January 2009
Science	March 2009
Social Studies	May 2010

Now that the CCRS have been incorporated into the TEKS content standards, these skills will be tested on STAAR assessments.

Readiness Standards and Supporting Standards

TEA has worked with educator advisory committees to design a more focused and rigorous assessment program. The TEKS content standards were reviewed by educator committees (K–12 and higher education) to determine which standards are assessable. Once that determination was made, committees then provided guidance as to which of these standards are readiness standards and should be emphasized in the assessments and which standards can be assessed on a supporting basis.

In general, readiness standards

- are essential for success in the current grade or course;
- are important for preparedness for the next grade or course;
- support college and career readiness;
- necessitate in-depth instruction; and
- address broad and deep ideas.

In general, supporting standards

- may be emphasized in a subsequent year or course although they are introduced in the current grade or course;
- may have been emphasized in a previous year or course although they are reinforced in the current grade or course;
- play a role in preparing students for the next grade or course but not a central role; and
- address more narrowly defined ideas.

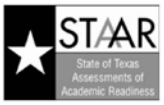
For more information about readiness standards and supporting standards, see Chapter 2.

Test Blueprints

In addition to making recommendations about readiness and supporting standards, educator advisory committees assisted TEA in making decisions regarding test blueprints. STAAR blueprints were

structured to emphasize student performance on readiness standards. Blueprints for STAAR grade 8 science and STAAR Algebra II are included here as examples. All of the STAAR blueprints are found on the TEA website at http://www.tea.state.tx.us/student.assessment/staar/. More detailed information regarding STAAR test development can be found in Chapter 2.

STAAR Grade 8 Science Blueprint



Scientific Investigation and Reasoning Skills is not a separate reporting category. These skills will be incorporated into at least 40% of the test questions from reporting categories 1–4 and will be identified along with the content standards.				
Reporting Categories	**Number of Standards**		**Number of Questions**	
Reporting Category 1: Matter and Energy	Readiness Standards	5	**14**	
	Supporting Standards	7		
	Total	12		
Reporting Category 2: Force, Motion, and Energy	Readiness Standards	2	**12**	
	Supporting Standards	6		
	Total	8		
Reporting Category 3: Earth and Space	Readiness Standards	5	**14**	
	Supporting Standards	10		
	Total	15		
Reporting Category 4: Organisms and Environments	Readiness Standards	3	**14**	
	Supporting Standards	11		
	Total	14		
Readiness Standards	**Total Number of Standards**	**15**	**60%–65%**	**32–35**
Supporting Standards	**Total Number of Standards**	**34**	**35%–40%**	**19–22**
Total Number of Questions on Test			**50 Multiple Choice 4 Griddable 54 Total**	

STAAR Algebra II Blueprint

Reporting Categories	Number of Standards		Number of Questions	
Reporting Category 1: Properties and Attributes of Functions	Readiness Standards	3	**8**	
	Supporting Standards	2		
	Total	5		
Reporting Category 2: Representational Tools to Solve Problems	Readiness Standards	3	**8**	
	Supporting Standards	2		
	Total	5		
Reporting Category 3: Properties of Quadratic Functions	Readiness Standards	4	**12**	
	Supporting Standards	3		
	Total	7		
Reporting Category 4: Representations of Quadratic Relations	Readiness Standards	1	**6**	
	Supporting Standards	6		
	Total	7		
Reporting Category 5: Properties of Square Root Functions	Readiness Standards	1	**5**	
	Supporting Standards	6		
	Total	7		
Reporting Category 6: Properties of Rational Functions	Readiness Standards	1	**5**	
	Supporting Standards	6		
	Total	7		
Reporting Category 7: Properties of Exponential and Logarithmic Functions	Readiness Standards	2	**6**	
	Supporting Standards	4		
	Total	6		
Readiness Standards	**Total Number of Standards**	**15**	**60%–65%**	**30–33**
Supporting Standards	**Total Number of Standards**	**29**	**35%–40%**	**17–20**
Total Number of Questions on Test			**45 Multiple Choice 5 Griddable 50 Total**	

Field-Test Plans for STAAR

Background

Field testing is an important part of the item and test development process. Field testing allows for the development of tests that are fair for all student groups, are of high quality, are legally defensible, and can withstand rigorous scrutiny when evaluated relative to professional standards. Field testing helps determine whether test items are valid and reliable measures of what students know and can do. Field-test data are necessary for constructing tests, setting performance standards, conducting validity studies, and implementing growth measures required by statute.

In 2007, the 80th Texas Legislature passed SB 1031, which required TEA to conduct a study to review the sample size and sampling procedures used in field testing. In 2008, TEA incorporated the results of this study into a report for the 81st Texas Legislature titled “A Report on Field-Testing for the

Texas Assessment Program." The report outlined the actions that should be taken to reduce the field-test burden and can be found on the TEA website at http://ritter.tea.state.tx.us/comm/leg_reports/2008/08FieldTest.pdf. Subsequent efforts to reduce field testing for the new STAAR program are outlined in this chapter.

Field-Test Plans for STAAR Grades 3–8 Assessments

For the new STAAR grades 3–8 program, the overall field-test burden on students and districts has been reduced through the embedding of field-test items whenever possible. Field-test items for STAAR grades 3–8 mathematics, reading, writing (revising and editing items), social studies, and science will all be embedded. One exception to this new policy is that in spring 2011, stand-alone field tests are planned for the STAAR writing assessments for grades 4 (English and Spanish) and 7; however, after this single occurrence, stand-alone field tests will be conducted for grade 4 writing only every three years (beginning in 2014). Grade 7 writing field tests will be embedded after the initial 2011 stand-alone field test. Note that performance on field-test questions is not calculated as part of reported test scores and participation in the stand-alone field tests is required for selected campuses. A field-test timeline for all subjects and grades follows.

STAAR Grades 3–8 Field-Test Timeline

STAAR Grades 3–8 Assessment	Spring 2011	Spring 2012	Spring 2013	Spring 2014
Grades 3–8 Mathematics (English)	Embedded in operational TAKS*	Operational with embedded field test	→→→	→→→
Grades 3–5 (Spanish) Mathematics	Embedded in operational TAKS*	Operational with embedded field test	→→→	→→→
Grades 3–8 (English) Reading	Embedded in operational TAKS*	Operational with embedded field test	→→→	→→→
Grades 3–5 (Spanish) Reading	Embedded in operational TAKS*	Operational with embedded field test	→→→	→→→
Grade 4 Writing (English and Spanish)	Stand-alone field test	Operational with embedded field test for revising and editing	→→→	Stand-alone field test
Grade 7 Writing	Stand-alone field test	Operational with embedded field test	→→→	→→→
Grade 5 (English and Spanish) Science	Embedded in operational TAKS*	Operational with embedded field test	→→→	→→→
Grade 8 Science	Embedded in operational TAKS*	Operational with embedded field test	→→→	→→→
Grade 8 Social Studies	Embedded in operational TAKS*	Operational with embedded field test	→→→	→→→

* Last year for current TAKS

Field-Test Plans for STAAR EOC Assessments

Since no operational test forms exist initially in which to embed field-test items, stand-alone field testing is required during the first year an EOC assessment is introduced. In all subsequent years, field-

test items will be embedded in operational forms. By spring 2012, all twelve EOC assessments will be operational and will contain embedded field-test items. As previously noted in the STAAR grades 3–8 field-test section, performance on field-test questions is not calculated as part of reported test scores. In addition, embedded field testing will reduce the burden on students and districts while still providing the STAAR EOC program with the data necessary to create high-quality assessments.

The field-test timeline for implementation of the twelve EOC assessments—from stand-alone field tests to operational tests with embedded field-test items—follows.

EOC Field-Test Timeline

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

In addition to mandatory field-test requirements, several operational tests over the past few years have required mandatory participation, and some assessments will require mandatory participation in the next few years. This data collection is needed to ensure adequate data for future EOC test construction, standard setting, and other validity studies.

In summary, the overall field-test burden on students and districts will be reduced through the embedding of field-test items, whenever possible, in the STAAR program.

Number of Testing Days for TAKS and STAAR

The following chart compares the number of testing days for field-test and operational administrations for the TAKS assessment program and for STAAR. STAAR significantly increases the number of testing days at the high school level because of the increase in the number of assessments students will be taking. Currently on TAKS there is a total of 25 testing days, including exit level retest administrations. With three testing opportunities each year, STAAR will require up to 45 testing days when it is fully implemented.

	TAKS Assessment Program	STAAR Assessment Program
Number of Testing Days	• Grade 3 – reading and mathematics (2 days) • Grade 4 – reading, mathematics, and writing (3 days) • Grade 4 – writing field test (1 day) • Grade 5 – reading, mathematics, and science (3 days; up to 7 days for SSI retesting) • Grade 6 – reading and mathematics (2 days) • Grade 7 – reading, mathematics, and writing (3 days) • Grade 7 – writing field test (1 day) • Grade 8 – reading, mathematics, science, and social studies (4 days; up to 8 days for SSI retesting)	• Grade 3 – reading and mathematics (2 days) • Grade 4 – reading, mathematics, and writing (4 days; writing now a 2-day administration) • Grade 5 – reading, mathematics, and science (3 days; up to 7 days for SSI retesting) • Grade 6 – reading and mathematics (2 days) • Grade 7 – reading, mathematics, and writing (4 days; writing now a 2-day administration) • Grade 8 – reading, mathematics, science, and social studies (4 days; up to 8 days for SSI retesting)
	Total – 19 (27 with SSI retesting)	*Total – 19 (27 with SSI retesting)*
	• Grade 9 – reading and mathematics (2 days) • Grade 9 – reading field test (1 day) • Grade 10 – ELA, mathematics, science, and social studies (4 days) • Grade 10 – ELA field test (1 day) • Grade 11 (Exit Level) – ELA, mathematics, science, and social studies (4 days; up to 16 days for retesting) • Exit Level – ELA field test (1 day)	• English I (2 days) • English II (2 days) • English III (2 days) • Algebra I (1 day) • Geometry (1 day) • Algebra II (1 day) • World History (1 day) • World Geography (1 day) • U.S. History (1 day) • Biology (1 day) • Chemistry (1 day) • Physics (1 day) • 2 additional testing opportunities per year
	Total – 13 (25 with Exit Level retesting)	*Total – 15 (45 with retesting)*

STAAR Implementation Policies

Time Limits on Tests

As the state transitions from the TAKS program to the STAAR program, one of the implementation policies TEA is considering is a policy to limit the amount of time a student spends taking a STAAR assessment on a given day. This consideration is based on advice from advisory committees to align Texas's testing policies with other state and national assessments. SAT, ACT, and AP exams are all administered in timed settings, for example, so there is concern that high school students are not being

adequately prepared for testing in this environment. In addition, with the increased testing that will be necessary at high schools with twelve tests requiring retest opportunities as opposed to four exit level tests under TAKS for which retests are offered, there is interest by school district personnel in administering two EOC sessions in one day. Although an individual student would likely not take two EOC assessments in one day, a school district would be able to schedule two EOC testing sessions in a single day. For STAAR grades 3–8 assessments it is likely that schools, at a minimum, will not be allowed to administer the tests beyond the end of the regular school day, and additional time limits could be considered for tests administered at these grades. In making this decision, TEA will consider the appropriate time needed to complete the assessments, student fatigue, and other related factors.

Test Security

In June 2007, TEA introduced a comprehensive 14-point plan designed to assure parents, students, and the public that test results are meaningful and valid. Maintaining the security and confidentiality of the Texas state assessment program is crucial for ensuring valid test scores and providing standard and equal testing opportunities for all students. Given the high stakes associated with student performance and the increasing complexity of the STAAR program, test administration personnel will face new challenges in managing the testing requirements at the local level. TEA will continue to publish test administrator manuals, a test security supplement, conduct face-to-face training, and provide online security training modules as part of the 14-point security plan. These resources will continue to provide districts guidance in implementing test-security requirements and to foster best practices for maintaining a secure testing program. Aspects of this plan will be implemented as the transition to the STAAR program continues. The 14-point test security plan can be found on the TEA website at http://ritter.tea.state.tx.us/student.assessment/admin/security/14point_Recommendations_and_Timelines.pdf.

Use of Statistical Analysis

One component of the 14-point test security plan is the use of statistical analysis to identify irregular patterns of test answers that may indicate cheating to augment other detection methods already in use, such as multiple-mark analysis. TEA will outline a new process, with advice from experts and school district personnel, to address the use of statistical analysis with the STAAR assessment program. The use of statistical methods will take place within a larger investigative process that includes the collection of additional evidence, such as locally maintained seating charts, reports of testing irregularities, and records of test security and administration training for campuses.

TEA will pilot statistical measures with data from the 2009 and 2010 TAKS administrations and will generate statewide campus metrics in summer 2011 for TAKS grades 3–11 primary administrations. Beginning in 2012, statewide metrics will be applied to the STAAR grades 3–8 and EOC assessments.

Security Challenges Associated with the STAAR EOC Program—Expanded High School Testing

Once STAAR EOC assessments are operational in 2012, there will be three separate administrations—once each in the fall, spring, and summer. In each administration, there will be 12 different tests that will be administered in both paper and online modes. Further, all 12 of the EOC assessments will be available for retest opportunities, so there is the potential that students will be taking more than four tests in a given administration window beginning with the summer 2012 administration. Because current statute allows students to retest an EOC assessment for any reason, there are additional security challenges that must be addressed with the need to allow more days for testing.

Because of the number of high-stakes assessments to be administered at the same time, it is likely that the testing windows will span several weeks to allow districts sufficient opportunity to assess all students with multiple assessments. Secure test materials will be in districts for a longer duration than with the current program, potentially leading to more test-item exposure. In addition, the nature of the English I, II, and III writing tasks pose security challenges similar to those that exist with the current program. Based on these concerns, TEA will continue to implement the 14-point test security plan and will investigate other test-administration-specific policies that can provide the most secure testing program possible as well as support districts in their efforts to maintain test security. Such administration-specific policies include

- the scrambling of test items on different test forms during administrations;
- the use of multiple test forms during administrations; and
- assigned testing days versus testing windows for specific assessments.

As these new policies are being explored, they will be weighed against a significant cost increase and delays in the reporting of results due to additional time needed to ensure that reporting is accurate.

Testing Accommodations

Other implementation policies TEA will need to consider for the STAAR program relate to testing accommodations. Testing accommodations are practices or procedures that provide equitable access to grade-level content standards during instruction and assessment for all students. Testing accommodations have been part of the TAKS program and will be incorporated into the STAAR program; however, TEA is evaluating all accommodations to determine which ones will continue in the STAAR program and which accommodations will be added. As part of the TAKS program, an accommodated form has been offered for students receiving special education services who take the general assessment. With the STAAR program, TEA is considering the elimination of the separate accommodated form and building in specific accommodations to the general STAAR assessments for these students. Also, TEA is exploring the possibility of standardized oral administrations for the STAAR program, using an online testing format. If found feasible, information about standardized oral administrations will be communicated to districts by the end of the 2010–2011 school year.

Grades 3–8 Paper Administrations

Spring 2011 marks the last primary administration of TAKS for grades 3–8. STAAR grades 3–8 assessments will be operational beginning in spring 2012. These assessments will be administered on paper only due to multiple factors including

- lack of technology resources at the district level to administer a single-day administration for all students on a campus as noted in the technology survey found at the following link http://ritter.tea.state.tx.us/comm/leg_reports/2008/2009OnlineReadinessReport.pdf;
- the cost associated with developing and administering both paper and online assessments; and
- limited participation in previous optional online assessments at the middle school level for TAKS.

Student Success Initiative and Other Statutory Requirements in 2011–2012

Because performance standards for STAAR grades 3–8 will not be set until after the spring 2012 administration, only raw-score information (the number of questions correct out of the total number of questions on the test) will be available. Therefore, Student Success Initiative (SSI) retest opportunities for STAAR grades 5 and 8 reading and mathematics will not be available in May and June of 2012. For the 2011–2012 school year, districts must use other information in addition to raw-score information to make promotion/retention decisions. When making promotion decisions for students in grades 5 and 8, statute requires that districts consider the following academic information:

- the recommendation of the student's teacher(s),
- the student's grade in the subject or course, and
- the student's potential for achievement or proficiency in the subject or course.

More information about SSI procedures in the absence of passing scores on STAAR will be provided in the 2011–2012 Grade Placement Committee (GPC) Manual.

In a similar fashion, STAAR scores for grades 3–8 will not be available to meet statutory requirements such as those that call for the use of passing scores to determine eligibility of limited English proficient students to exit special language programs and be reclassified as English proficient. During the 2011–2012 school year, information will be disseminated to provide guidance to school districts about procedures to follow for this and other program-related purposes.

Grades 3–8 Administration Schedule

The following chart provides an overview of the administration schedule for STAAR grades 3–8 currently planned.

STAAR Grades 3–8 Operational Test Administrations

	2012	2013	2014	2015
March or Early April	Grades 4 and 7 writing	Grades 4 and 7 writing	Grades 4 and 7 writing Grade 4 writing field test	Grades 4 and 7 writing
Early April	Grades 5 and 8 reading and mathematics SSI administration	Grades 5 and 8 reading and mathematics SSI administration	Grades 5 and 8 reading and mathematics SSI administration	Grades 5 and 8 reading and mathematics SSI administration
Late April	Grades 3–8 administration	Grades 3–8 administration	Grades 3–8 administration	Grades 3–8 administration
May	N/A	Grades 5 and 8 reading and mathematics SSI retest administration	Grades 5 and 8 reading and mathematics SSI retest administration	Grades 5 and 8 reading and mathematics SSI retest administration
June/July	N/A	Grades 5 and 8 reading and mathematics SSI retest administration	Grades 5 and 8 reading and mathematics SSI retest administration	Grades 5 and 8 reading and mathematics SSI retest administration

EOC Administration Schedule

Once all EOC assessments are operational in spring 2012, they will be administered three times a year. English I, II, and III will be administered late in March or early in April to allow for the additional time needed to score essays and short-answer reading responses. For the spring administrations, the remaining nine EOC assessments will be administered in May, as legislatively mandated. All twelve EOC assessments will also be administered at the end of the summer and fall semesters. However, there will not be a fall administration in the 2011–2012 school year as the state makes the transition to STAAR. The plan for the phase-out of high school TAKS and the phase-in of EOC assessments is shown in Chapter 7. All EOC administrations will be offered on paper and online.

Online EOC Testing

As part of new EOC administration policies TEA has been investigating, a netbook computer study was conducted to determine if student results from an online test on a standard-sized screen (approximately 14 inches and larger) were comparable to those on a netbook-sized screen (less than 14 inches). Based on the results of the study, TEA will not preclude a district from using or purchasing netbooks for use in online testing, allowing districts to use current inventory or purchase a less expensive alternative. The results of the netbook study can be found at the TEA website at http://www.tea.state.tx.us/student.assessment/reports/.

STAAR Scoring and Reporting

Like the TAKS program, STAAR assessments will report student results in terms of raw scores, scale scores, and performance levels. However, the values of the scale scores, the performance-category labels, and the policy definitions associated with those labels for STAAR will be different from TAKS. See Appendix A for more information about the STAAR performance categories.

STAAR EOC assessment results will be reported in spring 2012 and grades 3–8 will be reported in late fall 2012 or early 2013. This timing is based on when performance standards will be established for these two components of the STAAR assessment system. Both a vertical and horizontal scale will be used in the STAAR program.

For STAAR EOC reporting after the spring administrations in 2012 and beyond, due to the legislative requirement to test no earlier than the first full week in May, reporting EOC assessment results in a timely manner will be a challenge. The chart below shows the difficulty of testing all students with nine different assessments (English I, II, and III will have been tested in late March/early April), returning testing materials to the contractor, processing the tests taken with a paper version, and making reports available to districts online as well as providing paper student reports to the districts for distribution to their students.

In 2012, districts cannot begin to administer EOC assessments, except the English assessments, before May 7. Other dates on the following chart are tentative due to decisions yet to be made regarding the methods required to determine final scale scores needed for reporting.

Draft 2012 EOC Testing and Reporting Timeline

Activity	Date
Test Administration	May 7–18
Shipment of Scorable Test Materials	May 21
Processing and Scoring	May 23–May 31
Posting of Online Individual Student Reports and Rosters	May 30–June 1
Posting of Online Summary Reports and Data Files	June 1–June 4
Paper Reports Distributed to School Districts	By June 8

In spring 2012, districts are required to use test results as 15% of students' course grades. However, this requirement may be challenging to implement since districts may not have online test results prior to the end of the school year and will receive paper results after the school year has ended. After spring 2012, the reporting timeline may be able to be compressed by a day or two; however, this timeline may still present challenges for reporting test results in a timely manner.

Texas Assessment Management System (Student Assessment Data Portal)

Students, parents, and teachers will be able to access results through the data portal legislated by HB 3. The portal is a secure system that will provide new abilities to view reports, track student progress, provide assessment data to institutions of higher education, and provide assessment information to the general public.

Student Portal

Via the portal, parents and students can access assessment information across administrations and years. Users can compare their results to aggregated campus, district, and state performance. Interpretive

information will also be available to parents and students and will provide explanations about the test results.

Two types of reports will be provided to parents and students:

- a report that will show historical and current assessment results for which comparisons can be made between a student's scores and the scores of the student's campus, district, and the state
- a report that will show progress toward graduation for high school students in relationship to the students' assessment results

Teacher Portal

The portal will also provide teachers access to student performance data. Teachers will be able to access their students' assessment data for use in developing strategies for improving student performance. District personnel will be able to compare information across campuses, including historical and growth information, and the portal will also allow a student's performance to be viewed in relation to other groups to which the teacher has access, including campus, district, and statewide averages. Specific features of the Teacher Portal include

- viewing student assessment results individually or by group;
- comparing student results among groups, campuses, districts, or statewide;
- examining a distribution of student performance; and
- accessing individual student scale scores and objective scores.

Analytic Reporting

In addition, the analytic reporting system through the portal allows users to analyze results in order to compare current and historical data as well as to perform comparisons of classes to classes, classes to campuses, campuses to campuses, campuses to districts, districts to districts, and districts to the state. Campuses and districts will also be able to disaggregate data so that different demographic and program information groups may be examined, enabling easy access to cross-section analysis of the assessment data. Other information, such as locally developed assessment results or norm-referenced test results, could also be loaded locally into the data portal, allowing analytic reporting that could compare those scores with STAAR results. The following table outlines specific components of the portal, including timelines for implementation.

Texas Assessment Management System (Data Portal)—Implementation Timeline

Students and Parents Portal access for students and parents is planned for December 2010. Login information will be provided to students and parents via a student's Confidential Student Report which is provided to students after each test administration.
School Districts and District Teachers Portal access for authorized district and campus personnel will be provided during fall 2010. The capability for teachers to access their students' information will require a link between the teachers and their students. District personnel will be able to compare like information across campuses. This will include historical and growth information. For district and campus personnel, the portal will also be used for other assessment functions, such as providing student information for assessment purposes, ordering materials for assessments, testing online and receiving assessment reports after each administration.
Public Institutions of Higher Learning Providing results to public institutions of higher learning is planned for summer 2011. Current plans call for use of the Texas Records Exchange (TREx) system which currently provides student record and transcript information for public institutions of higher learning. This new functionality will allow authorized employees of a public institution of higher education to readily access the individual assessment data of students applying for admission for use in developing strategies for improving student performance.
Graduation Requirements Students and parents will have the ability to track a student's progress on the assessment instrument requirements for graduation (same access as described above). This will include TAKS exit level as well as STAAR graduation requirements. This is available for TAKS in December 2010.
Public Access General student assessment data (non confidential information) is currently accessible and available to the public. Enhancements will be provided in the future so that the data can be disaggregated at multiple levels.
Data Comparisons Comparisons at the campus, district, and state levels will be available in spring 2011 through "analytic" reporting. Comparisons at the classroom level will not be possible until a link between teachers and students is made available for Texas student assessment data through the Public Education Information Management System (PEIMS). This linkage to PEIMS is planned for summer 2011.

Plans for STAAR Assessment Reporting System

The assessment reports that will be prepared for the STAAR program will be designed to provide comprehensive, easy-to-understand results for students, parents, educators, and the public. The reports for the STAAR 3–8 and EOC assessments will be designed with advice from advisory groups to ensure that the results are clearly communicated to their intended audiences. The development of the new STAAR reporting system also provides an opportunity to design assessment reports that are closely aligned with the reporting requirements for the state and federal accountability systems. To the extent possible, assessment reports and data files prepared by the test contractor will include information that will enable districts and campuses to more easily determine what assessment data will be used to determine their state and federal accountability ratings. These assessment results will also be included in the Texas Assessment Management System.

Why Scales are Used on Tests

The basic score on any test is the raw score, which is the number of items correct. However, the raw score alone does not present a broad picture of test performance because it can be interpreted only in terms of a particular set of test questions. When new test forms are administered in subsequent administrations, most questions on the new forms are different. The set of questions on one test may be slightly easier or slightly harder than the set of questions that were on another test. Because the overall difficulty of the tests may vary, the raw scores or percentage correct cannot be directly compared to indicate differences in student performance.

Unlike raw scores, scale scores do allow direct comparisons of student performance between specific sets of test questions from different test administrations. A scale score is a conversion of the raw score onto a "scale" that is common to all test forms for that assessment. The scale score takes into account the difficulty level of the specific set of questions. There are two types of scales, vertical and horizontal, used on assessments, and the two different types are chosen based on specific assessment attributes. Specific information about vertical and horizontal scales follows.

Vertical Scale

A vertical scale allows a student's scale score in one grade to be compared to the student's scale score in the next grade in the same content area. Currently a vertical scale for STAAR grades 3–8 in reading and mathematics is legislatively mandated, and this is the only vertical scale planned for the STAAR program. A vertical scale will be used for these assessments because they are administered each year in grades 3–8; have substantial content overlap for consecutive grades; and show incremental increases in difficulty from grade to grade. A vertical scale for the grades 5 and 8 STAAR science assessments is not planned due to the differences in science content covered at the different grades. A vertical scale for grades 4 and 7 STAAR writing assessments is not planned because the difficulty levels of the grades 4 and 7 writing assessments do not overlap sufficiently to create a meaningful vertical scale—the increase in difficulty is not "incremental."

A vertical scale cannot be established for other STAAR assessments because these assessments are not administered in consecutive years. A vertical scale for the STAAR EOC science and social studies assessments is not planned because there is not a mandated course sequence for students and the content for some subject areas is not necessarily related (i.e., biology differs from chemistry, which differs from physics). There currently are no plans to develop a vertical scale for the STAAR EOC English (English I/II/III) and Algebra (Algebra I/II) assessments. It may be feasible to do so after further consideration of the impact on score reporting and the implementation of the cumulative score requirement for high school graduation. If implemented, however, the vertical scales for STAAR EOC English and Algebra would be developed independently from those in grades 3–8 reading and mathematics due to the content shift in those areas from middle school to high school.

The STAAR vertical scale for grades 3–8 in reading and mathematics will be developed using data from a study that will be conducted as part of the first administration in 2012 and will be reported in fall 2012 along with other assessment results.

Horizontal Scale

A horizontal scale will be established for all other STAAR assessments. A horizontal scale for STAAR will be used for grades 5 and 8 science, grades 4 and 7 writing, grade 8 social studies, and EOC assessments. A horizontal scale allows for direct comparisons of performance across different test administrations within a grade and subject area but not across grades or subjects. A horizontal scale is appropriate for all subjects and does not require substantial overlap in content across grades. A horizontal scale allows comparisons of student scores to the performance standards at each grade level, but does not indicate student academic progress across grade levels.

Unique STAAR EOC Attributes for Scoring and Reporting

English I, II, and III

The English content area under the STAAR EOC assessment program consists of three assessments: English I, English II, and English III. Each English EOC assessment has two components: writing and reading. The writing component consists of multiple-choice items and essays, and the reading component consists of multiple-choice and short-answer items. Because of the length of these tests and the desire to embed field-test items to eliminate stand-alone field testing, each of the English EOC assessments will be administered over two days. The current plan is for students to complete the writing section on the first day and the reading section on the second day. The test design for English I, II, and III will allow for the reading and writing components to be calibrated, equated, and scaled separately so that the score on the reading and writing components can be reported separately. This allows a student to retake only the portion of the English EOC assessment on which he or she did not meet the minimum score requirements. All other EOC assessments will be administered on one day only during a scheduled assessment window.

Measuring College and Career Readiness

College- and career-readiness standards have been incorporated into the TEKS that will be assessed by STAAR. HB 3 states that the college- and career-readiness skills assessed on the Algebra II and English III assessments will be used in determining the level of performance necessary to indicate college readiness. Note that this indicator of readiness will be only one piece of information used in making readiness determinations. See Chapter 3 for more information about measuring college and career readiness.

In addition, HB 3 mandates that TEA conduct research studies to evaluate the correlation between performance on appropriate science and social studies EOC assessments and college/career readiness. If the commissioner of education, in collaboration with the commissioner of higher education, determines that the research studies substantiate an empirical relationship between a certain level of performance by

students on specific science and/or social studies EOC assessments and college and career readiness, the commissioners may establish college- and career-readiness performance standards for science and/or social studies EOC assessments as soon as possible. The research studies examining the extension of the concept of college and career readiness to science and/or social studies EOC assessments will be completed by December 1, 2012, when a report is due to the state legislature.

STAAR EOC Online Versus Paper Comparability

STAAR EOC assessments will be offered both online and on paper. Online and paper versions of the test forms are built to the same blueprints and specifications. To evaluate statistical comparability between online and paper tests, individual items are studied to determine whether they perform differently depending on the administration mode. The plan for comparability studies is specified in the following table.

Data Collection for Comparability Evaluation

Spring 2009	Spring 2010	Spring 2011
World Geography field test	Geometry	Algebra I
	Algebra II field test	Physics
	Biology	U.S. History
	Chemistry	World History field test
	English I field test	English II field test
		English III field test

If tests are found to be comparable, online and paper statistics and test scores can be used interchangeably. In this case, inferences drawn from students testing online are the same as for students testing on paper. If tests are not found to be comparable, estimated differences in test scores due to administration mode can be estimated. These estimated differences can inform what adjustment, if any, should be made to the online or paper score conversion tables. Once appropriate adjustments are made, inferences drawn from students testing online are the same as for students testing on paper.

TAKS vs. STAAR: A Comparison of Assessment Attributes

The summary table below compares the current TAKS assessment program with the new STAAR assessment program. The STAAR program will differ significantly from the TAKS program in a number of ways, including the curriculum assessed, how the passing standards will be set, how test items will be field tested, the number of days that will be devoted to testing, and the methods used to equate the difficulty of the tests from year to year.

A Comparison of Assessment Attributes Texas Assessment of Knowledge and Skills (TAKS) to State of Texas Assessment of Academic Readiness (STAAR)

Assessment Attributes	TAKS Assessment Program	STAAR Assessment Program
Assessed Curriculum	During initial TAKS development, Texas Essential Knowledge and Skills (TEKS) student expectations to be assessed were determined by Texas educators. Test objectives that matched the student expectations were developed. Blueprints for each assessment—the number of items per objective and on the overall test—were developed, with test lengths ranging from 30–60 items. At grades 3–8, content areas assess grade-specific content, with the exception of science at grades 5 and 8, which assess multiple grades of science curriculum. At grades 9–11, grade-level assessments assess content from multiple courses.	Educator committees identify which TEKS cannot be assessed on a paper/pencil assessment, which TEKS should be emphasized because they are necessary both for success in the current subject/grade or course and for preparedness in the next subject/grade or course, and which TEKS are considered supporting and should be assessed but receive less emphasis. New test blueprints will emphasize the assessment of the curriculum standards that best prepare students for the next grade or course. The assessments will encompass only the curriculum for that grade or course, with the exception of science at grades 5 and 8. The science assessments at these two grades will emphasize the 5th and 8th grade curriculum standards that best prepare students for the next grade or course; in addition, these assessments will include curriculum standards from two lower grades (i.e., grades 3 and 4 or grades 6 and 7) that support students' success on future science assessments.
Rigor of Assessment	The item-development process has been consistently followed once item-writer guidelines were developed in 2001. Performance standards were recommended by standard-setting committees and approved by the SBOE in November 2002. Because performance standards have remained consistent since the first operational administration in 2003 and after the phase-in of standards, students have "outgrown" the assessments. Measuring students' growth within the "Commended" performance category is difficult because too few items are rigorous enough to reflect this performance category and many students "top out" on the assessments.	Assessments will increase in length at most grades and subjects. Overall test difficulty will be increased by including more rigorous items. The rigor of items will be increased by assessing skills at a greater depth and level of cognitive complexity. In this way, the tests will be better able to measure the growth of higher-achieving students. In science and mathematics, the number of open-ended (griddable) items on most tests will increase to allow students more opportunity to derive an answer independently. Students will be required to respond to two writing tasks (including personal narrative, literary, expository, persuasive, and analytic) rather than one task. Performance standards will be set using empirical data gathered from studies that link performance year to year from grades 3–8 to high school and from specific courses to college readiness. Empirical studies will be conducted comparing students' performance on the new assessments with nationally norm-referenced assessments. Performance standards will be reviewed at least once every three years and, if necessary, adjusted to ensure that the assessments maintain a high level of rigor. Performance standards will be set so that they require a higher level of student performance than is required on the current TAKS assessments.
Field-Testing Process	From 2003–2007, stand-alone field testing for grades 4 and 7 writing, grade 9 reading, grade 10 and exit level English language arts, (ELA), and grade 5 Spanish reading and mathematics occurred annually; however, in 2008, stand-alone field testing moved to every other year. For all other subject areas, field-test items have been embedded in operational assessments.	For grade 7 writing and for each end-of-course assessment, there is a one-time only stand-alone field test. Once STAAR assessments are operational, all field testing will be embedded, with the exception of grade 4 writing, which will require an abbreviated stand-alone field test every three years.

Assessment Attributes	TAKS Assessment Program	STAAR Assessment Program
Performance Standards	Performance standards were set separately for each grade and subject. Performance standards were set based on the examination of test content.	Performance standards will be set as an aligned system across grades and courses within a content area (from grades 3–8 through high school). Performance standards will be set based on data from empirical studies of other state, national, and international assessments as well as on the examination of test content.
Test Administration Procedures	All assessments are currently administered within a one-day time frame. Online testing is offered for exit-level retests only.	Grades 4 and 7 writing as well as English I, II, and III will be administered over two days to assess writing more comprehensively and allow for the inclusion of embedded field-test items. End-of-course assessments will be made available on paper and online.
Measures of Student Progress	Measures of student progress for the growth model were developed and implemented after the TAKS program was established. Growth measures are projections to the "Met Standard" performance level at the next high-stakes grade (5, 8, and 11). Growth measures provide information about whether students are on track to meet the passing standard in the next high-stakes grade.	Measures of student progress for the growth model will be developed and implemented as STAAR assessments are developed and implemented. Progress measures will be based on the new, more rigorous standards for STAAR assessments. Progress measures will be phased in over several years as data for the new program become available. Progress measures may provide an early-warning indicator for students that are not on track to meet the passing standard, may not be successful in the next grade or course, may not be ready for advanced courses in mathematics and English in high school, or may not be college or career ready in mathematics and English.
Number of Testing Days	*Total – 19 (27 with SSI retesting)* Grade 3 – reading and mathematics (2 days) Grade 4 – reading, mathematics, and writing (3 days) Grade 4 – writing field test (1 day) Grade 5 – reading, mathematics, and science (3 days; up to 7 days for SSI retesting) Grade 6 – reading and mathematics (2 days) Grade 7 – reading, mathematics, and writing (3 days) Grade 7 – writing field test (1 day) Grade 8 – reading, mathematics, science, and social studies (4 days; up to 8 days for SSI retesting) *Total – 13 (25 with Exit Level retesting)* Grade 9 – reading and mathematics (2 days) Grade 9 – reading field test (1 day) Grade 10 – ELA, mathematics, science, and social studies (4 days) Grade 10 – ELA field test (1 day) Grade 11 (Exit Level) – ELA, mathematics, science, and social studies (4 days; up to 16 days for retesting) Exit Level – ELA field test (1 day)	*Total – 19 (27 with SSI retesting)* Grade 3 – reading and mathematics (2 days) Grade 4 – reading, mathematics, and writing (4 days; writing now a 2-day administration) Grade 5 – reading, mathematics, and science (3 days; up to 7 days for SSI retesting) Grade 6 – reading and mathematics (2 days) Grade 7 – reading, mathematics, and writing (4 days; writing now a 2-day administration) Grade 8 – reading, mathematics, science, and social studies (4 days; up to 8 days for SSI retesting) *Total – 15 (45 with retesting)* English I (2 days) English II (2 days) English III (2 days) Algebra I (1 day) Geometry (1 day) Algebra II (1 day) World History (1 day) World Geography (1 day) U.S. History (1 day) Biology (1 day) Chemistry (1 day) Physics (1 day) 2 additional testing opportunities per year

Assessment Attributes	TAKS Assessment Program	STAAR Assessment Program
Assessments for English Language Learners (ELLS) at Grades 3–8 and High School	The majority of ELLs participate in TAKS in English (grades 3 through exit level) or TAKS in Spanish (grades 3–5) Grades 3–10: Eligible recent immigrant ELLs may, however, be granted a limited English proficiency (LEP) exemption for up to three years under state law. Students exempt under Texas law are required to test in federally mandated grades and subjects (grades 3–8 and 10 mathematics and reading; grades 5, 8, and 10 science). In these grades and subjects, they take TAKS with linguistic accommodations, as permitted by federal regulations. In other grades and subjects, they do not take TAKS while exempt under state law. Exit level: ALL ELLs must pass exit level TAKS to meet graduation requirements. There are no exemptions. Exit level testing, however, may be postponed during an eligible immigrant ELL's first 12 months in U.S. schools.	The vast majority of ELLs will participate in STAAR in English (grades 3 through high school) or STAAR in Spanish (grades 3–5). State exemption policies and linguistically accommodated assessment methods for immigrant ELLs are under review, with the goal of expanding valid and reliable linguistic accommodation methods and including more recent immigrant ELLs in the state assessment system.
Assessments for Students Receiving Special Education Services	Assessments for students receiving special education services—an accommodated form, a modified assessment, and an alternate assessment—were developed. All these assessments are aligned to the TEKS as well as to the TAKS objectives, but the test blueprints for the modified and alternate assessments differ from TAKS. Separate performance standards were set on the modified and alternate assessments. However, performance standards for the accommodated form are the same as TAKS. These assessments were developed after the TAKS program was well established.	For students receiving special education services, modified and alternate versions of the STAAR assessments will be developed, although it is possible that all 12 end-of-course assessments may not be developed due to the nature of the coursework actually taken by students who are eligible to participate in these assessments. The modified and alternate assessments will be aligned to the TEKS as well as to the reporting categories for STAAR, although the test blueprints for these assessments will differ from the general assessments. Separate performance standards will be set on the modified and alternate versions of STAAR. The alternate assessments will be developed at the same time and in coordination with STAAR development activities, providing for greater continuity and alignment between the general and alternate assessments.
Equating	The TAKS program has used both pre- and post-equating models to verify that the assessments maintain the same level of difficulty from year to year. Post-equating has been done using the base test items as the linking items to maintain difficulty from year to year.	TEA is considering using both pre- and post-equating models to verify that the STAAR assessments maintain the same level of difficulty from year to year. A new post-equating design that uses embedded linking items on a subset of test forms is being considered for assessments at grades 3–8 as well as for English I, II, and III.