

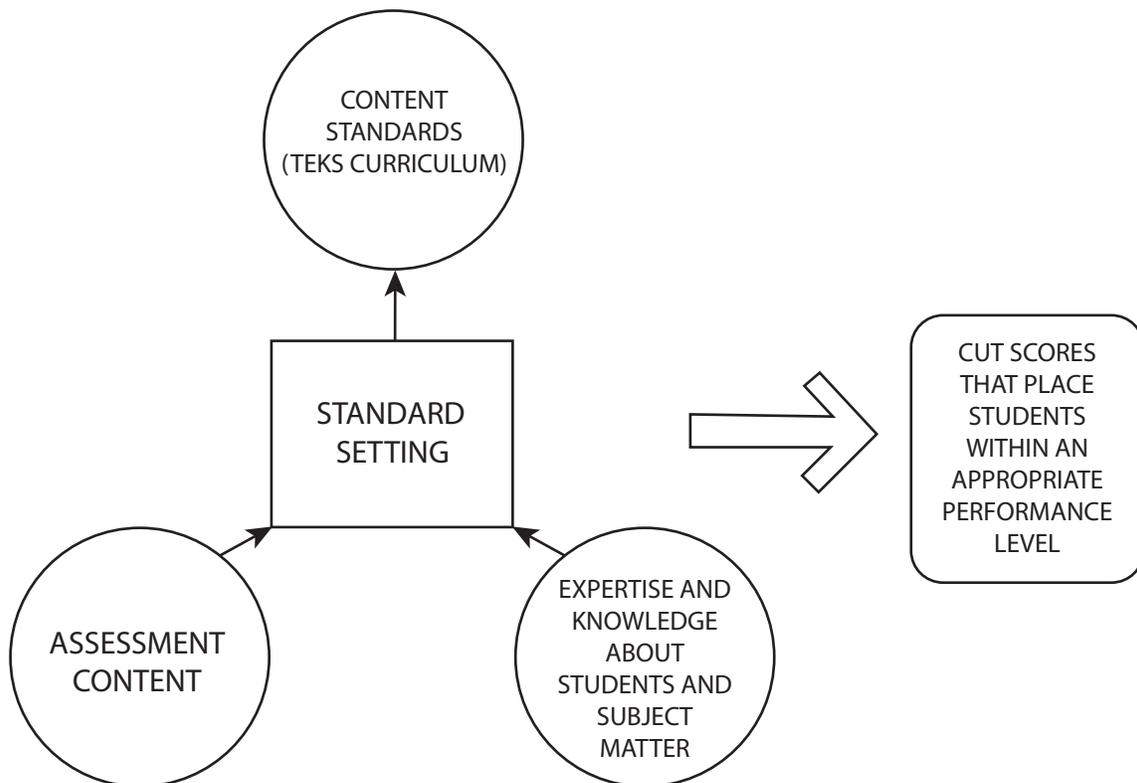
Chapter 14: Standards

Overview

One of the most critical aspects of a statewide testing program is the establishment of performance levels that provide an additional frame of reference for interpreting the test scores. Performance standards relate test performance directly to the Student Expectations expressed in the state curriculum, the Texas Essential Knowledge and Skills (TEKS), in terms of what students are expected to learn by the completion of each grade level.

Once an assessment is given, students, parents, educators, administrators, and policymakers want to know, in clear language, how students performed on that assessment. Performance standards provide that clarity and facilitate the translation of student scores into decisions about individual students. As depicted in Figure 6, standard setting is the process of linking content standards (the TEKS), educator knowledge about what students should know and be able to do, and the assessment to set cut scores that most appropriately place students within one of several performance levels.

Figure 6. Standard-Setting Process



Because setting performance standards involves professional judgment, an organized process is used that incorporates all the critical elements of the standard-setting process so that high-quality judgments are made on behalf of Texas students.

The Standard-Setting Process

The general process of setting performance standards includes the following steps:

1. Texas educators use the TEKS curriculum to establish content standards.
2. Policymakers set general performance level descriptors based on the content standards. In Texas, for example, the State Board of Education (SBOE) determined that there would be three descriptors for Texas Assessment of Knowledge and Skills (TAKS)—Commended Performance, Met Standard, and Did Not Meet Standard.
3. Standard-setting panelists take the general descriptors and make them specific by elaborating on what they mean for students in a particular content area and grade level. They describe what students should know and be able to do in each of the performance categories on the assessment.
4. Using the content-specific performance level descriptors, standard-setting panelists complete the standard-setting process and produce a recommendation for cut scores, or score differentiations on the assessment, that indicate how the general performance-level descriptors map onto the assessment.
5. The performance standards recommendation from the standard-setting committee is then submitted to either the Commissioner of Education or the SBOE depending on which is responsible for determining the final performance standards.

The state curriculum—the TEKS—is used to establish test objectives that guide the development of test items and the assessment content. Performance standards therefore are based on the content standards for the assessment. The product of the standard-setting process is a set of cut scores that classify students into an appropriate performance level.

Historical Perspective on Assessment and Standard Setting in Texas

The following section provides a brief history of testing in Texas with an emphasis on the standard-setting process.

Texas Assessment of Basic Skills (TABS)

The first formal assessment to explicitly link student assessment results to curriculum statewide was the TABS. In 1979 the Texas Legislature enacted a bill amending the Texas Education Code (TEC) to require the Texas Education Agency (TEA) to adopt and administer a series of criterion-referenced assessments designed to assess basic skills competencies in mathematics, reading, and writing for students in grades 3, 5, and 9. Since no mandated statewide curriculum existed at that time, the learning objectives for TABS were developed by TEA and represented a small portion of the many skills students were expected to learn in Texas public schools. The TABS objectives were reviewed and revised by educator committees, including teachers, principals, and curriculum and psychometric specialists, to ensure that they were appropriate before they were submitted to the SBOE. Careful consideration was given to the mandated requirement for TABS to assess “minimum basic skills competencies.”

The TEC was amended by the Texas Legislature in 1983 to require grade 9 students failing the TABS test to retake the exam each year thereafter. Although individual students were not held accountable for failing the exam (that is, they were not denied diplomas), this legislative action increased the incentive for campuses to provide remedial support for students falling below minimum expectations on TABS. In addition, for the first time in Texas, the TABS test results for each campus and district were released to the public. The publication of campus and district results regarding specific performance relative to the learning objectives represented the beginning of high stakes accountability for large-scale assessment in Texas.

Texas Educational Assessment of Minimum Skills (TEAMS)

In 1984 the Texas Legislature amended the wording of the TEC from “basic skills competencies” to “minimum basic skills.” This change was seen as a mandate to increase the rigor of the assessments and to add individual student sanctions related to performance on the test. TEAMS replaced TABS as the new state-mandated, criterion-referenced achievement test in the subjects of reading, mathematics, and writing. TEAMS was administered to students in grades 1, 3, 5, 7, 9, and 11, with the grade 11 test being an exit level assessment. The introduction of TEAMS reflected an increased emphasis on educational reform as required by changes in law. In addition to moving the assessment from the measurement of “basic skills competencies” to the measurement of “minimum basic skills,” these changes included an increase in the number of grades tested and mandatory remediation and retesting of students failing the exit level assessment. For the first time in Texas, individual students were denied diplomas based, in part, on their performance on the TEAMS test.

Texas Assessment of Academic Skills (TAAS)

In the late 1980s the SBOE directed TEA to make a number of changes to the assessment program. These changes included an expansion of the content being measured and a greater emphasis on the assessment of problem-solving skills, with more content directly

linked to the core curriculum, the TEKS. The new assessment program, the TAAS, was implemented in 1990. The TAAS testing program reflected the desires of both the SBOE and the Commissioner of Education for Texas students to attain higher levels of academic achievement. The primary purpose of assessment in Texas had evolved from the collection of school-level information (TABS) to assessment of curriculum-specific minimum skills (TEAMS) to school accountability for student performance (TAAS).

Beginning in fall 1990, TAAS was administered to students in grades 3, 5, 7, 9, and 11, with grade 11 serving as the exit level test. In addition to meeting attendance and course requirements, students were required to meet a passing standard established by the SBOE to receive a high school diploma. The subject areas tested included reading, mathematics, and writing.

TAAS testing was moved to the spring in the 1993–1994 school year, and testing was expanded to include grades 3–8 in reading and mathematics and grades 4 and 8 in writing. The exit level test was moved to grade 10 to provide more time for remediation and retesting. After 1994 the TAAS tests remained similar in scope and sequence. However, additional assessments were added to the Texas Student Assessment Program through legislative changes or SBOE directives. For example, tests in social studies and science were added to TAAS at grade 8, and Spanish-version tests were added in grades 3 through 6. End-of-course assessments in Algebra I, biology, English II, and U.S. history were also added. Students were given the option of passing the Algebra I, English II, and either the biology or U.S. history tests as an alternative to the TAAS graduation test requirements.

Texas Assessment of Knowledge and Skills (TAKS)

As mandated by the 76th Texas Legislature in 1999, TEA developed a new assessment program to replace TAAS. This new assessment program, the TAKS, was first administered to students in the 2002–2003 school year. TEC Section 39.024(a) charged the SBOE with establishing the required performance standards on the new TAKS. Because the standard-setting process was critically important, TEA convened a national Technical Advisory Committee of educational testing experts to provide SBOE members with the expertise necessary to perform this function. The Technical Advisory Committee provided guidance in developing a standard-setting process whereby TEA could provide information and training to the SBOE in support of the task of setting the TAKS performance standards in November 2002. For further information about the guidance provided in selecting a standard-setting process and drafting the standard-setting plan, see Chapter 11 of the *2003–2004 Technical Digest*.

The goal of the standard-setting activities was to provide the SBOE with the necessary information to facilitate setting performance standards that would meet the needs of public education in Texas. The procedures, recommendations, and suggestions provided by the national Technical Advisory Committee plan were designed to ensure that those decisions were based on careful and deliberate expert analysis. Every effort was made to ensure that the TAKS performance standards were legally defensible (see Chapter 11

of the *2003–2004 Technical Digest*). Ultimately the SBOE had the authority and responsibility for establishing performance standards on the TAKS assessments, as required by statute.

Coordination with the State Board of Education

On October 31, 2001, the SBOE Committee on Instruction held a special work session, during which a subgroup of national measurement experts from the TAC presented information critical to an understanding of the standard-setting process. The Committee on Instruction then considered components to be included in the standard-setting plan. Information collected at this October 31 work session was presented to the SBOE at its November meeting for further input and review and then to the TAC as a basis for the development of the standard-setting plan. The final standard-setting plan was presented to the SBOE for approval in January 2002. Detailed components of the standard-setting plan can be found in Chapter 11 of the *2003–2004 Technical Digest*.

General Performance Level Descriptors

The SBOE decided to create three performance categories for TAKS. The task of selecting and describing the various student performance “labels” to report student performance on TAKS was assigned to a specially convened advisory panel. This advisory panel was composed of 19 individuals representing state professional, educational, and public policy organizations. Its primary function was to select appropriate descriptive labels for the TAKS student performance categories and to set forth broad, generic descriptors for each label. After discussion and selection of three preferred labels, panelists generated various phrases and descriptors for inclusion with the labels. The labels and descriptors chosen were intended to be equally applicable to all grade levels and content areas covered by TAKS. The labels recommended by the advisory panel, with the generic descriptors added by TEA based on the recommendations of the panel, were as follows:

- **Commended Performance**—This category represents high academic achievement. Students in this category performed at a level that was considerably above the state passing standard. Students demonstrated a thorough understanding of the knowledge and skills measured at this grade.
- **Met Standard**—This category represents satisfactory academic achievement. Students in this category performed at a level that was at or somewhat above the state passing standard. Students demonstrated a sufficient understanding of the knowledge and skills measured at this grade.
- **Did Not Meet Standard**—This category represents unsatisfactory academic achievement. Students in this category performed at a level that was below the state passing standard. Students demonstrated an insufficient understanding of the knowledge and skills measured at this grade.

These labels and a description of this process were presented to the SBOE for its review at the July 2002 meeting. See Chapter 11 of the *2003–2004 Technical Digest* for further elaboration on the process used to select the recommended labels.

Content-Specific Performance Level Descriptors

When each of the standard-setting advisory panels met, a significant activity that took place prior to the panels making any recommendations was for them to extend the general descriptors listed above to the specific grade level and content area on which they were to recommend standards. This activity permitted panelists to conceptualize more clearly the various labels in terms of specific grade- and content-based academic behaviors. Information about the formalization of these descriptors is discussed on the following pages in the Formalization of Content-Specific Performance Level Descriptors section of this chapter.

Standard-Setting Advisory and Review Panels

The bulk of the effort that led to the TAKS performance standards as recommended to the SBOE was conducted by 25 performance standards advisory and review panels. A total of 15 English TAKS advisory panels, four TAKS review panels, and six Spanish TAKS advisory panels were convened between August 26, 2002, and October 23, 2002. Details on the process followed by the advisory and review panels can be found in Chapter 11 of the *2003–2004 Technical Digest*.

Item-Mapping Methodology

A variety of research-based methods are available for setting standards on academic subject tests such as the TAKS. As recommended by the standard-setting contractor and approved by the TAC and SBOE, the general methodology used with panels of Texas educators and stakeholders as they generated recommended standards for the SBOE was a modified item-mapping method. The item-mapping method (Lewis, Green, Mitzel, Baum, & Patz, 1998) was chosen for several reasons. First, it is a method for high-stakes educational assessment standard setting that has been used in many other statewide testing programs. Second, it is a procedure well suited for assessments that contain both selected-response (multiple-choice) and constructed-response (open-ended) items. While only some of the TAKS tests contain constructed-response items, it was considered desirable to use a single methodology for all tests.

For the TAKS standard setting, panelists were trained to examine items that had been ordered in a review booklet from least to most difficult. Panelists progressed through the booklet until they reached the point at which they believed a threshold student who minimally Met Standard should be able to answer the item correctly. That is, panelists placed a cut point at the item at which they judged that a student who just barely performed at the Met Standard level would answer correctly. The median cut point across panelists was used as the recommended standard. A similar process was then followed for the Commended Performance cut. For additional information and data concerning methodological procedures and session outcomes, see the “Setting Standards on the TAKS Tests: A Modified Item Mapping Procedure” (2002) report in the 2007 Texas Education Agency Technical Report Series which can be found at <http://www.tea.state.tx.us/student.assessment/resources/techdigest/>.

Performance Standards Phase-In Recommendation

The primary purpose of performance standards is to communicate to students, schools, parents, and the public the expected level of achievement. As such, these standards serve as goals to which all students can aspire. It would be of little use to establish performance standards that would result in 100 percent classification of all students into the highest category, for example. Thus, if the performance standards established by the SBOE are goals for students to aspire to, presumably not all students will reach the standard. Furthermore, most standards are set on new assessments aligned to a rigorous curriculum. As such, the national TAC of 2002 recommended that the SBOE consider establishing a graduated (that is, phased-in) passing standards system that moves from a lower standard the first year to a more rigorous standard at some point in the future. It should be noted that the advisory panels also recommended that the SBOE establish a phase-in period for the standards—that is, adopt the recommended standards but provide a period of some small number of years during which these standards would become effective. TEA provided several options to the SBOE regarding how rapidly the panel-recommended performance standards might become effective.

Performance Standards Established

At a November 2002 meeting, members of TEA staff, along with select members from the TAC, provided the SBOE with an overview of the TAKS performance standards advisory process and the recommendations offered. The SBOE spent time reviewing all proposed TAKS test materials and carefully considered the results of the impact data resulting from the fall 2001 study, the spring 2002 field test, information from the standard-setting advisory panels, and data that link TAKS field-test data to TAAS.

On November 15, 2002, the SBOE voted to adopt the TAKS performance standards originally proposed by the approximately 350 educators and citizens who served on TAKS standard-setting advisory panels. Additionally, the SBOE adopted a transition plan whereby the recommended Met Standard cut score would be phased in over a three-year period, expressed in terms of standard error of measurement (SEM) units below the panel-recommended performance standards. The transition plan, thus, uses the SEM to phase in the panels' recommended standards over time. It was noted that even at the initial phase-in level standard, TAKS was a more challenging test for students than its predecessor, TAAS.

The transition plan adopted by the SBOE is depicted in Figure 7.

Figure 7. SBOE-Adopted Transition Plan for TAKS Performance Standards

Grade	2003	2004	2005	2006	2007	2008	2009
3	2 SEM	1 SEM	PR	PR	PR	PR	PR
4	2 SEM	1 SEM	PR	PR	PR	PR	PR
5	2 SEM	1 SEM	PR	PR	PR	PR	PR
6	2 SEM	1 SEM	PR	PR	PR	PR	PR
7	2 SEM	1 SEM	PR	PR	PR	PR	PR
8	2 SEM	1 SEM	PR	PR	PR	PR	PR
9	2 SEM	1 SEM	PR	PR	PR	PR	PR
10	2 SEM	1 SEM	PR	PR	PR	PR	PR
11*	TAAS	2 SEM	1 SEM	PR	PR	PR	PR
12*	TAAS	TAAS	2 SEM	1 SEM	PR	PR	PR

SEM = Standard Error of Measurement

PR = Panel recommendation

*For grade 11 (exit level), the TAKS exit level standard in place at the time a student begins grade 10 is the standard that will be maintained throughout the student's high school career. Therefore, grade 11 students tested in spring 2005 must meet the 1 SEM TAKS passing standard because that was the standard in place in 2004 when they were in grade 10.

For the spring 2003 TAKS test administration, students had to achieve a score equivalent to, or exceeding, a cut score that was two SEM below the panel-recommended performance standard for that grade and content area. With the exception of the exit level tests, the performance standard cut score in spring 2004 shifted to one SEM below the panel-recommended standard (see Table 7). For the exit level tests, students retained the performance-level standard that was in effect when they took the TAKS grade 10 test as the standard that they must meet to satisfy the TAKS graduation requirements. By spring 2006, all the TAKS tests were at the performance-level cut scores recommended by the standard-setting advisory panels.

Under the transition plan adopted, grade 3 students, for example, were required to correctly answer 20 of 36 questions on the spring 2003 English reading test to meet the passing standard. Because future exams might vary slightly in difficulty (possibly affecting raw score cuts), it was understood by the SBOE that these recommended cut scores would be translated into underlying scale score cuts with the first operational administration of the TAKS (when the scale score system was established). The underlying scale cut scores associated with the performance standards as adopted by the SBOE on the spring 2003 test became the thresholds defining the TAKS performance standards for all future test administrations. For all subject-area/grade-level tests at the panel-recommended standard, a scale score of 2100 is required for Met Standard, and a scale score of 2400 is required to achieve Commended Performance.

Table 7. TAKS Scale Scores Required to Achieve Met Standard and Commended Performance Levels

Grade	Subject	2 SEM Scale Score	1 SEM Scale Score	Panel Recommended** Scale Score	Commended Performance Scale Score
3	Mathematics	1986	2041	2100	2400
4	Mathematics	1997	2047	2100	2400
5	Mathematics	1978	2037	2100	2400
6	Mathematics	1994	2046	2100	2400
7	Mathematics	2023	2061	2100	2400
8	Mathematics	2015	2057	2100	2400
9	Mathematics	2000	2050	2100	2400
10	Mathematics	2007	2054	2100	2400
11	Mathematics	2015	2058	2100	2400
3	Reading	2029	2064	2100	2400
4	Reading	2039	2069	2100	2400
5	Reading	2025	2062	2100	2400
6	Reading	1989	2044	2100	2400
7	Reading	2009	2053	2100	2400
8	Reading	2006	2051	2100	2400
9	Reading	2021	2059	2100	2400
4	Writing*	2023	2060	2100	2400
7	Writing*	2035	2067	2100	2400
10	ELA*	2045	2071	2100	2400
11	ELA*	2045	2072	2100	2400
8	Social Studies	2016	2058	2100	2400
10	Social Studies	2020	2060	2100	2400
11	Social Studies	2033	2067	2100	2400
5	Science	1940	2016	2100	2400
8	Science**	1984	2041	2100	2400
10	Science	1993	2046	2100	2400
11	Science	2035	2068	2100	2400
Spanish-Version Tests					
3	Mathematics	1986	2042	2100	2400
4	Mathematics	1954	2025	2100	2400
5	Mathematics	1957	2026	2100	2400
6	Mathematics	1979	2038	2100	2400
3	Reading	2019	2059	2100	2400
4	Reading	1988	2043	2100	2400
5	Reading	1970	2034	2100	2400
6	Reading	1976	2038	2100	2400
4	Writing*	2000	2049	2100	2400
5	Science	1941	2017	2100	2400

*An essay score of 2 or higher is required for Met Standard on the grades 4 and 7 writing tests and the grades 10 and 11 English language arts tests. An essay score of 3 or higher is required for Commended Performance on the grades 4 and 7 writing tests.

**Students taking the assessment in 2007–2008 needed to obtain the panel recommended scale score to meet the standard.

Following the initial administration of the TAKS tests in spring 2003, the performance standard cuts as established by the SBOE in terms of total score (raw score) points were translated to the underlying scale score system so that these performance standards would be generalizable across future test forms and administrations. Table 7 shows the TAKS performance-level cuts expressed in terms of scale scores for the three phase-in levels of the Met Standard performance level as well as for the Commended Performance level. These tables were reviewed by the SBOE at its meetings on July 11, 2003, and September 12, 2003. At its September meeting, the SBOE affirmed, by a vote of 8–7, the phase-in plan for TAKS performance standards was adopted on November 15, 2002.

It should also be noted that the TEC requires establishment of a standard that reflects a student's readiness to enroll in an institution of higher learning to fulfill the higher education component of the TAKS exit level assessment. This performance standard is separate from the TAKS performance standards and was established by the Texas Higher Education Coordinating Board in spring 2004. Students who receive a scale score of 2200 or above on the mathematics test and/or a 2200 or above on the English language arts (ELA) test with a composition subsection score of a least 3 will be exempt from the requirements of the Texas Success Initiative in the relevant subject area. See the "TAKS Higher Education Readiness Component (HERC) Contrasting Groups Study" report in the 2007 Texas Education Agency Technical Report Series for more information about the setting of this readiness standard. The technical report series can be found at <http://www.tea.state.tx.us/student.assessment/resources/techdigest/>.

Grade 8 Science Standards

On June 29–30, 2005, TEA and Pearson convened a group of 22 panelists to recommend Commended Performance and Met Standard standards for the TAKS grade 8 science assessment. The recommendations from the standard-setting panel, their associated standard errors, and the statewide impact data by group for the final panel-recommended cut points are summarized in Appendix 15 of the *2004–2005 Technical Digest*. The SBOE reviewed these recommendations at the July 14–15, 2005, meeting. The SBOE reviewed the recommendations again at the October 18, 2005, meeting and set the final standards for this assessment, voting to phase in the panel recommendations at the Met Standard level.

Formalization of Content-Specific Performance Level Descriptors

Overview

As part of the requirements of the federal No Child Left Behind Act (NCLB), all states must develop descriptions of the knowledge and skills necessary for their students to achieve specific performance levels on their state assessments. To comply with this requirement, TEA invited educators from across the state to convene in Austin in January 2006 to formalize performance level descriptors for the TAKS. These educators may have served either on TAKS external review committees in 2005 or participated in the original 2002 standard-setting meetings for TAKS, and represented the state geographically, regionally, and ethnically.

The committee members were initially divided into working groups for single grades or grade-level clusters. Each group brainstormed competency-based, content-specific phrases that describe the classroom behaviors that students exhibit at each of the TAKS performance levels: Commended Performance, Met Standard, and Did Not Meet Standard. Next, representatives from these small groups met as one committee to vertically align the descriptors across the grade levels in each content area. The descriptors then were reviewed by internal and external content experts. This process resulted in a set of comprehensive descriptors that Texas districts, campuses, and teachers can use to guide them in teaching and assessing students.

Districts, campuses, and teachers can use the descriptors to

- review district content standards and classroom unit and lesson plans to ensure that they promote the behaviors at Commended Performance and Met Standard,
- develop new units and lessons and their accompanying assessments, integrating and reinforcing knowledge and skills across the content areas when applicable,
- guide planning for professional development on instructional strategies that encourage the behaviors at Commended Performance and Met Standard, and
- serve as the focus of professional dialogue among teachers about teaching specific knowledge and skills (for example, at grade level, department, or staff meetings).

In addition, the performance level descriptors can enhance parents' understanding of their child's academic strengths and weaknesses and can help the community at large better understand state test scores and the level of performance required of students to achieve Commended Performance and Met Standard on TAKS.

The TAKS performance level descriptors for Met Standard exemplify what Texas wants all students to minimally achieve at each grade level for each foundation subject area. Those for Commended Performance are the goals for the majority of our students. Acquiring this set of knowledge and skills will provide the foundation for students' success in the future.

The performance level descriptors can be found online at <http://www.tea.state.tx.us/student.assessment/resources/perlevel/>.

Algebra I End-of-Course (EOC) Assessment

On November 3–4, 2005, TEA and Pearson convened a group of 22 panelists to recommend performance standards—specifically, Commended Performance and Met Standard—for the Algebra I EOC assessment, which was administered in the current online format for the first time in spring 2005. The two-day standard-setting event included sessions in which panelists (1) reviewed the test items online, (2) clarified what the performance level descriptors meant for the Algebra I course, and (3) applied an item-mapping procedure to set recommended cut points. During the item-mapping procedure, panelists reviewed the content assessed by the test items, engaged in table and whole-group discussions, and considered the impact on students when making their cut-score recommendations.

The final recommendations from the standard-setting panel, their associated standard errors, and impact data are presented in the “Algebra I EOC Standard Setting Report” report in the 2007 Texas Education Agency Technical Report Series, which can be found at <http://www.tea.state.tx.us/student.assessment/resources/techdigest/>. Panel recommendations were presented to TEA and approved for use with the fall 2005 administration.

Texas Assessment of Knowledge and Skills–Modified (TAKS–M)

Standard Setting for TAKS–M occurred during the first two weeks of August 2008. Standards were set for the following grades and subjects: reading grades 3–8, ELA grade 10; mathematics grades 3–8 and 10; and science grades 5, 8, and 10. The first operational administration for these assessments was in spring 2008. Standards for the remaining tests (reading grade 9, ELA grade 11, writing grades 4 and 7, mathematics grades 9 and 11, social studies grades 5, 8, 10, and 11, and science grade 11) will be set in the fall 2009.

Immediately following the spring 2008 TAKS–M administration, raw scores were reported back to the students taking the assessment. Once standards have been set, a second wave of reports will be sent out with the standards applied to student scores.

The approach for setting standards will be an item-mapping method (Lewis, Mitzel, & Green, 1996), the same as the approach originally used to set standards for TAKS. Panelists will receive an ordered item booklet containing items ordered by difficulty from easiest to most difficult. Panelists will look through the booklet and mark the location in the booklet where they believe that minimally proficient (or advanced) students will be able to correctly answer the items appearing before that location. Panelists will make three rounds of judgments and have the opportunity to work independently and in table groups as part of the standard-setting process.

Texas Assessment of Knowledge and Skills–Alternate (TAKS–Alt)

During September 2007, the TEA Student Assessment Division and Pearson convened ten standard-setting panels to recommend performance standards—specifically, Commended Performance and Met Standard—for TAKS–Alt for the following grades and subjects: reading grades 3–9; ELA grades 10 and 11; writing grades 4 and 7; mathematics grades 3–11; science grades 8, 10, and 11; and social studies grades 8, 10, and 11. The ten, three-day, standard-setting meetings included sessions in which panelists (1) were provided with an overview of the assessment, (2) became familiar with the TAKS–Alt performance level descriptors, and (3) applied a “body of work” procedure (Kingston, Kahl, Sweeney, & Bay, 2001) to set recommended cut points. During the body of work procedure, panelists reviewed the content of student folders, engaged in small group and whole group discussions, and considered the impact on students when making their cut-score recommendations.

Once standards were recommended for all grades in a subject area, select panelists from each standard-setting panel formed a smoothing committee to make cross-grade comparisons. This committee evaluated whether the recommended standards were reasonable given the standards set in the other grades for the same subject area. To make this evaluation, the smoothing committee reviewed TAKS–Alt impact data for all grades, the TAKS–Alt scoring rubric and performance level descriptors, and TAKS impact data across several years to demonstrate the expected shift in student performance over time. Based on their evaluation, this committee recommended adjustments to the results in one or more grades. In this manner, the standard-setting results from the independent grade panels were aligned or “smoothed.”

The final recommendations from the standard-setting panels, their associated standard deviations, impact data, and smoothing results are presented in the “TAKS–Alt Standard-Setting” report in the 2007 TEA Technical Reports Series which can be found at <http://www.tea.state.tx.us/student.assessment/resources/techdigest/>. TEA applied the results of the standard setting and smoothing before presenting their final cut score recommendations to the Commissioner of Education. The standards were approved for use with the spring 2008 administration.

TELPAS Reading for Grades 2–12

The second version of the Texas English Language Proficiency Assessment System (TELPAS) reading test was administered for the first time in spring 2008. In implementing this version of the assessment, proficiency level cut points needed to be appropriately transitioned from the former assessment, the TELPAS Reading Proficiency Tests in English (RPTE). To accomplish this, TEA implemented a plan that had been reviewed by the Texas Technical Advisory Committee (TTAC) and three national experts in language acquisition. The plan included a process for locating cut points that best categorize students into the proficiency levels that have fundamentally guided the test item development and the construction of the proficiency level mini-tests that comprise the assessment. The process was designed to identify the cut points that best classify students in accordance with the proficiency level definitions that underlie the test and included two phases:

Phase I. From June to August 2008, an internal work group reviewed item-level data, test-level data, and impact data. The group consisted of approximately 12 staff members from TEA, Questar Assessment, and Pearson. The work group had expertise in TAKS reading assessments, psychometrics, and second language acquisition in reading and item development. The group also had expertise in TAKS reading assessments and psychometrics. After a series of eight meetings, the group recommended a set of cut points for each grade/grade cluster assessment.

Phase II. In August 2008, TEA convened an external review group of state educators who had experience reviewing field-test items from the former and new program and using RPTE test results to guide instruction. The group reviewed the cut-point recommendations from Phase I, the test forms on which the cut points were based, and impact data. Impact data was shown for a range of cuts in the vicinity of the recommended cut points. The group then recommended one set of cut points for each grade/grade cluster assessment.

The final cut-point recommendations from Phase II was presented to the Commissioner of Education for approval in August 2008. Complete results of the proficiency level setting will be available in the “TELPAS Reading Proficiency Level Setting” report in the 2007 TEA Technical Reports Series which can be found at <http://www.tea.state.tx.us/student.assessment/resources/techdigest/>.

Ensuring the Quality of the Standard Setting

To help ensure that all data used in the standard-setting meetings are accurate and that the standard-setting methods employed are those considered best practice in the field, several quality assurance steps are followed. To promote accuracy in the data, all data analyses are replicated. Estimations of the item parameters, students’ abilities, and impact data are calculated independently by two psychometricians. Results from the two

analyses are then compared, and discrepant results are discussed and resolved. Only after agreement is reached are data used in the standard-setting meeting.

As an additional check that the general planned standard-setting methods are considered best practice in the field, the standard-setting plans used with Texas assessments have been presented to the TTAC. This committee includes national standard-setting experts. The committee reviews the purpose for the meeting, the methods, the panelist selection process, and the evaluation methods.

