

11/8/2002

Standard Setting Implementation Plan

for the

Texas Assessment of Knowledge and Skills
(TAKS)

Drafted by

Beck Evaluation & Testing Associates, Inc (BETA)
& NCS Pearson, in cooperation with Texas Education Agency
July 2002

TAKS TEST OVERVIEW BY GRADE & CONTENT AREA

A total of 36 separate tests will make up the new TAKS system – 26 in English and 10 in Spanish. Tests that will eventually become required for promotion or graduation (i.e., “high-stakes” tests) are underlined in the table below. Panels will be convened to recommend standards for each of the 36 tests. Most panels will recommend standards for more than one, but no more than two, tests. These pairings are shown in the chart on page 4. However, some panels will recommend standards for only one test; these panels are indicated in boldface below.

The process used by both types of panels to set the performance standards for the TAKS tests will be essentially the same. The time needed by the two types of panels will obviously differ. Those panels considering two tests (15 in number) will require three days, while those considering only a single subject and grade (six such panels) require only two days. (See “Session Outlines for Two- and Three-Day Sessions) Note that both types of panels will make judgment-based recommendations. The “mechanics,” or specific activities involved in the panels’ work, will be comparable; the number of tests they must consider, however, will differ.

A third type of panel will be convened after the work of the first two types is complete. These six “review” panels will meet for one day each for the purpose of supporting TEA staff as they review the recommendations of the 21 standard setting panels to minimize any irregularities and maximize the consistency in the recommended standards to be sent forward to the State Board of Education in November 2002. There will be one review panel for each content area. These panels will be composed primarily, if not exclusively, of panelists from the regular standard setting panels.

TABLE

Gr. 3	Gr. 4	Gr. 5	Gr. 6	Gr. 7	Gr. 8	Gr. 9	Gr. 10	Gr. 11
<u>TAKS ENGLISH</u>								
<u>Reading</u>	Reading	<u>Reading</u>	Reading	Reading	<u>Reading</u>	Reading		
	Writing			Writing			ELA	<u>ELA</u>
Math	Math	<u>Math</u>	Math	Math	<u>Math</u>	Math	Math	<u>Math</u>
		Science					Science	<u>Science</u>
					Soc.St.		Soc.St.	<u>Soc.St.</u>

TAKS SPANISH

<u>Reading</u>	Reading	<u>Reading</u>	Reading
	Writing		
Math	Math	<u>Math</u>	Math
		Science	

METHODOLOGY & PROCESS ISSUES

Number of Cuts

Key Decisions

1. TEA Student Assessment determined, with advice from the National Technical Advisory Committee, that 2 cuts should be set for the TAKS tests.
2. With the possible exception of the Grade 11 tests, the same number of cuts should be used for all tests.
3. The recommended number of cuts should be shared and discussed with the appropriate subject area specialists in Curriculum and Accountability.
4. **The recommendation is being taken to the SBOE for their review** in July 2002 for use by Standard Setting Panels in August, September, and October 2002.

Labels and Descriptors

Standard Setting Advisory Panel

1. Composition: 19 high-level individuals representing organizations such as the TX Business and Education Coalition, the NAACP, the TX Association of School Administrators, the TX Council of Teachers of Mathematics, and the TX Public Policy Foundation.
2. Convened on May 23, 2002, in Austin, TX, to identify **labels** for the categories with broad, generic descriptors.
3. This session lasted for 6 hours and was facilitated by M. Beck and M. Caswell of BETA, Inc. Presented the panel with a broad range of choices for labels. After discussion and selection of preferred labels, began discussion of descriptors and generated phrases and ideas believed to be key for inclusion in descriptors. Recognized that descriptors will be generic ones that apply to all grades and content areas. Labels that were adopted by TEA:
 - a. Commended Performance – Performance Well Above the Standard
 - b. Met the Standard – Performance Above the Standard
 - c. Did Not Meet the Standard – Performance Below the Standard

Subsequent to Initial Standard Setting Meeting

4. TEA – working to define descriptors in terms of the TEKS student competencies. At this stage, descriptors will be “generic”—crossing grades and content areas.
5. Standard Setting Panels – Each individual panel uses the information from the Advisory Panel and then further defines the labels in terms of concrete student behaviors for their session’s grade level and content area.

Using Labels from Other Sources (e.g., NAEP)

1. TAKS cut-score labels have now been developed specifically for TX.
2. Adoption of NAEP labels and descriptors in particular has been avoided. Standard Setting Advisory Panel members agreed, after discussion, that adoption of NAEP labels could create confusion and interpretive difficulties on the part of the public and parents.

STANDARD-SETTING MEETINGS

Overall Methodology – Item Mapping. See below for specifics.

Why Not 36 Panels?

The primary reason for limiting the number of “regular” panels is to minimize the irregularities that are almost guaranteed to occur if 36 separate such panels are used and to maximize the consistency of the standards set across grades and subjects in order to more accurately track the educational progress of Texas students. Consider Reading as an example. We propose consolidating the typical item-mapping panel sessions for collecting recommended cuts for reading for Grades 3 & 4, 5 & 6, 7 & 8, and 10 & 11 ELA. This will allow educators and others on these panels to benchmark their cuts in a consistent manner as they consider a second test in the same content area with related objectives and student expectations. Because panels will have received thorough training in the item-mapping procedure at the beginning of their three-day session, they will be readily able to repeat the procedure with the second test in this three-day period. Grade 9 reading will require a separate two-day standard setting session panel.

Number of Sessions Required

The following chart indicates the fifteen pairs of tests that will require three-day standard setting sessions. These were selected by TEA as logical pairs of grades within a common subject, where educators could most readily consider student performance on tests at two different grades and most probably arrive at cut score recommendations within the required time frame.

Grade Combinations

<i>3/4</i>	<i>4/7</i>	<i>5/6</i>	<i>7/8</i>	<i>10/11</i>
Reading		Reading	Reading	ELA
	Writing			
Math		Math	Math	Math
				Science
				Social Studies

Spanish TAKS

3/4	5/6
Reading	Reading
Math	Math

There are six subjects at various grades, as shown in the Table on page 2 in boldface, that are “stand alone” subjects to be addressed by single panels in two day sessions by six panels.

Taken together with the 15 three day panels described above, BETA will convene **21 regular** standard setting panels to provide TEA with the requisite cut scores for the new TAKS tests. In addition, review panels will be convened to assist TEA with the work necessary for finalizing these recommendations prior to their presentation to the State Board of Education in November 2002. BETA and NCS staff will assist with these latter panels as well.

Number of Panelists per Panel – target of 20-25 panelists per session

Rationale

1. Fewer than about 20 results in insufficient representation and diversity, and more than 25 is unwieldy and slows down the process.
2. Panel size of 20-25 seems advisable for either the “regular” or “consolidated” panels.

Composition

1. Mix of active educators (teachers, principals, administrators, curriculum specialists, counselors, testing coordinators) and all other involved stakeholders, with the former making up about 2/3 of the panel and teachers making up most of this group.
2. Appropriate ethnic, gender, and demographic (service center and super region), representation are also important.
3. Involve only educators who **haven’t** participated in TAKS item or data review sessions.
4. Participants are required to be present for entire two or three days of sessions; confidentiality of items and data is paramount.

Schedule

1. The schedule has been set, with the first sessions beginning the end of August. See the “Approved Schedule for Texas TAKS Standard-Setting Sessions.” In most instances, 2-3 sessions are being held concurrently using only experienced BETA facilitators.

2. The item mapping books will be assembled for panelists' use based on the proposed final TAKS tests, put together by TEA and NCS during the early summer from the spring 2002 field test materials.
3. Sessions begin August 26, 2002, and run through October 18, 2002, if necessary.
4. BETA estimates needing approximately one week to complete follow-up data analysis necessary to provide the final set of 36 Panel recommendations to TEA.

Location

1. Downtown Austin. Ample working space for panelists and materials has been given consideration as arrangements were made at the Embassy Suites Hotel.
2. TEA, NCS, and BETA have worked and will continue to work together to arrange the site, personnel, and logistics using the highly successful model developed to arrange and hold item and data review meetings.

Session Facilitation

Moderators

1. Experienced BETA staff will facilitate all Panel sessions. All panelists will use the same script for sessions.
2. BETA moderators: M. Beck, M. Caswell, J. Moyer, S. Potter, W. Brand. All of these BETA staff members have conducted multiple high-stakes sessions of this type. J. Slinde will assist as available with data entry and interpretation during the sessions; NCS Pearson psychometricians will provide this support when Dr. Slinde is unavailable.

Roles & Responsibilities During Panel Sessions

TEA

1. Arrange for a videotaped welcome by a SBOE member and/or the Commissioner to standardize the message received by all Panels about the importance of the Standard Setting activity.
2. Observe the standard setting activity.
3. Act as a resource in each session as requested to answer questions about the TEKS, policy issues, or the testing program. Either Assessment or Curriculum/Instruction staff members are appropriate as TEA desires.
4. Create, distribute, and collect evaluation form as well as provide onsite evaluation of facilitators.

NCS

1. Prepare all materials for Panelists and distribute these at the facility.
2. Logistics at facility – before and during the sessions.
3. Data analysis during and following each session. “Data analysis” demands are minimal in the item mapping procedures, but we will want to track cuts across sessions and across rounds within session.

TAC Members – observe, as available or interested

MECHANICAL ISSUES

Time Needed Per Panel

For a single test, 2 full days per panel (see attached “Session Outline, TAKS Two- Day Standard-Setting Session”). For the consolidated-panels, 3 full days will be used for their consideration and judgments about tests at multiple grades (see attached “Session Outline, TAKS Three- Day Standard-Setting Session”).

Number of Rounds of Work

Three are proposed (see outline referenced above). BETA’s experience is that very few panelists require more than 2 rounds to be comfortable with their judgments, and only rarely would a panel’s overall recommendation shift significantly after Round 3. (As noted above, 2 rounds of recommendations are proposed for the consolidated-panels.)

Data to be Shared with Panels

1. Panelists’ individual cuts will be displayed graphically (and anonymously) after each round of ratings.
2. Spring, ’02 field test data will be provided and explained before Round 2 cuts are made by Panelists.
3. Impact data (likely percents of students in each category) to be provided before Round 3 cuts are made by Panelists.

Variances from the typical CTB “Bookmark” Process

1. “Subgroup” discussions used in typical Bookmark process are not used in the BETA item mapping process. We strongly prefer to have all discussions held by the complete group – much like a jury approach.
2. Estimation of various error types are needed, but “subgroups” aren’t required to obtain these.

Large-Group Opening Session

The number of panelists attending concurrent sessions and facility constraints will likely preclude holding a large-group opening session. Hence, we recommend for a video taped welcome that could be seen by all panelists in individual session rooms. (Since a TEA representative will be present in each session, a personal TEA welcome and “thank you” can be given if desired.)

Standard Setting Professional Development

BETA conducted a 1/2-day "prototype" or "mini" standard-setting session for TEA and NCS staff members on June 12, 2002. An abbreviated version of the actual sessions was used so that TEA and NCS staff members could become familiar with the procedures and technical aspects of standard setting and have the opportunity to have their questions answered before the TAKS sessions get underway.

Although unlikely at this point, if BETA conducts standard-setting sessions for another state before August, it still seems desirable for 1 or 2 selected TEA staff members to observe.