

**Student Assessment Advisory Committee Report
Recommendations Concerning Field-Testing Procedures for the
Texas Student Assessment Program for 2008 and Beyond
April 2007**

Introduction

As state and federal testing requirements have continued to grow, the Texas Education Agency (TEA) has become increasingly aware of the need to keep district field testing to a minimum while maintaining the high quality of the student assessment program. As part of this effort, TEA has taken some steps to minimize the burden and reduce the amount of time that classroom instruction is interrupted for field-testing purposes. However, districts continue to express their concerns about the amount of field testing that is required each year.

To address district concerns and to explore further potential reductions in field testing, the commissioner convened an advisory committee comprised of superintendents, district testing coordinators, central office staff in leadership positions, and representatives from education service centers. This committee was called on to provide the agency with recommendations addressing major policy and design issues related to field testing for the 2007–2008 school year and beyond under current federal and state statute. The recommendations the committee made were constrained by the following requirements.

1. The validity and integrity of the state testing program are not compromised.
2. The state tests are accurate and reliable measures of student achievement.
3. The state testing program is legally defensible.
4. The state testing program is in compliance with agency interpretation of current statute.

Overview

Field testing is conducted in high-stakes testing programs as an important part of the item-review process. Field testing facilitates the development of tests that are fair for all student groups, of high quality, legally defensible, and would withstand rigorous scrutiny when evaluated relative to professional standards. Field testing is, in a sense, a “test of the test”; it is the only way to determine whether the items that have been developed are valid and reliable measures of what students know and can do.

For the Texas Assessment Program, field testing in the 2006-2007 school year is being conducted for the Texas Assessment of Knowledge and Skills (TAKS), Spanish TAKS, Reading Proficiency Tests in English (RPTE II), TAKS-Alternate (TAKS-Alt), and End-of-Course (EOC) tests. With promotion and graduation requirements for individual students as well as accountability ratings for schools and districts linked to performance on state assessments, TEA has to ensure that state tests are fair, appropriate, and defensible measures of student achievement. To do that, the items developed each year must be field-tested.

Two needs must be balanced when field testing is conducted: minimizing burden on students and districts and administering tests that meet recommended industry test development and construction standards. Practically, it is important to minimize field testing so that as little instructional time as possible is compromised. Statistically, it is important that Texas collect

enough field-test data from a representative sample of students to have reliable information to evaluate the fairness of questions and build high-quality tests.

Background Information on Field Testing in Texas

Why field-test?

- Texas uses field-test data to evaluate questions for bias and fairness.
- Texas field-tests enough questions each year so that a sufficient number of questions are available to
 - develop tests that meet content requirements
 - include questions that cover a range of difficulty on each test
 - create tests of equal difficulty from year to year
 - replace questions that are released
- Texas development and test construction processes are designed to meet professional testing standards recommended for high-stakes testing.

What are the different types of field testing?

TEA uses two approaches to administer field-test questions to samples of students: embedded questions and separate field-test forms.

Embedded Field Tests in Texas

TEA embeds the majority of items it field-tests each year.

- Embedded field-test questions are administered at the same time as “live” questions.
- Embedding field-test questions into the operational test is a process typically used with tests that have multiple-choice questions.
- Texas includes field-test questions—usually about 10 per form—on most multiple-choice tests.

Separate Field Tests

Separate field tests are used when test structure, small student populations, new tests, or method of test delivery preclude embedding field-test items.

- *Test Structure*
For some TAKS assessments—grade 4 writing (in English and in Spanish), grade 7 writing, grade 9 reading, and grade 10 and exit level English language arts—separate field testing is necessary due to the structure of the tests and the inclusion of performance tasks (a composition and/or short answer reading items) to which students must respond.

Embedding field-test items in the live test would put unrealistic demands on students and would likely be detrimental to student performance. At the fourth and seventh grade, it would require students to write on two prompts (the actual prompt and a field-test prompt) in one day. At the high school level, students would have to read two triplets (six literature pieces) instead of one (three literature pieces), answer twice as many short answer reading questions, and write on two prompts.

- *Small Student Populations*
When the numbers of students in a population are small, such as for Spanish TAKS, it may be necessary to field-test the entire tested population (referred to as census field testing) rather than a sample to develop a sufficient number of items to support test construction.
- *New Tests*
When TEA must institute new tests to satisfy state or federal requirements, separate field testing is needed initially. The data from this field test are used to review test questions for appropriateness and fairness, inform the standard-setting process, and build the initial operational test.
- *Method of Test Delivery*
When tests are administered in a new mode, such as online, separate field tests are often needed. The field-test data are used to evaluate whether test questions given on paper vary in difficulty when administered online.

How Are Field-Test Data Used to Evaluate Item Appropriateness and Fairness?

The analysis of field-test data is a critical aspect of the test development process. After items are field-tested on a large sample of ethnically diverse students from across the state, Pearson Education Measurement, the primary contractor for the testing program, conducts a variety of statistical analyses of the data based on student performance. Then data review committees composed of Texas educators as well as content specialists from TEA's curriculum and assessment divisions meet to review the statistics and field-tested items. These review committees are carefully selected to represent the state demographically with regard to ethnicity, gender, type and size of district, and geographical region. They also represent the various populations of students, including students with disabilities, English language learners, and at-risk students. The data review committees examine the field-test statistics for each item relative to that item's content. If the educators on the committee believe that the item is unfair or inappropriate in any way, that item is not eligible for inclusion on a future test. However, if they believe that an item is a fair and accurate measure of student learning, the item becomes eligible for inclusion.

Field testing ensures that the items eligible for assessment represent a range of difficulty. For each administration, tests are constructed with a combination of items that equal the approximate level of difficulty of preceding years. This method certifies that students are tested at the same relative level of difficulty each year.

Overall, field testing is a central part of the rigorous test-development process, which has been instituted to ensure that every item on a state-developed test is aligned with the curriculum, grade-level appropriate, unbiased, and clearly written and that the testing program is, in fact, based upon reasonable expectations for Texas students. In this way, valid and reliable tests can be constructed each year.

Summary of Efforts to Reduce the Amount of Stand-Alone Field Testing

As required testing demands have increased over the last several years, TEA has taken some steps to minimize the burden and reduce the amount of time in which classroom instruction is interrupted for field-testing purposes. For more specific information about TEA's efforts to reduce districts' field-testing burden, please see the report entitled *Annual Spring Field-Test Sampling for TAKS Mixed-Format Assessments: Alternative Sampling Guidelines to Reduce Campus Participation Burden* at the following link:

http://www.tea.state.tx.us/student.assessment/resources/field_test/fieldtest_summary.pdf.

Field-Testing Options for Modifying Stand-Alone Field Tests

There are a number of different field-testing options that can be pursued starting with the 2007–2008 school year. Options are discussed in more detail below, but in general, they include maintaining the current model while further reducing the field-test burden through additional changes to sampling procedures, alternating field-test years, and moving the existing field-test window to another time during the year. Other factors, such as test design and pending legislation related to time lines for releasing state assessments, could also influence options for field testing.

Option 1: Maintain the current model

One option is to maintain the current field-testing model and continue to examine ways to reduce the field-test burden. This would require that sampling procedures in each subsequent year be examined so that TEA could minimize the amount of field testing to the extent possible.

Pros:

- Stability—districts are familiar with this model and have developed local calendars to accommodate these field-testing dates.

Cons:

- Continued concerns—district concerns about the amount of field testing under this model would likely continue, as this option may not result in substantial changes in the amount and types of field testing currently being done.

Option 2: Alternate field-testing years

Under this option, stand-alone field tests could be conducted at intervals, such as every other year. One consideration under this option is that if field testing were moved to alternating years, it would likely require field tests to be conducted with more students in a grade and subject. In other words, additional students would be needed to field-test enough items to support test construction in the interim years. Legislative actions, such as new testing requirements, would require exceptions to alternate-year field testing.

Pros:

- Decreased testing—districts would be assured that they would be required to field-test only every other year.

Cons:

- Larger field-test samples—in the field-test years, the number of students required to participate in field testing might increase.

Option 3: Conduct stand-alone field testing at a different time of year

Listed below are implications associated with moving stand-alone field-testing to a window after the primary administration or to the fall.

Administer field tests after the primary April TAKS administration

- The testing window for a May field test would need to occur prior to school ending to allow for enough time for return of materials.
- Materials would most likely arrive in districts at the same time as the primary April administration materials are being packaged and shipped for return.
- Grade 5 mathematics (and, in 2008, grade 8 mathematics) retests would likely occur during a May field-test window.
- EOC assessments may occur at the same time as a May field-test window.
- The field tests may overlap with other testing programs such as Advanced Placement.

Administer some or all field tests in the fall

- Field testing in the fall would mean that items would be tried out more than six months after the operational tests have been given.
- If some or all field tests are moved to the fall, students in the subsequent grades would take the field-test items associated with the previous grade. For example, students in grade 5 would take the TAKS grade 4 writing field test, as these would be the students who have had grade 4 instruction.
- Giving students the field test in fall at the subsequent grade may cause “opportunity to learn” issues. Students might perform better in the fall because they would have more instruction than students responding to field-test questions in January. On the other hand, students might perform worse because of a summer learning loss.

Pros:

- Less concentrated testing—moving some or all field testing to the late spring would reduce the amount of testing close to the primary administration in February
- Less spring testing—moving some or all field tests to the fall would place these tests outside of the busy spring semester of testing

Cons:

- Potential instability of data—data may be unstable given the lack of student motivation to perform well after testing has occurred and the increase in time between the operational test and the field test.
- More complicated logistics—managing test materials for multiple administrations at the same time would place an additional burden on testing coordinators and could cause errors in the handling of these materials.

Other Considerations

In addition to the options discussed above, other considerations could impact the way in which field testing is conducted:

Change in Release Policy

Under the current law, tests are released every other year (e.g., 2006 and 2008). If the law were changed so that only a percentage of questions were released or that retests and the TAKS grade 10 ELA make-up test were not released, then less field testing would be required. Further field-testing relief could be realized if tests were released less often, for example, every third year.

Move to EOC Assessments at High School

If changes in law require that Texas move to EOC assessments at high school, there would initially be an increased field-testing burden during the transition from TAKS to EOC assessments. For instance, each new EOC assessment (multiple-choice as well as those tests with essay/short answer responses) would require an initial stand-alone field test. At the same time, TAKS assessments will continue to require field testing until the transition to EOC assessments is complete.

Online Test Delivery

Administering TAKS and RPTE II on computer requires additional field testing during the transition to the online mode. Halting or slowing the move to online testing for some or all of these tests would reduce the field-test burden. This change would result in use of paper tests only, would not make use of previously collected online data, and might require repeated field testing if Texas decided to move to online testing in the future.

Summary of Options

Options	Description	Advantages	Disadvantages
Option 1: Maintain the current model	Maintain the current field-testing model and continue to examine ways to reduce the field-test burden through sampling procedures.	<ul style="list-style-type: none"> Stability 	<ul style="list-style-type: none"> Continued concerns about burden May not be possible for Spanish TAKS
Option 2: Alternate field-testing years	Stand-alone field testing could be conducted at intervals, such as every other year or every third year. This option would perhaps be more viable for Spanish TAKS if tests were released less frequently than now.	<ul style="list-style-type: none"> Decreased testing 	<ul style="list-style-type: none"> Larger field-test samples
Option 3: Conduct stand-alone field-tests at a different time of year	Move stand-alone field testing to a window after the primary administration or to the fall.	<ul style="list-style-type: none"> Less concentrated testing 	<ul style="list-style-type: none"> Potential instability of data More complicated logistics

Student Assessment Advisory Committee Recommendations to Reduce Field Testing Required for the Texas Assessment Program

After the Student Assessment Advisory Committee reviewed and discussed the background information on field testing in Texas, they offered recommendations for ways to reduce field-test burden. These recommendations are organized in three general categories and are summarized below.

When to Field-Test

Alternate-year stand-alone field tests

The primary recommendation from the group was to field-test in alternate years. The group acknowledged that moving to an alternate-year model would likely mean that, during the years in which field testing would be conducted, there would be

- an increased number of students selected for participation
- field testing in more grades at a campus
- more campuses selected for participation within a district

Different models for operationalizing this recommendation would need to be examined. One model would be to conduct stand-alone field testing for all school districts one year and then offer relief for all school districts in a subsequent year. A second model would be to select some districts for stand-alone field testing in one year and a different set of school districts for stand-alone field testing in a subsequent year. However, under the second model it may not be possible to ensure the representativeness of the field test data if certain large school districts such as Dallas and Houston are excluded from the samples.

Spring field tests

The general consensus of the group was that they would not support moving stand-alone field testing to the spring after the primary TAKS administrations.

Fall/winter field tests

The general consensus of the group was to request the development of a draft implementation schedule for moving the majority of stand-alone field testing to the fall. The group characterized the fall as a time during the school year that could more easily accommodate field testing than the spring.

Who to Field-Test

Sample size and sampling procedures

The group recommended that sampling procedures be re-examined to ensure that adequate but not excessive numbers of students are included in field-testing activities. Issues to be considered include

- determining the minimum number of students and test forms required to maintain reliability and validity as well as legal defensibility of mandated assessments
- examining the feasibility of testing samples of English language learners (ELLs) rather than conducting census field testing for the stand-alone Spanish TAKS field tests
- conducting an independent verification of sampling procedures to verify that minimum yet adequate numbers of students are included in field-test samples and to ensure that certain districts and campuses are not overselected for field testing

Campuses implementing state/federal interventions

The group recommended that consideration be given to exclude from field testing alternative education campuses or campuses that are implementing state or federally required interventions due to academic performance concerns. Issues of particular importance to these campuses include

- the impact of field testing on scheduling
- the number of instructional days that are used for field testing
- the implications of being designated as a non-participating campus for field testing

April TAKS retest

The general consensus of the group was to remove field-test items from the TAKS exit-level retests whenever feasible. To accomplish this TEA will explore

- eliminating embedded field-test items for seniors
- removing field-test items from April TAKS retests

How to Field-Test

Test design

The group recommended that consideration be given to changing the fundamental design of the state assessments, particularly the English language arts and writing assessments. Issues to be considered include

- examining the implications of moving to timed tests as the untimed nature of the current assessments directly contributes to logistical burden on districts regarding scheduling
- exploring test design options that would minimize the need for stand-alone field testing, such as separating out reading and writing tests
- evaluating whether the benefits of preparing students for college-level expectations regarding written compositions outweighs the field-testing relief provided by using multiple short writing samples instead

Online testing

The consensus of the group was that online testing represents a positive direction for the state assessment program, although the aggressive implementation timeline should be reevaluated. Issues to be considered include

- reviewing the amount of time required to reconfigure machines for each online assessment
- standardizing communication to districts and making administrative materials available much sooner for online testing
- evaluating the impact of online field testing on students and the additional administrative burden on professional staff
- considering timed tests to allow for testing more students on each computer, e.g., a three-hour test would allow a campus to schedule multiple administrations in one day

Incentives/feedback

The consensus of the group was that participation in field testing could be increased if some form of incentives were offered. The group recommended that TEA

- explore the feasibility of reporting summary performance information from stand-alone field tests
- investigate using incentives for stand-alone field-test participation

Student Assessment Advisory Committee Recommendations for Topics to Address at Future Meetings

After the Student Assessment Advisory Committee discussed field testing in Texas, they offered recommendations for topics for future meetings. These recommendations include

- online testing and training
- assessments for special populations (English language learners and students with disabilities)
- end-of-course issues
- security issues, including future statistical procedures for identifying potential instances of security violations
- review plans for implementation of recommendations made about field testing

Commissioner's Preliminary Response to the Field-test Recommendations of the Student Assessment Advisory Committee by Program and Year

This preliminary response addresses possible field-testing changes to the student assessment program in Texas based on the recommendations of the Student Assessment Advisory Committee noted on the previous pages. The response is organized by program for the 2007–2008, 2008–2009, and 2009–2010 school years. Overall field-test recommendations are listed before the preliminary recommendations proposed for each individual assessment program.

For 2007–2008

1. Overall
 - Notify districts prior to the start of the school year of their field-test participation
 - Conduct an independent verification of sample size and sampling procedures
 - Consider removing field-test items from the April exit level administration of mathematics, science, and social studies for senior retesters
 - Consider the feasibility of reporting summary field-test results to those districts that participated in the stand-alone field test
 - Consider excluding alternative education campuses and campuses that are implementing state or federally required interventions due to academic performance concerns
2. TAKS Reading and ELA (grades 9, 10, 11)
 - Develop a plan to transition to an alternate year, stand-alone field-test schedule
 - Continue efforts to further reduce field-test sample sizes by
 1. considering the number of students needed per form
 2. considering sampling procedures
 - Explore the feasibility of reducing the number of field-test forms
3. TAKS Writing (grades 4 and 7) and Spanish TAKS Writing (grade 4)
 - Develop a plan to transition to an alternate year, stand-alone field-test schedule
 - Explore the feasibility of moving the field test to a fall time frame
 - Evaluate the number of composition prompts that need to be field-tested
4. TAKS Spanish Reading and Mathematics (grades 5 and 6)
 - Examine the possibility of reducing the size of the field-test sample
 - Explore the feasibility of moving the field test to a fall time frame
 - Explore the feasibility of transitioning to an alternate year, stand-alone field test schedule
5. RPTE II
 - Conduct the RPTE II field test online as planned (since this is the last year needed for a stand-alone field test)
6. EOC
 - Consider field testing only one subject per campus

For 2008–2009

1. Overall
 - Notify districts prior to the start of the school year of their field-test participation
 - Continue to conduct an independent verification of sample size and sampling procedures
2. TAKS Reading and ELA (grade 9, 10, 11)
 - No stand-alone field tests
3. TAKS Writing (grades 4 and 7) and Spanish TAKS Writing (grade 4)*
 - No stand-alone field tests
4. Spanish TAKS Reading and Mathematics (grades 5 and 6)*
 - No stand-alone field tests for grade 5

* If feasible, 2007–2008 will be the first of the alternate years not to have these stand-alone field tests. If so, stand-alone field tests would occur in 2008–2009 and would not occur in 2009–2010.

For 2009–2010

1. Overall
 - Notify districts prior to the start of the school year of their field-test participation
 - Continue to conduct an independent verification of sample size and sampling procedures
2. TAKS Reading and ELA (grade 9, 10, 11)
 - Conduct stand-alone field tests
3. TAKS Writing (grades 4 and 7) and Spanish TAKS Writing (grade 4)
 - Conduct stand-alone field tests
4. Spanish TAKS Reading and Mathematics (grades 5 and 6)
 - Conduct stand-alone field tests

**Student Assessment Advisory Committee
March 29, 2007, Meeting**

Attendees or Substitutes Provided

- David Anthony, Superintendent, Cypress-Fairbanks ISD
- Sara Arispe, Director of Assessment and Accountability, Fort Worth ISD
- Mac Bernd, Superintendent, Arlington ISD
- Johnny Brown, Superintendent, Port Arthur ISD*
(Paula Beaty, District Testing Coordinator)
- Cathy Bryce, Superintendent, Highland Park ISD*
(Gena Gardiner, Assistant Superintendent)
- H.D. Chambers, Superintendent, Stafford ISD
- John Folks, Superintendent, Northside ISD*
(Sandra Poth, Director of Testing and Evaluation)
- Hector Gonzales, Superintendent, Brownsville ISD
- Keith Haffey, Executive Director of Accountability and Assessment, Spring Branch ISD
- Roland Hernandez, Superintendent, Waco ISD
- Michael Hinojosa, Superintendent, Dallas ISD*
(Cecilia Oakley, Associate Superintendent)
- Rick Howard, Superintendent, Comanche ISD
- Pam Leftwich, Student Assessment Coordinator, Lubbock ISD
- Patti Lyle, District Testing Coordinator, Bay City ISD
- Richard Middleton, Superintendent, North East ISD
- Thomas Negri, District Testing Coordinator, Alief ISD
- Kaye Orr, Coordinator of Accountability, Region 18 ESC
- Connie Player, Leadership and School Improvement, Region 7 ESC
- Judy Pollan, Superintendent, Pittsburg ISD
- Pat Schmitz, Director of Testing, San Antonio ISD
- Rod Schroeder, Superintendent, Amarillo ISD*
(Charles Ritchie, District Testing Coordinator)
- Ann Smisko, Assistant Superintendent for Curriculum and Instruction, Austin ISD
- Sue Thompson, Director of Testing, Ysleta ISD
- Martha Vannoy, Director of Planning, Research, and Evaluation, Garland ISD
- Leland Williams, Superintendent, Dickinson ISD

Non-attendees

- David Baum, Executive Director, Abilene ISD
- D. Scott Eliff, Superintendent, Corpus Christi ISD
- Roberto Fernandez, Superintendent, San Felipe-Del Rio ISD
- Dawson Orr, Superintendent, Wichita Falls ISD
- Don Stockton, Superintendent, Conroe ISD

* Did not attend – name of substitute in parentheses