

# CHAPTER 14: VALIDITY

Validity is a process of collecting evidence to support inferences made from the scoring results of an assessment. In the case of the TAKS and SDAA exams, test results are used to make inferences about students' knowledge and understanding of the TEKS. For RPTE, test results provide a measure of progress, indicating annually where each LEP student is on a continuum of English language development designed for second language learners.

## CONTENT VALIDITY

Standards-referenced assessments, such as the TAKS, SDAA, and RPTE tests, are based on an extensive definition of the content they assess. Test validity is therefore content based and tied directly to the statewide curriculum. In order to ensure the highest level of content validity, the process of aligning TAKS, SDAA, and RPTE to the curriculum was carefully approached and included numerous committees of Texas educators.

When TAKS was designed as the standards-referenced assessment for the TEKS, advisory committees consisting of educators from school districts across the state were formed for each subject area at each grade level. Teachers, test development specialists, and TEA staff members worked together in these committees to develop test objectives, instructional targets, specifications, and test item types. Similarly, committees met during the 2000–2001 school year to identify TEKS statements eligible for testing. In addition, committees met starting in 2001–2002 to review and edit TAKS items for content and bias and to review data from field testing. A similar process was conducted for both SDAA and RPTE when they were developed.

## RELATION TO THE STATEWIDE CURRICULUM

The item writers as well as the reviewers for each stage of development verify the alignment of test items with the objectives to ensure that the items measure appropriate content. The sequential stages of item development and item review provide many opportunities for Texas educators to offer suggestions for improving or eliminating items and to offer insights into the interpretation of the statewide curriculum. The nature and specificity of these various review procedures provide additional strong evidence for the content validity of the TAKS, SDAA, and RPTE tests.

## EDUCATOR INPUT

Not only do Texas educators provide valued input on the content and the match between the items and the statewide curriculum, but many current and former Texas educators and some educators from other states also work as independent contractors to write items specifically to measure the objectives. This provides for a system of checks and balances for item development and review that reduces single source bias. In other words, because test items are written by many different people with different backgrounds, it is less likely that items will suffer from a bias that might occur if items were written by a single author. The direct input from educators offers additional evidence regarding the content validity of constructed TAKS, SDAA, and RPTE tests.

## TEST DEVELOPER INPUT

The staff at TEA, as well as professional test developers from Harcourt Educational Measurement, Pearson Educational Measurement, and BETA, Inc., provide a wealth of test-building experience, including content-related expertise. Each internal review of an item by these experts increases the probability of the item being an accurate measure of the intended objective. Hence, these reviews are offered as additional evidence for the content validity of the TAKS, SDAA, and RPTE tests.

## CONSTRUCT VALIDITY

Content and construct validity are traditionally separate categories, but when discussing the validity of achievement tests, these distinctions become somewhat obscured. Content validity describes whether the test objectives adequately represent what students should be able to do and whether the items, which are based on test objectives, measure intended responses. Construct validity is the extent to which a test can be said to measure a theoretical construct or trait.

In the case of TAKS, SDAA, and RPTE tests, the construct tested is the academic content required by the statewide curriculum. With curriculum-based achievement tests, both types of validity are intertwined. The construct validity is grounded in the content validity of the test.

Further evidence of construct validity for TAKS was provided as part of the Higher Education Readiness Component of TAKS. TEA gathered performance data for exit level mathematics and ELA tests on a sample of college students at two- and four-year institutions throughout the state on which a contrasting groups study was performed. The percent of college students at the “Met Standard” and “Commended” performance on these tests was compared to that of high school students (see “Higher Education Readiness Component” on p. 12).

## CRITERION-RELATED VALIDITY

Criterion validity indicates the relationship between test performance and performance on some other measure. This other measure can be evaluated concurrently or at a future point in time and is then correlated with the test score. In this way, the test score is compared with a criterion that is thought to be a reasonable estimate of the same construct the original test purports to measure.

As part of the TAKS Higher Education Readiness Component, a concurrent validity study is planned for 2003 and 2004 to correlate performance on exit level TAKS with performance on national testing programs.

## VALIDITY EVIDENCE FOR SUBPOPULATIONS

The primary evidence for the validity of the TAKS test lies in the content being measured. Since the tests assess the Texas state-mandated curriculum, which is required to be taught to all students, the tests are not more or less valid for use with one subpopulation over another subpopulation. In other words, because the TAKS test is measuring what is required to be taught to all students and is given under the same standardized conditions to all students, the test has the same validity for all students. Correspondingly, the evidence for the validity of SDAA and RPTE also lies in the content of the state-mandated curriculum being measured. The format of the tests is designed to be appropriate for use with all SDAA and RPTE subpopulations.

Moreover, great care has been taken to ensure that the items that make up the TAKS, SDAA, and RPTE tests are fair and representative of the content domain as provided by the required curriculum and expressed in the measurement specifications and objectives. Additionally, much scrutiny is applied to the items and their possible impact on minority or subpopulation groups making up the population of the state of Texas. Every effort is made to eliminate items that may have ethnic or cultural biases. These procedures are described in Chapter 6, Annual Test Development Activities.