

# CHAPTER 15: VALIDITY

Validity is a process of collecting evidence to support inferences made from the scoring results of an assessment. In the case of TAKS and SDAA II, test results are used to make inferences about students' knowledge and understanding of the TEKS. RPTE assessment results provide both a measure of progress in English language acquisition and information about students' knowledge and understanding of the reading TEKS.

## Evidence Based on Test Content

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

When TAKS was designed as the standards-referenced assessment for the TEKS, advisory committees consisting of educators from 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 identify important-to-assess TEKS student expectations and to develop test objectives, item development guidelines, and test-item types. 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 II and RPTE when they were developed.

The development of the TAKS Equivalency Standard as part of the SDAA II-to-TAKS linking study provides additional validity evidence for SDAA II. The SDAA II testing program was designed to align with the TAKS and to cover the same range of the TEKS curriculum as the TAKS. The TAKS Equivalency Standard is the score on each SDAA II test given on grade level at which a student demonstrates the same level of performance as a student who met the passing standard on the corresponding TAKS grade-level test. Further information on the TAKS Equivalency Standard can be found in Chapter 12: Standards and in Appendix 22. The fact that the TAKS and SDAA II tests could be successfully linked provides evidence of the content validity of the SDAA II.

## 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 II, and RPTE assessments.

## **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 validity of constructed TAKS, SDAA II, and RPTE tests.

## **Test Developer Input**

The staff at TEA, as well as professional test developers from Educational Testing Service, Pearson Educational Measurement, and BETA, Inc., provide a wealth of test-building experience, including content 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 II, and RPTE tests.

## **Test Expert Input**

TEA, in conjunction with Pearson Educational Measurement, receives ongoing input from a panel of national testing experts regarding all plans for collecting validity evidence for the Texas assessments.

## **Evidence Based on Relations to Other Variables**

Another way to provide validity evidence is by analyzing the relationship between test performance and performance on some other measure. This other measure can be evaluated concurrently or in the future 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 was conducted in 2004–2005 to correlate performance on exit level TAKS with performance on national testing programs.

## **TAKS**

Further criterion-related evidence of validity for TAKS was provided in a study conducted by TEA and PEM to fulfill the Senate Bill 103 requirement that TEA implement a college readiness component as part of the TAKS. The research, called the Higher Education Readiness Component study, included two parts: a contrasting groups study and a performance data correlation study. The contrasting groups study examined the performance of high school juniors on the first administration of the TAKS exit level mathematics and English language

arts tests in 2003 as compared to performance on the same TAKS assessments by a sample of second semester college freshmen who had demonstrated college readiness through successful completion of their first semester courses.

The performance data correlation study examined student performance on TAKS in relation to performance on three college readiness measures used statewide for making college readiness and placement decisions: the Texas Academic Skills Program (TASP), the American College Test (ACT), and the Scholastic Assessment Test I (SAT I). The TAKS to TASP, TAKS to ACT, and TAKS to SAT I comparisons incorporated data collected from Texas public high school juniors who took the exit level TAKS and one or more of these other assessments in spring 2003. ACT and SAT I data were also collected for high school juniors who took the TAKS in spring 2004.

Results of the study indicated that the TAKS scale scores at the Met Standard performance level predicted ACT scale scores of approximately 20 for mathematics. Based on a national study of high school graduates from 2002–2004, 45 percent of students scored at or above this ACT score. The TAKS scale scores at the Met Standard performance level predicted ACT scale scores of approximately 18 for English. Of the high school students in the national study, 67 percent scored at least this high on the ACT English test.

Results of the study also indicated that the TAKS scale scores at the Commended Performance level predicted ACT scale scores of approximately 27 for mathematics. Based on a national study of high school graduates from 2002–2004, 15 percent of students scored at or above this ACT score. The TAKS scale scores at the Commended Performance level predicted ACT scale scores of approximately 24 for English. Of the high school students in the national study, 67 percent scored at least this high on the ACT English test.

Results of the study indicated that the TAKS scale scores at the Met Standard performance level predicted an SAT I scale score of approximately 470 for mathematics. Based on a national study of high school graduates, 50 percent of students scored at or above this SAT I score. The TAKS scale scores at the Met Standard performance level predicted an SAT I scale score of approximately 460 for English. Based on a national study of high school graduates, 50 percent of students scored at or above this SAT I score.

Results of the study indicated that the TAKS scale scores at the Commended Performance level predicted an SAT I scale score of approximately 620 for mathematics. Based on a national study of high school graduates, 25 percent of students scored at or above this SAT I score. The TAKS scale scores at the Commended Performance level predicted an SAT I scale score of approximately 580 for English. Based on a national study of high school graduates, 25 percent of students scored at or above this SAT I score. See Appendix 21 for more details on the Higher Education Readiness Component study.

Another study that provided criterion-related validity evidence for the TAKS tests was the 2004 grade Correlation Study. The study compared the passing credit/no passing credit rates of grade 9 Texas students in their 2003–2004 academic year English I course with their pass/fail rates on the spring 2004 TAKS grade 9 reading test. Results indicated that 87 percent of students in the study passed the TAKS grade 9 reading test, while 85 percent received passing

credit in their English I course. Of all students in the sample, 77 percent passed the TAKS grade 9 reading test and received passing credit in their English I course, whereas 5 percent failed the TAKS grade 9 reading test and did not receive passing credit in their English course. A small percentage (10 percent) passed the TAKS grade 9 reading test but did not receive passing credit in their English I course; a smaller percentage (8 percent) received passing credit in their English I course but failed the TAKS grade 9 reading test. See Appendix 30 for more details on the grade correlation study.

## **Relationship Between RPTE and TAKS Reading/English Language Arts**

RPTE is designed to provide both a measure of progress in English language acquisition in reading and information about a LEP student's knowledge and understanding of grade-appropriate TEKS in the content area of reading. One way to evaluate whether the skills assessed on RPTE align with grade-level TEKS expectations in reading is to relate RPTE performance by proficiency level to mean scale score performance on TAKS reading/English language arts tests. The results of this analysis are located in Chapter 4: TELPAS.

In the 2004–2005 and 2005–2006 school years, other analyses between RPTE and TAKS were conducted. These analyses were conducted to provide additional evidence concerning the suitability of RPTE as an alternative reading assessment for LEP-exempt recent immigrants in Adequate Yearly Progress (AYP) performance measures.

One set of analyses was conducted to examine how similar the percentages of LEP-exempt students who meet the RPTE incremental progress standard are to the percentages of nonexempt LEP students who pass TAKS. The results of these analyses indicated that the percentages of students who met the RPTE incremental progress standard and TAKS standard were very similar. Specifically, approximately 66 percent of LEP-exempt recent immigrants met the RPTE incremental progress standard in spring 2004 while approximately 65 percent of nonexempt LEP students met the TAKS standard. The performance results from the 2005–2006 school year were even more similar. In 2005–2006 the percentage of LEP-exempt recent immigrants who met the RPTE incremental progress standard and the percentage of nonexempt LEP students who met the TAKS standard was the same, 68%. The results of these analyses suggest that the AYP performance hurdle is similar whether the RPTE progress standard or TAKS standard is used. Additional RPTE-TAKS analyses and data are provided in Appendix 31.

A study was also conducted in 2005–2006 which established links between RPTE scores and the TAKS performance categories of Met Standard (passing level) and Commended Performance (highest performance level). TAKS Met Standard concordance scores on RPTE were located near the cut point between the RPTE Advanced and Advanced High proficiency levels. TAKS Commended Performance levels were located nearer the top of the RPTE Advanced High score range. This study, which showed that the range of reading proficiency assessed on RPTE is broad enough to link to both the passing category on TAKS and the higher commended level category, can help school districts gain a better understanding of the way in which high performance on RPTE relates to high performance on TAKS. It also provides validity evidence of the alignment

between the reading skills assessed on RPTE and TAKS. The results of this study are consistent with second language acquisition theory in that one would expect a strong relationship between high scores on an English language reading proficiency test aligned to grade-level reading skills and passing rates on a grade-level criterion-referenced reading achievement test. More information about this linking study may be found in Appendix 31.

