

CHAPTER 4: TEXAS ENGLISH LANGUAGE PROFICIENCY ASSESSMENT SYSTEM (TELPAS)

Overview

Title III, Part A of the No Child Left Behind (NCLB) Act of 2001 requires states to conduct annual statewide English language proficiency assessments that measure the progress limited English proficient (LEP) students in grades K–12 make in learning the English language. In accordance with NCLB, these assessments must assess four language domains: listening, speaking, reading, and writing.

To meet these federal requirements, the Texas English Language Proficiency Assessment System (TELPAS) was developed. This assessment system consists of the Reading Proficiency Tests in English (RPTE), which have been administered in Texas since the 1999–2000 school year, and the Texas Observation Protocols (TOP), which were benchmarked in spring 2004 and implemented fully in spring 2005. This assessment system is used to show the extent to which districts and the state meet federal Annual Measurable Achievement Objective (AMAO) accountability indicators. Composite ratings, rather than individual language domain ratings, are used in the AMAO accountability measures. For information on how the TELPAS composite ratings and composite scores are generated, see Appendix 7.

RPTE Development

The impetus for the development of the Reading Proficiency Tests in English came from state legislation passed in 1995, which required the commissioner of education to propose a way to evaluate the progress of LEP students eligible by state law for exemption from the state-mandated assessments. After seeking input from national experts and state stakeholders, the commissioner proposed to develop a test that would measure the progress LEP students make annually in learning to read in English. Such a test would not only be useful for measuring student progress in the state reading curriculum but would also provide educators with valuable information about how well English learners were able to understand the written English used in other content areas and on the state assessments.

The Reading Proficiency Tests in English enable Texas schools to monitor whether LEP students are making steady annual progress in English development during the time they qualify for an exemption from TAKS under state law. RPTE has been carefully designed to assess what LEP students can read and comprehend at distinct stages of learning English. Educators who understand the stages of second language development are better able to help English learners progress from one stage to the next.

In May 1998 TEA convened an advisory committee of Texas educators, assessment specialists, and state and national experts in second language acquisition to provide direction regarding the design of RPTE.

In subsequent steps of test development, groups of educators helped identify which TEKS student expectations to assess. In addition, national consultants as well as Texas bilingual and English as a second language (ESL) teachers articulated what second language learners can and cannot read and understand as they progress from one stage of second language development, or proficiency level, to the next. Afterward, the testing contractor, in collaboration with TEA staff, developed prototype test items. Texas educator committees reviewed these prototypes and provided the necessary input to allow the test developer to draft a large number of items, which were field-tested in spring 1999. Additional committees of educators reviewed the field-test items before and after the field test and determined which items were appropriate for inclusion on tests. Finally, before implementing RPTE in spring 2000, TEA conducted a field study to determine the appropriate length of the tests and configuration of test items. Numerous districts from the previous spring’s field test assisted with this study by administering pilot test forms and providing input.

The Reading Proficiency Tests in English were developed for each of the following grade clusters: grade 3, grades 4–5, grades 6–8, and grades 9–12. Each RPTE test was constructed with items that assess the characteristics and abilities associated with three levels of second language reading proficiency—beginning, intermediate, and advanced. (Recent changes to RPTE grade clusters and the addition of an advanced high proficiency rating are discussed below.) Following the spring 2000 baseline administration, statistical analyses were performed to determine the range of raw scores representing each proficiency level. Subsequent reviews by educators and psychometricians concluded that these ranges appropriately identified the stages of language development assessed. The table below shows the raw score ranges from the spring 2000 baseline administration. An equating process is used to ensure that subsequent tests maintain the same levels of difficulty as the baseline tests.

Table 4. Spring 2000 RPTE Raw Score Ranges and Proficiency Levels

RPTE Test	Proficiency Rating		
	Beginning	Intermediate	Advanced
Grade 3	0–37	38–47	48–56
Grades 4–5	0–37	38–47	48–56
Grades 6–8	0–39	40–48	49–57
Grades 9–12	0–39	40–49	50–60

The next step in developing the RPTE assessments was a scaling study conducted in fall 2000. This study generated data used to develop a cohesive scaled score system for documenting students’ annual progress.

Prior to the 2004–2005 school year, RPTE did not assess the upper range of second language development in reading. English language learners who reached the advanced RPTE proficiency level in a given year did not take RPTE in subsequent years. The TAKS tests were used to determine whether LEP students had attained the full range of English reading skills expected at their grade level. To fulfill NCLB requirements for annually assessing the English language proficiency in reading of all LEP students, the state of Texas implemented two RPTE

policy changes in the 2004–2005 school year. First, an advanced high RPTE proficiency rating was provided with RPTE test results. Second, LEP students were required to take RPTE annually, until they met the state criteria for reclassification as non-LEP. These policy changes would enable RPTE to show whether LEP students who demonstrated performance at the advanced level in a previous year were continuing to make progress in second language acquisition.

Several options were considered when proposing cut scores for the advanced high proficiency category. It was determined that the advanced high rating should identify English language learners who needed minimal second language acquisition support when reading grade-appropriate texts and strongly correlate with the Met Standard level of performance on the TAKS reading and English language arts tests. Options included setting the cut scores:

1. halfway between the cut score for the advanced proficiency level and the top of the scale (maximum scale score attainable on the 2005 test);
2. three-fourths of the way between the cut score for the advanced proficiency level and the top of the scale; or
3. at the scale score for which the predicted TAKS score is just above 2100 (that is, predicting passing TAKS).

The impact of setting different scale score cuts was investigated. Specifically, the percentage of students who would have been placed in the advanced high proficiency category had it existed in previous years and how those students performed on TAKS the following year was noted. An English language learner (ELL) focus group that assists the Texas Education Agency with LEP student assessment matters reviewed the data and collaborated with the test contractor and agency staff in establishing the final cut scores. Appendix 8 provides the technical information associated with establishing the advanced high RPTE cut scores.

Each year, new RPTE test items are written, reviewed by Texas educators, and field-tested in the spring as part of the ongoing test development needed to support annual test construction and the release of these tests to the public.

RPTE’s Role in the State Assessment Program

The Reading Proficiency Tests in English are used in a number of ways in the Texas Student Assessment Program. These assessments help schools monitor whether recent immigrant students are making steady progress in English reading proficiency during the time they may qualify under state law for a LEP exemption from other state assessments. RPTE also provides valuable diagnostic information about the English development of LEP students who are not eligible for an exemption and who participate in English or Spanish versions of the state assessments.

The Reading Proficiency Tests in English were developed originally to align with the previous state assessment program, the Texas Assessment of Academic Skills (TAAS). As of spring 2004,

RPTE began to more closely reflect the design of TAKS reading selections and test questions. As such, RPTE provides a strong content link to the Texas curriculum and TAKS assessments.

More information about the RPTE test design and test items can be found in the following resources, which are available on the TEA Student Assessment Division website.

- *Reading Proficiency Tests in English (RPTE) Information Booklet*
- *LPAC Decision-Making Process for the Texas Assessment Program (Grades 3–12): Procedural Manual*

Relationship Between Performance on RPTE and TAKS in Spring 2006

Every spring a large number of LEP students take both the RPTE and the TAKS reading or English language arts tests. The table on the following page shows the relationship between the performance of LEP students on both assessments. It is important to keep in mind that a small number of recent immigrants qualify under state law for a limited English proficiency exemption from TAKS during their initial years in the U.S. and that some LEP students in grades 3–6 take Spanish versions rather than English versions of TAKS. Included in the following table are only those nonexempt LEP students who took both RPTE and the English versions of TAKS.

Table 5. 2006 TAKS Scale Score Performance by RPTE Proficiency Rating for Students Who Participated in Both Assessments

TAKS Grade	RPTE Proficiency Rating	N	Mean TAKS Reading/ELA Scale Score
3	Beginning	1,174	1935.72
	Intermediate	6,241	2043.39
	Advanced	12,463	2160.83
	Advanced High	26,067	2324.55
4	Beginning	349	1867.87
	Intermediate	2,267	1951.96
	Advanced	12,732	2073.03
	Advanced High	13,516	2213.34
5	Beginning	202	1836.88
	Intermediate	1,190	1866.57
	Advanced	9,306	1982.07
	Advanced High	17,717	2133.48
6	Beginning	147	1876.32
	Intermediate	858	1924.20
	Advanced	8,826	2057.48
	Advanced High	9,857	2212.75
7	Beginning	192	1796.68
	Intermediate	912	1836.21
	Advanced	8,645	1949.58
	Advanced High	8,709	2080.44
8	Beginning	139	1852.59
	Intermediate	642	1851.80
	Advanced	6,122	1937.60
	Advanced High	9,009	2087.97
9	Beginning	1,010	1857.06
	Intermediate	2,002	1910.79
	Advanced	10,054	2024.82
	Advanced High	4,806	2156.14
10	Beginning	437	1892.08
	Intermediate	1,034	1945.90
	Advanced	6,047	2029.57
	Advanced High	4,185	2126.48
11	Beginning	292	1908.60
	Intermediate	732	1959.25
	Advanced	4,658	2036.21
	Advanced High	3,815	2139.62

Second Edition of RPTE

A second edition of the Reading Proficiency Tests in English (RPTE II) is being developed to add a grade 2 assessment and more fully align the RPTE advanced high proficiency level with the academic reading demands of core content areas of the Texas curriculum. RPTE II will address NCLB Title III goals for evaluating whether English language learners develop the academic English language proficiency that enables them to understand and process the English encountered in mathematics and science written materials. In addition, the grade clustering of RPTE II tests will be narrower at the secondary level. RPTE has grade cluster tests for grades 3, 4–5, 6–8, and 9–12. The planned grade clusters for RPTE II are grades 2, 3, 4–5, 6–7, 8–9, and 10–12.

In the 2004–2005 school year, educator committees convened to review proposed assessment objectives, TEKS student expectations, and prototype items for the redesigned assessment. The prototype test items were piloted in spring 2005. In summer 2005 Texas educators convened to review the prototype items in conjunction with the student performance data gathered during the pilot. Using information from the pilot, the test developer generated large numbers of items that were field-tested in spring 2006.

RPTE II field-test items for grades 3–12 were embedded in the spring 2006 operational RPTE test booklets. Because certain grade clusters for RPTE II will change, the spring 2006 test booklets were reconfigured to allow both the RPTE operational tests and embedded RPTE II field-test items to be administered to the appropriate students. In summer 2006 field-test data review committees composed of Texas educators convened to review these items and make determinations about which items to place in RPTE II item banks.

Because it is anticipated that RPTE II will be developed exclusively as an online assessment, the spring 2006 grade 2 RPTE II field test included an online pilot component designed to examine the feasibility of online administrations for limited English proficient second graders. On a small number of volunteer campuses, grade 2 students were randomly assigned to either an online or paper version of the RPTE II field test. As part of the study, test administrators and coordinators completed survey questions about online test administration procedures and their perceptions about how well grade 2 students interacted with the online field test. The survey results from this small pilot provided valuable feedback and indicated that the participating second grade English language learners were able to engage in the online testing process. In addition, student performance data on the online and paper versions of the grade 2 RPTE II field test were compared. Item-level analyses indicated that item statistics were very similar across online and paper testing. This pilot, though small in scale and voluntary, generated findings that were clearly positive enough for the state to move forward with plans for larger scale RPTE II online field testing. For more detail about this grade 2 online pilot study, see Appendix 9. In spring 2007, online field testing of RPTE II items for grades 2–12 is planned.

Texas Observation Protocols (TOP) Development Activities

Title III, Part A of NCLB mandates that states conduct annual English language proficiency assessments to demonstrate the progress of LEP students in grades K–12 in the domains of listening, speaking, reading, and writing. RPTE assesses the domain of reading in grades 3–12. In spring 2007 TEA implemented the Texas Observation Protocols (TOP) to measure the other domains and grades, as shown below.

- grades K–2 listening, speaking, reading, and writing
- grades 3–12 listening, speaking, and writing

TOP was developed by TEA in collaboration with test development contractors, bilingual/ESL consultants, and members of an English language learner focus group composed of teachers, bilingual/ESL directors, assessment directors, campus administrators, and university professors. TOP is aligned with the Texas academic content and achievement standards and draws upon second language acquisition research, research-based standards, the experience of Texas practitioners, and observational assessment processes used in other states. TOP allows teachers to holistically rate each LEP student’s English language proficiency based on classroom observations and daily interactions with the student.

TEA chose to implement an observation protocol assessment instrument for a number of reasons.

- This assessment method, which is integrally linked with sound classroom instructional practices, has the potential for making a significant, positive impact on student learning. The training teachers receive in the holistic rating and classroom observation process helps them better understand and meet the needs of English language learners.
- TOP does not require annual field testing.
- The administration of TOP does not require students to lose instructional time.
- TOP does not have the logistical drawbacks associated with tape-recorded or individually administered speaking and listening assessments required by other types of English language proficiency tests.

TOP Rating and Training Process

For the administration of TOP, districts identify their K–12 LEP students and designate teachers to observe and rate the students’ English language proficiency. TOP is designed to capture an overall level of English language acquisition and is a holistically scored assessment rather than an assessment of isolated skills. Teachers are trained to use holistic rating rubrics called TOP Proficiency Level Descriptors to assign one of four proficiency ratings—beginning, intermediate, advanced, or advanced high—in each domain assessed. The ratings are submitted to the state’s testing contractor, and performance reports are generated and sent

back to the districts. The TOP Proficiency Level Descriptors are available on the TEA Student Assessment Division website.

TOP raters received holistic rating training through a training-of-trainers model. In response to field input, the TOP training model was changed during 2005–2006 to move the majority of TOP holistic rating training to the fall of the year. The goals of fall training are to

- train teachers to use the TOP classroom observational processes throughout the school year to maximize student learning;
- ensure that teachers receive a thorough holistic rating training; and
- reduce the amount of training required in the spring.

Training materials included annotated examples of student performance and authentic student writing samples for use in training raters to apply the holistic rating rubrics.

Also in 2005–2006, a qualifying component was added to the TOP training model to ensure that teachers are adequately prepared to rate grades 2–12 writing collections. All raters, except those who rated K–1 students only, were assigned particular grade clusters in which to qualify online between January 30 and March 3, 2006. Individuals successful on the qualifying activities received a training certificate qualifying them to rate student writing collections independently.

Finally, beginning in the 2005–2006 school year, TOP raters were required to base the writing ratings of students in grades 2–12 on the students' writing collections only, rather than on additional classroom observations. TOP ratings for approximately 630,000 English language learners were assigned during the March 6–31, 2006, TELPAS testing window. District, campus, and student TELPAS reports indicating TOP and RPTE performance were provided by the test contractor in May 2006.

Spring 2006 TOP Audit

Following the 2006 TOP administration, a state audit of selected student writing collections and other TOP assessment processes was conducted to examine the effectiveness of the training and application of the holistic rating rubrics. As part of the spring 2006 audit, 13,357 total writing collections originally rated by district raters were re-rated by the state. District ratings of the audit writing collections were in perfect agreement with the state ratings 77% of the time. This overall perfect agreement rate of 77% was found to be acceptable by the state for the spring 2006 audit. According to the International Standards Organization (ISO), a perfect agreement rate of 70% is considered acceptable for a holistically scored writing assessment that uses a 4-point rating scale. Furthermore, the audit revealed an adjacent agreement rate of 98%, which means that 98% of the time the district and state audit ratings were either the same or one proficiency level apart.

Teachers who rated the audited writing collections completed survey questions about the TOP rating process for all language domains. Results of the survey indicated that 97% of the

raters felt they had enough information about the writing abilities of the students to make judgments about their proficiency levels. Approximately 93% of the raters felt that they received sufficient training for the online qualification activities, and 92% of the raters reported that they successfully completed the online qualifying component of the TOP training. TEA plans to conduct TOP audits annually. A report with more information on the spring 2006 TOP audit is included as Appendix 10.

