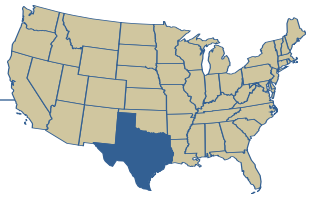
Texas



Introduction

This study linked data from the 2003 and 2006 administrations of Texas's reading and math tests to the Northwest Evaluation Association's Measures of Academic Progress (MAP) assessment, a computerized adaptive test used in schools nationwide. We found that Texas's definitions of proficiency are relatively less difficult than the cut scores set by the other 25 states in this study in reading and mathematics. In other words, Texas's tests are below average in terms of difficulty.

Still, the level of difficulty has increased from 2003 to 2006—the No Child Left Behind era—though more so for some grades than others. Texas is one of the few states in this study whose cut scores have become more challenging over time. Even so, the state's expectations are not consistent from one grade to the next and policymakers should consider more closely calibrating them to ensure equivalent difficulty at all grades. In this way, parents and schools can be assured that elementary school students scoring at the proficient level are truly prepared for success later in their educational careers.

What We Studied: Texas Assessment of Knowledge and Skills (TAKS)

Texas currently uses the Texas Assessment of Knowledge and Skills (TAKS), which tests students in reading in grades 3 through 9; in writing in grades 4 and 7; in English/language arts in grades 10 and 11; in mathematics in grades 3 through 11; in science in grades 5, 10, and 11; and social studies in grades 8, 10, and 11. The Spanish TAKS is administered in grades 3 through 6. Satisfactory performance on the TAKS at grade 11 is prerequisite to a high school diploma. TAKS was first administered in the 2002-2003 school year.

To determine the difficulty of Texas's proficiency cut scores, we linked data from state reading and math tests from a group of elementary and middle schools to the NWEA assessment. (A "proficiency cut score" is the score a student must achieve in order to be considered proficient.) This was done by analyzing a group of schools in which almost all students took both the state's assessment and the NWEA test. (The methodology section of this report explains how performance on these two tests was compared.)

Part 1: How Difficult are Texas's Definitions of Proficiency in Reading and Math?

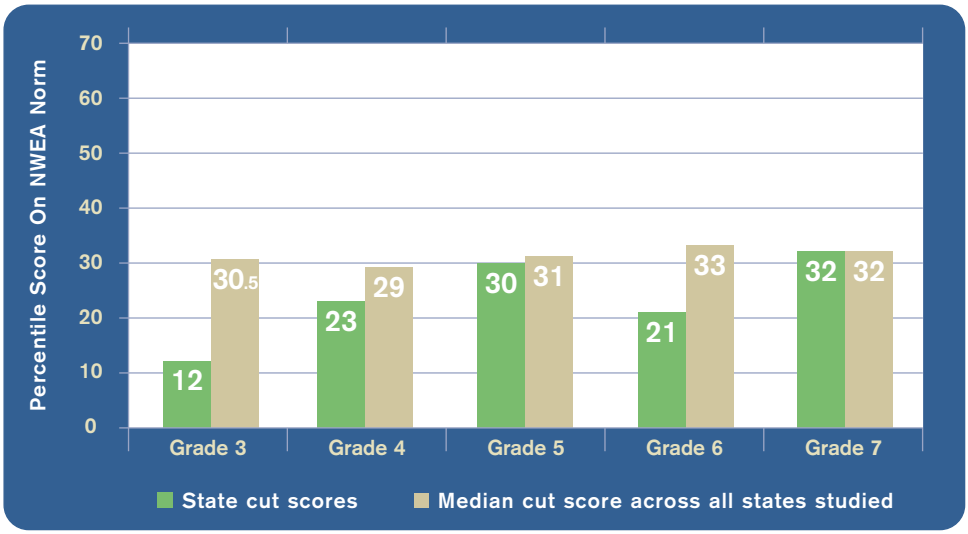
One way to evaluate the difficulty of a standard is to determine how many people attempting to attain are likely to succeed. How do we know that a two-foot high bar is easy to jump over? We know because, if we asked 100 people at random to attempt such a jump, perhaps 80 would make it. How do we know that a six-foot high bar is challenging? Because only one (or perhaps none) of those same 100 individuals would successfully meet that challenge. The same principle can be applied to academic standards. Common sense tells us that it is more difficult for students to solve algebraic equations with two unknown variables than it is for them to solve an equation with only one unknown variable. But we can figure out exactly how much more difficult by seeing how many eighth graders nationwide answer both types of questions correctly.

Applying that approach to this assignment, we evaluated the difficulty of Texas's proficiency standards by estimating the proportion of students in NWEA's norm group who would perform above the Texas standard on a test of equivalent difficulty. The following two figures show the difficulty of Texas's proficiency cut scores for reading (Figure 1) and mathematics (Figure 2) in 2006 in relation to the median cut score for all the states in the study. Sample sizes were sufficient to generate cut score estimates for reading and math in grades 3 through 7. Grade-8 cut scores were not available. The proficiency cut scores for **reading** in Texas ranged between the 12th and 32nd percentiles nationally, with the seventh grade being most challenging. In **mathematics**, the proficiency cut scores ranged between the 24th and 41st percentiles with the seventh grade again being most challenging.

For most grade levels, Texas's cut scores in both reading and mathematics are below the median level of difficulty among the states studied. Note, though, that Texas's cut scores for reading are generally less difficult than the corresponding mathematics cut scores within a given grade. Thus, reported differences in achievement between the two subjects may be more a product of differences in cut scores than in actual student achievement. In other words, Texas students may be performing worse in reading and better in mathematics than is apparent by looking at the percentage of students passing state tests in those subjects.

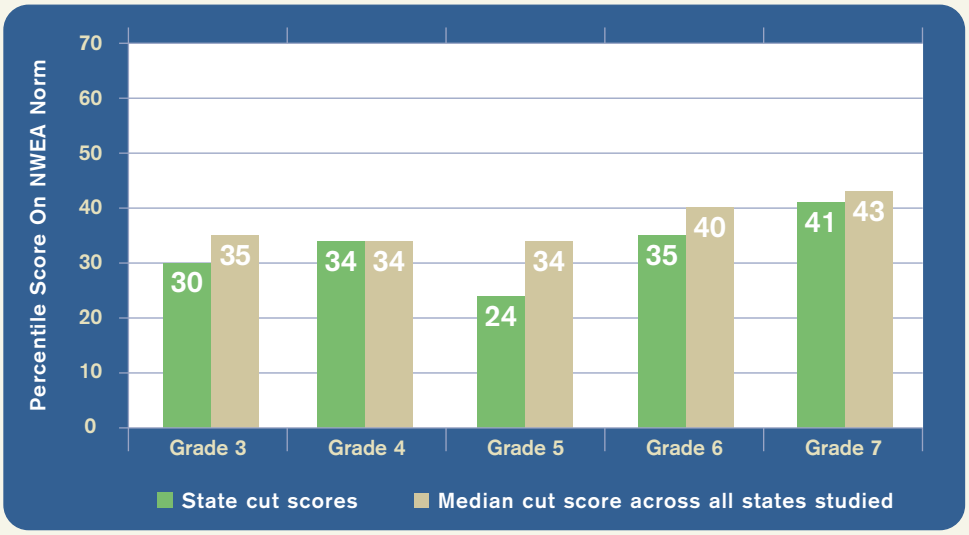
Another way of assessing difficulty is to evaluate how Texas's proficiency cut scores rank relative to other states. Table 1 shows that the Texas cut scores generally rank in the lower half for reading and the upper half for mathematics, among the 26 states studied for this report. Texas's third- and fourth-grade reading cut scores are particularly low, besting only two and six other states in the study, respectively. On the other hand, Texas ranks relatively high in third- and fourth-grade math.

Figure 1 – Estimate of Texas Reading Cut Scores in Relation to All 26 States Studied, 2006 (Expressed in MAP Percentiles)



Note: This figure compares reading test cut scores ("proficiency passing scores") as percentiles of the NWEA norm. These percentiles are compared with the median cut score of all 26 states reviewed in this study. Only in grades 5 and 7 do Texas's cut scores approach or equal the median.

Figure 2 – Estimate of Texas Mathematics Cut Scores in Relation to All 26 States Studied, 2006 (Expressed in MAP Percentiles)



Note: Texas's math-test cut scores are shown as percentiles of the NWEA norm and compared with the median cut score of all 26 states reviewed in this study. Only in fourth grade does Texas's cut score reach the median.

Table 1 – Texas Rank for Proficiency Cut Scores Among 26 States in Reading and Mathematics, 2006

	Ranking (Out of 26 States)				
	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7
Reading	24	20	14	22	13
Mathematics	14	13	20	16	15

Note: This table ranks Texas's cut scores relative to the cut scores of the other 25 states in the study, with 1 being highest and 26 lowest.

Part 2: Differences in Cut Scores over Time

In order to measure their consistency, Texas's proficiency cut scores were mapped to their equivalent scores on NWEA's MAP assessment for the 2002-3 and 2005-6 school years. Cut score estimates for both years were available for grades 3 through 7 for reading and grades 4 and 7 for mathematics.

States may periodically re-adjust the cut scores they use to define proficiency in reading and math, or may update the tests used to measure student proficiency. Such changes can impact proficiency ratings, not necessarily because student performance has changed, but because the measurements and criteria for success have changed.

This was certainly the case for Texas. When the Texas Assessment of Knowledge and Skills (TAKS) was introduced in 2002-03, the Texas Education Agency formally adopted cut scores that would increase in difficulty over the first three years of testing. This was meant to give schools and students an opportunity to adjust to the new test and its expectations.

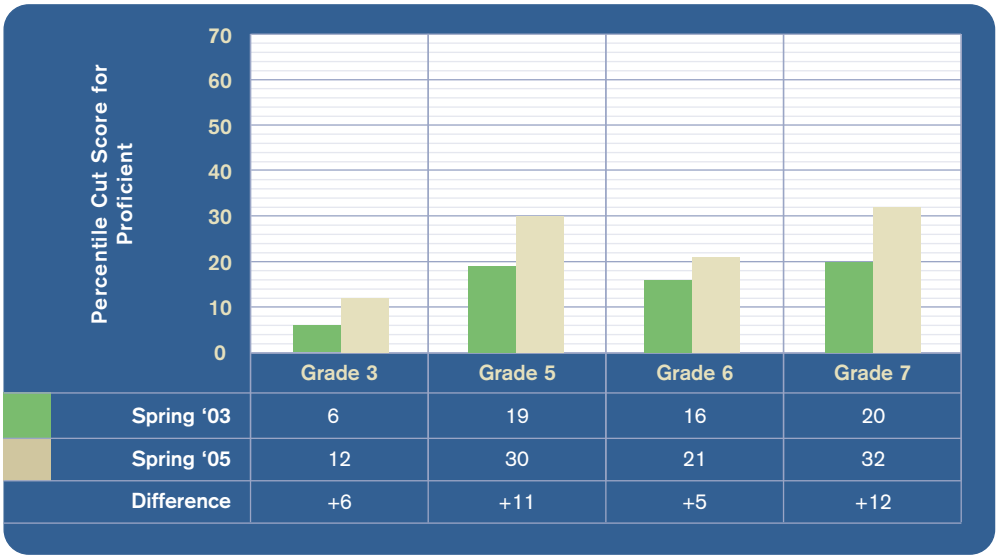
Is it possible, then, to compare the proficiency scores across this three-year period? Yes. Assume that we're judging a group of fourth graders on their high-jump prowess and that we measure this by finding how many in that group can successfully clear a three-foot bar. Now assume that we change the measure and set a new height. Perhaps students must now clear a bar set at one meter. This is somewhat akin to adjusting or changing a state test and its proficiency requirements. Despite this, it is still possible to determine whether it is more difficult to clear one meter than three feet, because we know the relationship between the measures. The same principle applies here. The measures or scales used by the TAKS in 2003 and 2006 can both be linked to the scale that was used to report MAP, which has remained consistent over time. Just as one can compare three feet to one meter and know that a one-meter jump is slightly more difficult than a three-foot jump, one can estimate the cut score needed to pass the TAKS in 2003 and 2006 on the MAP scale and ascertain whether the test may have changed in difficulty.

Texas's estimated **reading** cut scores indicate that, as intended by the state, the proficiency cut scores increased in difficulty over this three-year period for all available grades (see Figure 3). Consequently, even if student performance stayed the same on an equivalent test like NWEA's MAP assessment, one would expect the reading proficiency rates in 2006 to be lower than they were in 2003. These more difficult cut scores would likely yield 6 percent, 11 percent, 5 percent, and 12 percent decreases in the proficiency rates for third, fifth, sixth, and seventh grade students, respectively. (Texas reported an 8-point decline for grade 7, although proficiency rates in grades 3, 5 and 6 actually increased by 4, 1, and 5 points, respectively.)

Texas's estimated **mathematics** cut scores showed similar patterns, with increases over three years in the difficulty of the proficiency cut scores for grades 5 and 7 (see Figure 4). Consequently, even if student performance stayed the same on an equivalent test like NWEA's MAP assessment, these higher proficiency cut scores would likely yield decreases of 11 percent and 16 percent in the math proficiency rates for fifth and seventh graders, respectively. (Texas reported a 5-point decline for fifth graders and a 3-point decline for seventh graders over this period.)

Thus, one could fairly say that Texas's tests were harder to pass in 2006 than in 2003. As a result, improvements in actual student performance were been masked somewhat by the increased difficulty of the state's proficiency cut scores.

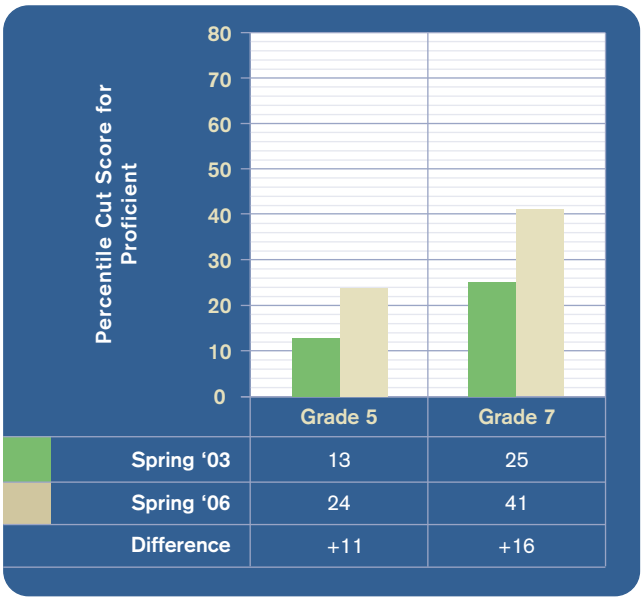
Figure 3 – Estimated Differences in Texas's Proficiency Cut Scores in Reading, 2003-2006 (Expressed in MAP Percentiles)



	Grade 3	Grade 5	Grade 6	Grade 7
Spring '03	6	19	16	20
Spring '05	12	30	21	32
Difference	+6	+11	+5	+12

Note: This graphic shows how the degree of difficulty in achieving proficiency in reading has changed. For example, third-grade students in 2003 had to score at the 6th percentile on the NWEA norm group in order to be considered proficient, while in 2006 third graders had to score at the 12th percentile of the NWEA norm group to achieve proficiency.

Figure 4 – Estimated Differences in Texas's Proficiency Cut Scores in Mathematics, 2003-2006 (Expressed in MAP Percentiles)



	Grade 5	Grade 7
Spring '03	13	25
Spring '06	24	41
Difference	+11	+16

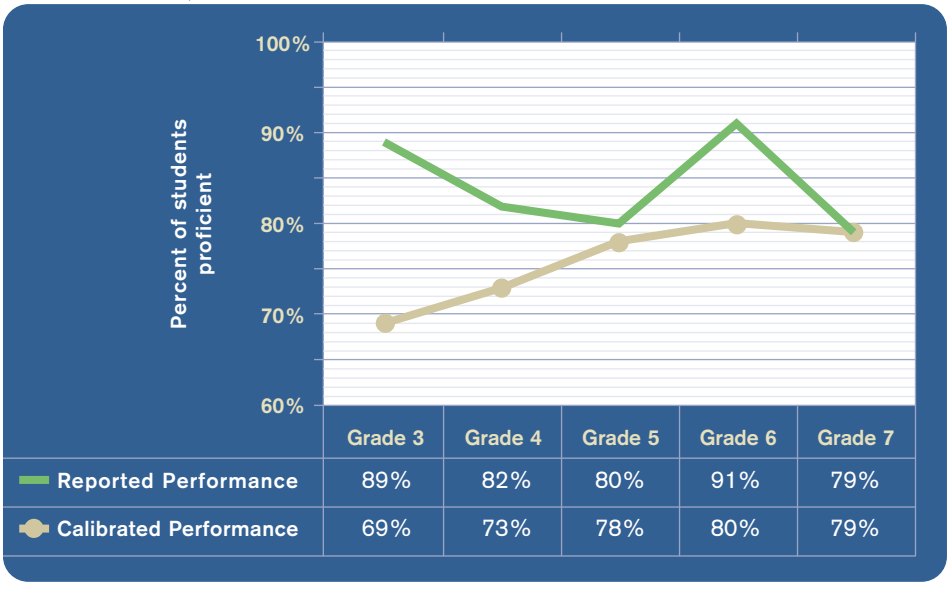
Note: This graphic shows how the degree of difficulty in achieving proficiency in math has changed. For example, fifth-grade students in 2003 had to score at the 13th percentile on the NWEA norm group in order to be considered proficient, while in 2006 fifth graders had to score at the 24th percentile of the NWEA norm group to achieve proficiency.

Part 3: Calibration across Grades

Calibrated proficiency cut scores are those that are relatively equal in difficulty across all grades. Thus, an eighth-grade cut score would be no more or less difficult for eighth graders to achieve than a third-grade cut score is for third graders. When cut scores are so calibrated, parents and educators have some assurance that achieving the third-grade proficiency cut score puts a student on track to achieve the standards at eighth grade. It also provides assurance to the public that reported differences in performance across grades are a product of differences in actual educational attainment and not simply differences in the difficulty of the test.

Figures 1 and 2 showed that Texas's upper-grade cut scores in reading and mathematics were more challenging than the cut scores in the lower grades, particularly in grade 3. The two figures that follow show Texas's reported performance in reading (Figure 5) and mathematics (Figure 6) on the state test compared with the rate of proficiency that would be achieved if the cut scores were all calibrated to the grade-7 standard. When differences in grade-to-grade difficulty of the cut score are removed, student performance is more consistent at all grades. This would lead to the conclusion that the higher rates of proficiency that the state has reported for elementary school students are somewhat misleading.

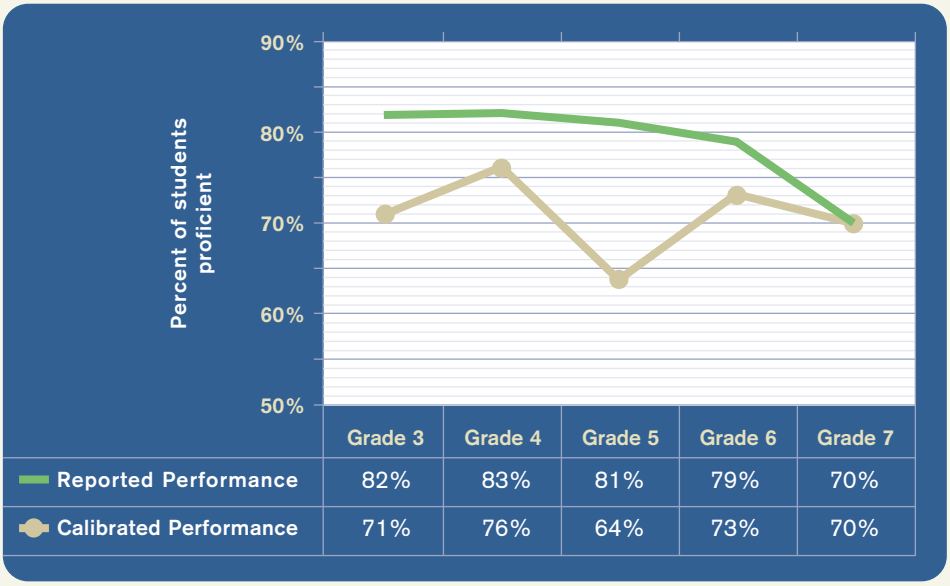
Figure 5 – Texas Reading Performance as Reported and as Calibrated to the Grade-7 Standard, 2006



	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7
Reported Performance	89%	82%	80%	91%	79%
Calibrated Performance	69%	73%	78%	80%	79%

Note: This graphic shows, for example, that if Texas's grade-3 reading cut score was set at the same level of difficulty as its grade-7 cut score, 69 percent of third graders would achieve the proficient level, rather than 89 percent, as was reported by the state.

Figure 6 – Texas Mathematics Performance as Reported and as Calibrated to the Grade-7 Standard, 2006



	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7
Reported Performance	82%	83%	81%	79%	70%
Calibrated Performance	71%	76%	64%	73%	70%

Note: This graphic shows, for example, that if Texas's grade-3 mathematics cut score was set at the same level of difficulty as its grade-7 cut score, 71 percent of third graders would achieve the proficient level, rather than 82 percent, as was reported by the state.

Policy Implications

When determining what constitutes proficiency, Texas is relatively low—more so in reading than in math—compared with the other 25 states in this study. This finding is consistent with the recent National Center for Education Statistics report, *Mapping 2005 State Proficiency Standards Onto the NAEP Scales*, which also found Texas's reading standards to be in the bottom third of the distribution of all 50 states, and the mathematics standards closer to the middle. In recent years, the difficulty of the proficiency cut scores has increased, though some grades have increased more than others. As a result, Texas's expectations are not smoothly calibrated across grades; students who are proficient in third grade are not necessarily on track to be proficient by the seventh grade. Texas policymakers might consider adjusting their cut scores across grades so that parents and schools can be assured that elementary school students scoring at the proficient level are truly prepared for success later in their educational careers.