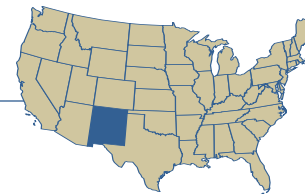


New Mexico



Introduction

This study linked data from the 2005 and 2006 administrations of New Mexico’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 New Mexico’s definitions of proficiency in reading are consistent with the cut scores set by the 25 other states in this study, while its definitions for mathematics proficiency are relatively more difficult. In other words, New Mexico’s reading tests are about average in terms of difficulty, while its math tests are above average.

However, the level of difficulty of New Mexico’s math tests declined somewhat from 2005 to 2006, although not for all grades. There are many possible explanations for these declines (see pp. 34-35 of the main report), which were caused by learning gains on the New Mexico test not being matched by learning gains on the Northwest Evaluation Association test. Additionally, New Mexico’s mathematics cut scores are now relatively less difficult for third-grade students than they are for eighth-grade students (taking into account the differences in subject content and children’s development). State policymakers might consider adjusting their math cut scores to ensure equivalent difficulty at all 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. Furthermore, state leaders need to be aware of the disparity between math and reading standards when evaluating differences in teacher and student performance across these domains.

What We Studied: New Mexico Standards Based Assessments (NMSBA)

New Mexico currently uses an assessment called the New Mexico Standards Based Assessments (NMSBA) which tests mathematics, language arts, and science in students in grades three through nine and math and language arts in grade 11. The tests were used in spring 2005. The current study linked reading and math data from spring 2005 and spring 2006 test administrations to a common scale also administered in the 2005 and 2006 school years.

To determine the difficulty of New Mexico’s proficiency cut scores, we linked data from NMSBA 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 elementary and middle 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 New Mexico’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 it 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 percent would make it. How do we know 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.

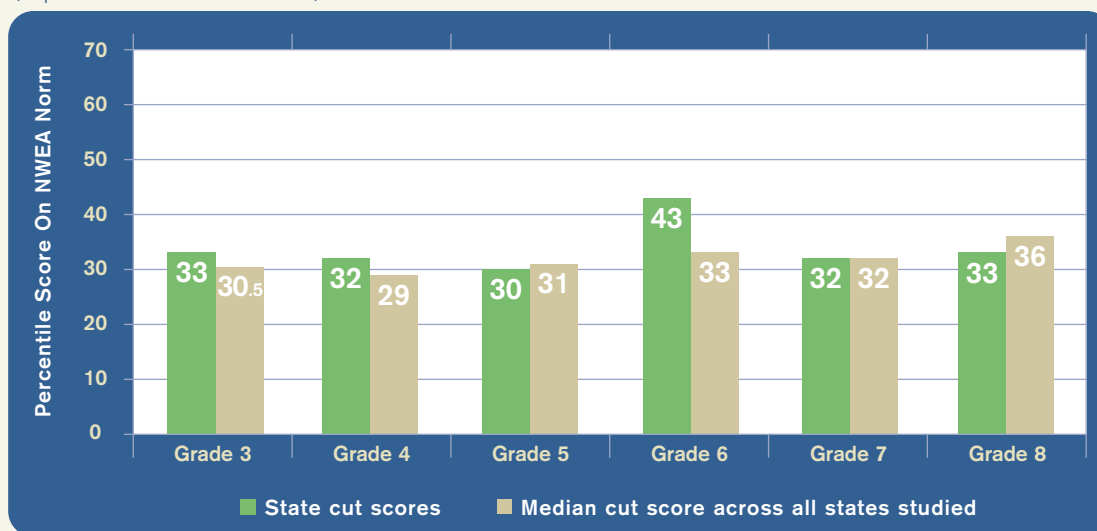
Applying that approach to this assignment, we evaluated the difficulty of New Mexico’s proficiency cut scores by estimating the proportion of students in NWEA’s norm group who would perform above the New Mexico cut score on a test of equivalent difficulty. The following two figures show the difficulty of New Mexico’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. The proficiency cut scores for **reading** in New Mexico ranged between the 30th and 43rd percentiles nationally, with sixth grade being most challenging. In **mathematics**, the proficiency cut scores ranged between the 46th and 61st percentiles, with seventh grade being most challenging.

Except in grade six, New Mexico’s reading cut scores are near the median difficulty of the states studied, whereas New Mexico’s mathematics cut scores are higher than the median in all grades. Note, too, that New Mexico’s cut scores for reading

are lower than the cut scores for mathematics. 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, New Mexico students may be performing worse in reading and/or better in mathematics than is apparent by just looking at the percentage of students passing state tests in those subjects.

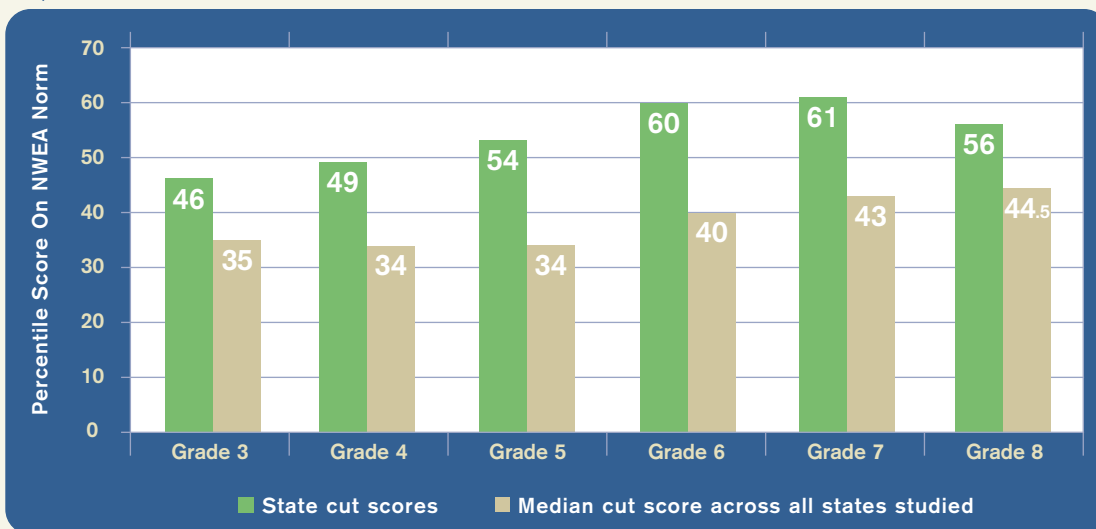
Another way of assessing difficulty is to evaluate how New Mexico’s proficiency cut scores rank relative to other states. Table 1 shows that the New Mexico reading cut scores generally rank in the top half in difficulty while math cut scores rank among the top three or four states in every grade.

Figure 1 – Estimates of New Mexico 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. New Mexico’s reading cut scores hover around the median, with the exception of grade 6, in which the state cut score is 10 percentile points higher.

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



Note: New Mexico'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. Across grades, New Mexico's math cut scores are above the median.

Table 1 – New Mexico Rank for Proficiency Cut Scores in Reading and Mathematics, 2006

Ranking (Out of 26 States)						
	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Reading	9	10	14	4	13	14
Mathematics	4	4	4	4	3	4

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

Part 2: Differences in Cut Scores over Time

In order to measure their consistency, New Mexico's proficiency cut scores were mapped to their equivalent scores on NWEA's MAP assessment for the 2005 and 2006 school years. Cut score estimates for reading and mathematics were available for both years in grades three through eight.

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. Plus, unintentional drift can occur even in states, such as New Mexico, that maintained their proficiency levels.

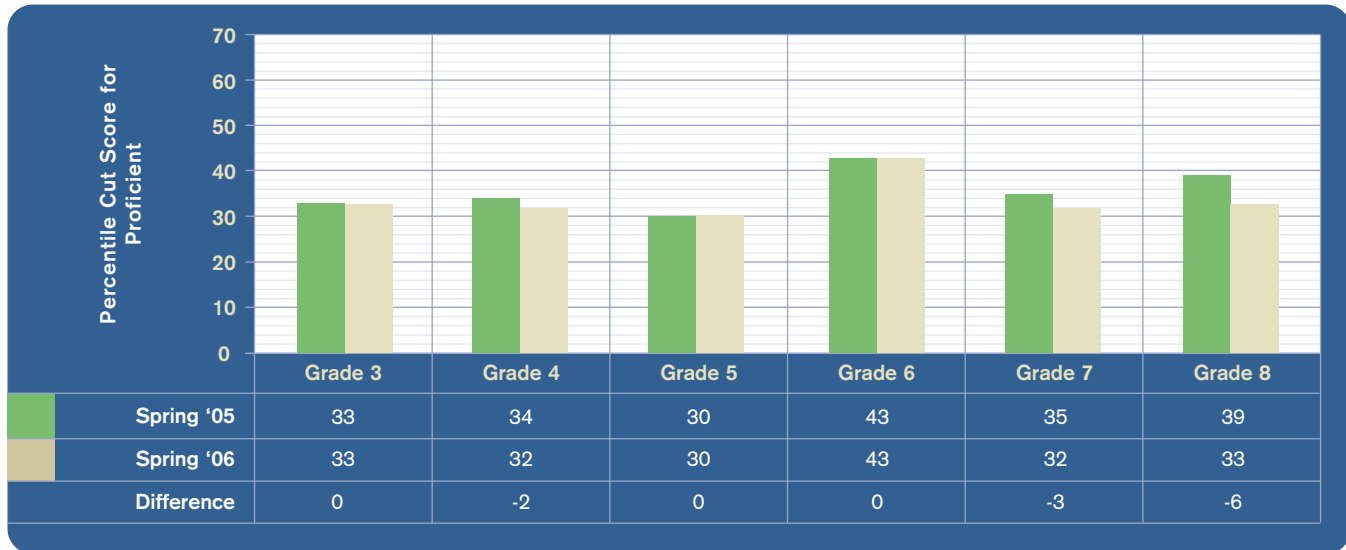
Is it possible, then, to make comparisons of the proficiency scores between earlier administrations of New Mexico tests and today's? Yes. Assume that we're judging a group of fourth graders on their high-jump prowess and that we measure this by finding how many in that group can successfully clear a 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 2005 and 2006 NMSBA can both be linked to the 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 scores needed to pass the NMSBA in 2005 and 2006 on the MAP scale and ascertain whether the state test may have changed in difficulty.

New Mexico's estimated **reading** cut scores indicate no substantive changes over this one-year period (see Figure 3). Consequently, one would expect that any changes in the reported reading proficiency ratings could be directly attributable to actual changes in student performance.

New Mexico's estimated **mathematics** cut scores show substantive decreases for grades six and eight (see Figure 4). Consequently, even if student performance stayed the same on an equivalent test like NWEA's MAP assessment, this would likely yield increases of seven and six percent, respectively, in the state-reported mathematics proficiency rates for those grades. (New Mexico reported a 2-point gain for sixth graders and a 2-point gain for eighth graders over this period.)

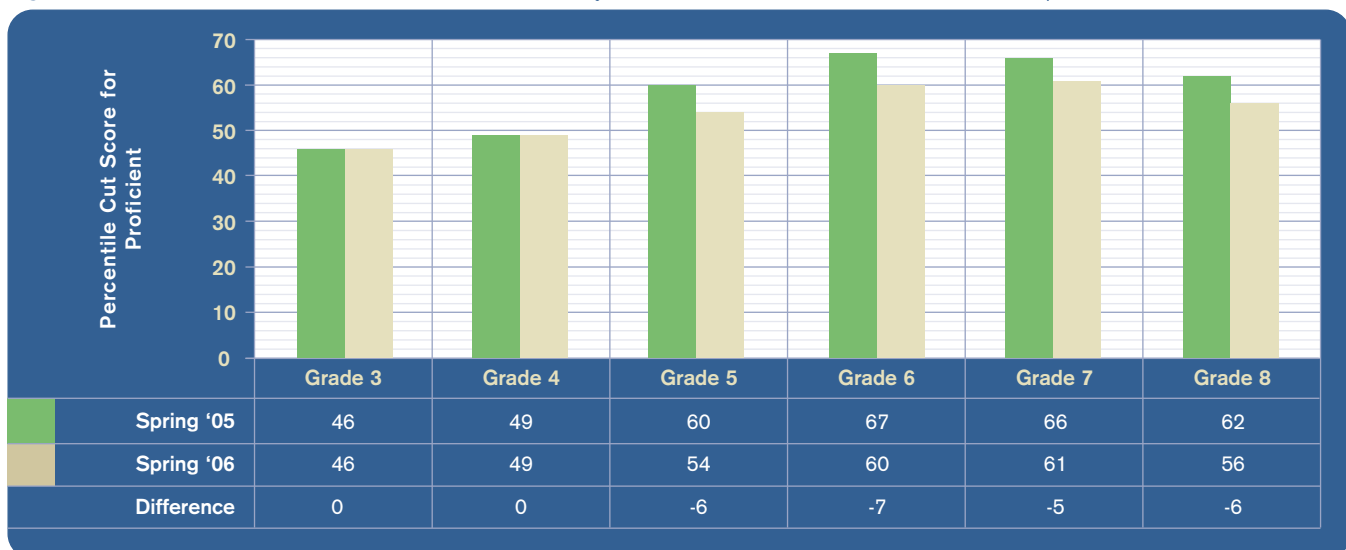
Thus, one could fairly say that New Mexico's reading tests remained about the same from 2005 to 2006, but that math tests for grades six and eight became easier to pass. As a result, some apparent improvements in the state's sixth- and eighth-grade mathematics proficiency rates during this period may not be entirely a product of improved achievement.

Figure 3 – Estimated Change in New Mexico's Proficiency Cut Scores in Reading, 2005-2006 (Expressed in MAP Percentiles)



Note: This graphic shows that the difficulty of achieving proficiency in reading has not changed. For example, third-grade students in 2005 had to score at the 33rd percentile on NWEA norms in order to be considered proficient, and in 2006 third graders still had to score at the 33rd percentile to achieve proficiency. The observed changes in all grades were within the margin of error (in other words, too small to be considered substantive).

Figure 4 – Estimated Difference in New Mexico's Proficiency Cut Scores in Mathematics, 2005-2006 (Expressed in MAP Percentiles)



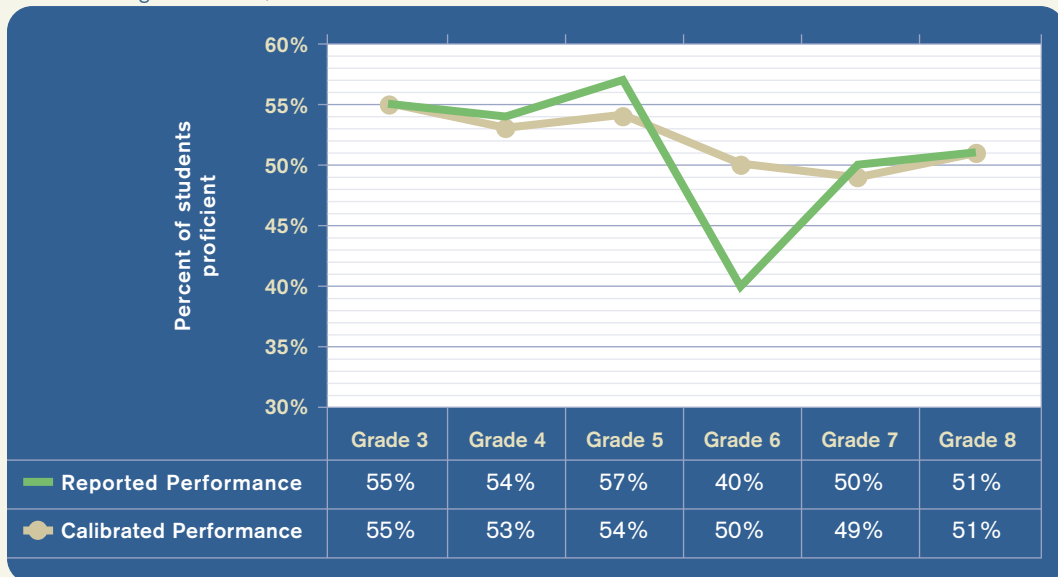
Note: This graphic shows how the difficulty of achieving proficiency in math has changed. For example, sixth-grade students in 2005 had to score at the 67th percentile on NWEA norms in order to be considered proficient, while in 2006 sixth graders had only to score at the 60th percentile to achieve proficiency. The changes in grades three, four, five, and seven were within the margin of error (in other words, too small to be considered substantive).

Part 3: Calibration across Grades

Calibrated proficiency cut scores are relatively equal in difficulty across all grades. Thus, the eighth-grade cut score is no more or less difficult to achieve for eighth graders than the third-grade cut scores is for third graders, respectively. When cut scores are all calibrated, to the grade-eight standard, parents and educators have some assurance that achieving the third-grade proficiency cut score puts a student on track to achieve the cut scores 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.

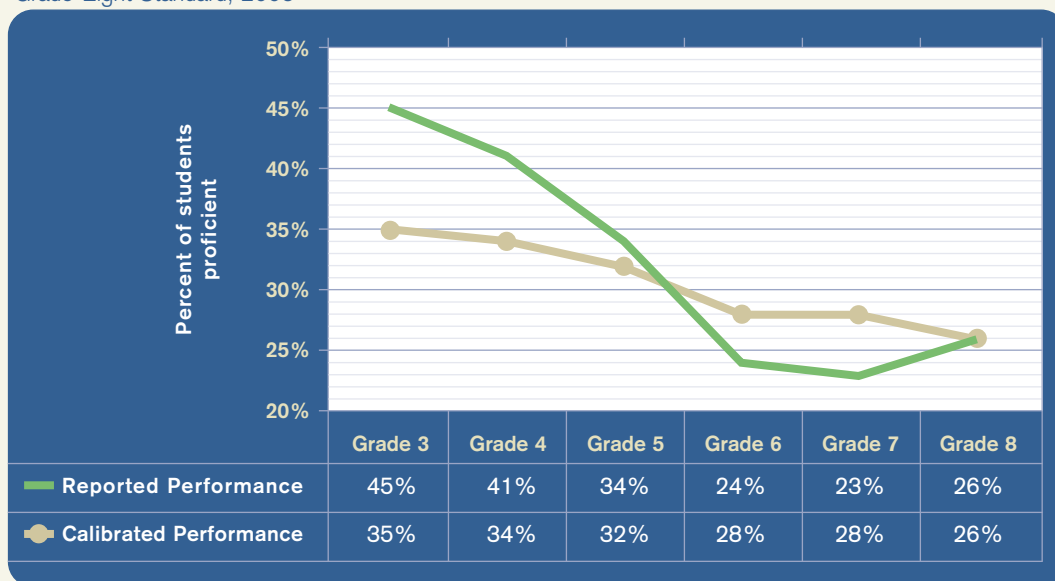
Figures 1 and 2 indicated the relative difficulty of the reading and math cut scores, showing that reading cut scores were consistent except in grade four, which was relatively more difficult. In mathematics, however, cut scores were less difficult in the lower grades than in the upper. (This pattern held true for most states studied.) The two figures that follow show New Mexico's reported performance in reading (Figure 5) and mathematics (Figure 6) on the state test, compared with the rates of proficiency that would be achieved if the cut scores were calibrated across grades. When grade-to-grade differences in difficulty of the cut score are removed, student performance is more consistent at all grades.

Figure 5 – New Mexico Reading Performance as Reported and as Calibrated to the Grade-Eight Standard, 2006



Note: This graphic shows, for example, that if New Mexico's grade-six reading cut score was set at the same level of difficulty as its grade-eight cut score, 50 percent of sixth graders would achieve the proficient level, rather than 40 percent, as was reported by the state.

Figure 6 – New Mexico Mathematics Performance as Reported and as Calibrated to the Grade-Eight Standard, 2006



Note: This graphic shows, for example, that if New Mexico's grade-three mathematics cut score was set at the same level of difficulty as its grade-eight cut score, 35 percent of third graders would achieve the proficient level, rather than 45 percent, as was reported by the state.

Policy Implications

New Mexico proficiency cut scores are relatively high in mathematics, at least compared to the other 25 states in this study. Its reading cut scores are about at the mid-point. This finding is fairly consistent with the recent National Center for Education Statistics report, *Mapping 2005 State Proficiency Standards Onto the NAEP Scales*, which also found New Mexico's standards to be in the upper-middle sector for reading and upper level for mathematics. Over the year-long span of time that cut scores were tracked for this study, the state's cut scores for mathematics have become less difficult in grades six and eight, although not in other grades.

Nonetheless, New Mexico's expectations in mathematics are still not smoothly calibrated across grades; students who are proficient in third grade are not necessarily on track to be proficient by the eighth grade. State policymakers might consider adjusting their math 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. Furthermore, state leaders need to be aware of the disparity between math and reading standards when evaluating differences in teacher and student performance across these domains.