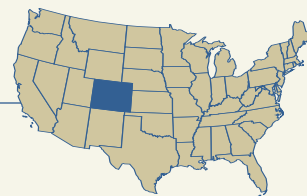


Colorado



Introduction

This study linked data from the 2002 and 2005 administrations of Colorado’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, for purposes of complying with the federal No Child Left Behind Act (NCLB), Colorado’s definitions of “proficiency” in reading and mathematics are much less difficult than the standards set by most of the other 25 states in this study. In other words, it’s easier to pass Colorado’s tests than those of almost all other states.

Moreover, the difficulty of Colorado’s tests decreased somewhat from 2002 to 2005—the NCLB era—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 Colorado test not being matched by learning gains on the Northwest Evaluation Association test. One finding of this study is that Colorado’s cut scores are now relatively less difficult at the lower grades than at the higher ones (taking into account the obvious differences in subject content and children’s development). Colorado policymakers might consider raising their standards in the earlier 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.

In this study, we used the proficiency cut scores that Colorado employs for purposes of NCLB to make comparisons. It’s well known that Colorado opted to use the state’s *partially proficient* level of academic performance as *proficient* for NCLB purposes. Hence we follow that practice here and subsequent references to “proficient” or “proficiency” in Colorado should be understood accordingly.

What We Studied: Colorado Student Assessment Program (CSAP)

Colorado currently uses an assessment called the Colorado Student Assessment Program (CSAP) which tests reading, writing, and math in grades 3-10 and science in grade 8. The same sets of tests were used in spring 2002 in which reading and writing were administered in grades 3-10, while math was administered in grades 5-10, and science was administered in grade 8. The current study linked data from spring 2002 and spring 2005 CSAP administrations to MAP, which was also administered in the 2002 and 2005 school years and has an unchanging scale.

To estimate the difficulty of Colorado’s proficiency cut scores, we linked data from Colorado’s reading and math tests from a group of elementary and middle schools to the NWEA assessment. (A “proficiency cut score” is the test score that a student must achieve in order to be considered proficient.) This was done by analyzing a group of schools in which almost all students had taken 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 Colorado's Definitions of Proficiency in Reading and Math?

One way to assess 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 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. How do we know that solving differential equations is more difficult than adding fractions? Because if you ask a group of tenth graders to do both tasks, far more will be able to add fractions than will be able to solve differential equations.

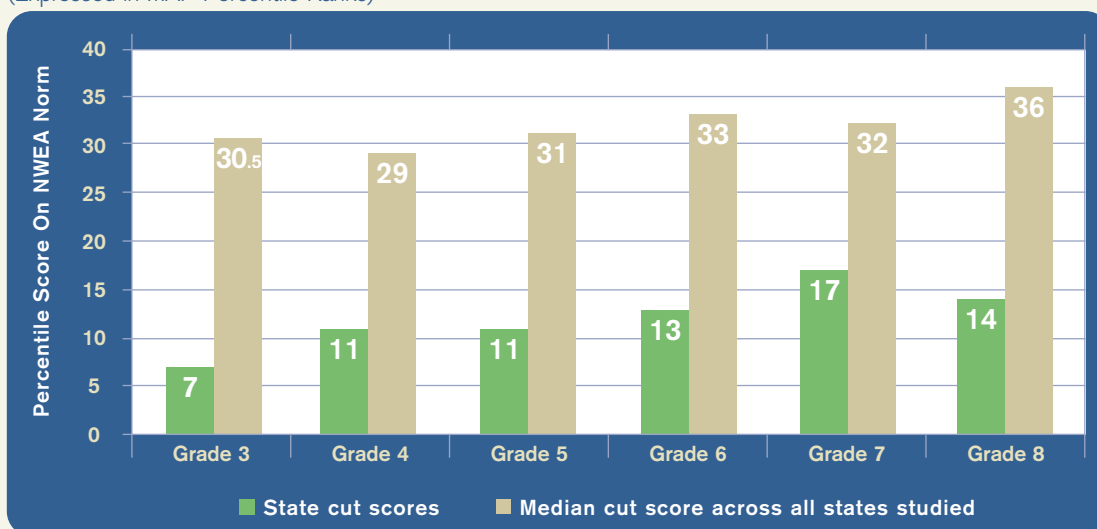
Applying that approach to this task, we evaluated the difficulty of Colorado's NCLB proficiency cut scores by estimating the proportion of students in NWEA's norm group who would perform above the Colorado cut score on a test of equivalent difficulty. The following two figures show the difficulty of Colorado's proficiency cut scores for **reading** (Figure 1) and **mathematics** (Figure 2) in 2005 in relation to the median cut score for all the states in the study. The NCLB proficiency cut scores for reading in Colorado ranged between

the 7th and 17th percentiles for the norm group, with the seventh grade being most challenging. In mathematics, the NCLB proficiency cut scores ranged between the 6th and 25th percentiles for the norm group with the eighth grade being most challenging.

Colorado's NCLB cut scores in both reading and mathematics are well below average in difficulty among the states studied. Note, too, that in middle school, Colorado's cut scores for reading are lower than those for mathematics. Thus, reported differences in achievement on the CSAP between reading and mathematics might be more a product of differences in cut scores than in actual student achievement. In other words, Colorado students might be performing worse in reading and better in mathematics than is apparent by just looking at the percentage of students passing state tests in those subjects.

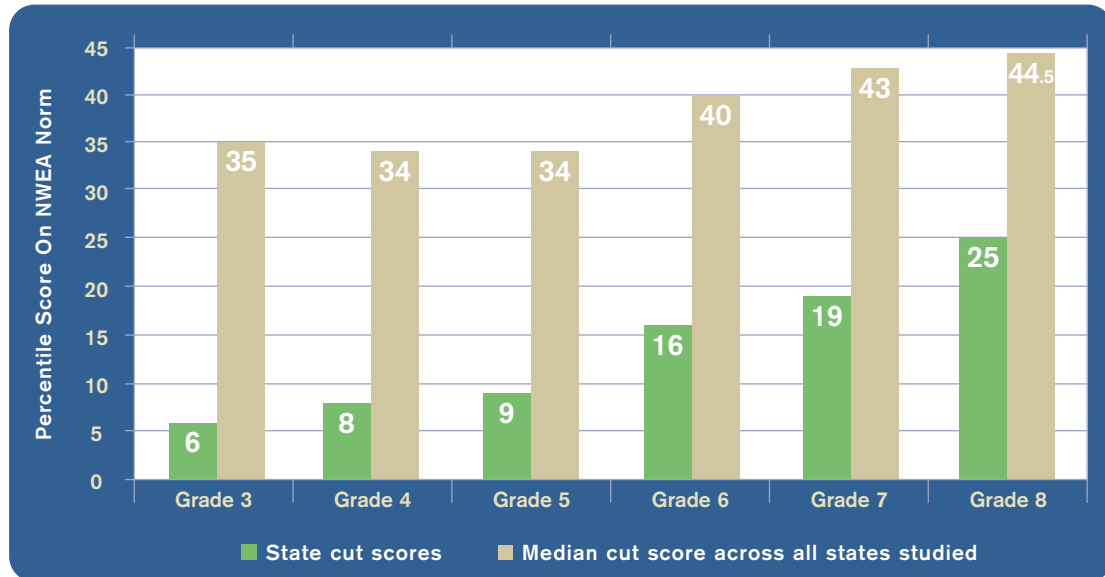
Another way of assessing difficulty is to evaluate how Colorado's NCLB proficiency cut scores rank relative to other states. Table 1 shows that the Colorado cut scores generally rank among the lowest of the 26 states studied for this report. In third and fifth grade reading, Colorado's cut scores rank; the state is second-to-last in fourth, sixth, and seventh grade reading and fifth grade mathematics.

Figure 1 – Estimate of Colorado Reading Cut Scores in Relation to the 25 Other States Studied, 2006 (Expressed in MAP Percentile Ranks)



Note: This figure compares reading test cut scores (“proficiency passing scores”) as percentiles of the NWEA norm. These percentiles are compared with the media cut scores of all 26 states reviewed in this study. Colorado's cut scores are consistently 15 to 23.5 percentile points below the median in grades 3 to 8.

Figure 2 – Colorado Mathematics Cut Scores in Relation to the 25 Other States Studied, 2006
(as Expressed in MAP Percentile Ranks)



Note: Colorado's math test cut scores are shown as percentiles of the NWEA norm and compared with the median cut scores of other states reviewed in this study. Colorado's cut scores are 29 to 19.5 percentiles below the median across grades 3-8.

Table 1 – Colorado 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	Grade 8
Reading	26	25	26	25	25	23
Mathematics	24	24	25	24	23	19

Note: This table ranks Colorado's cut scores relative to the cut scores of the other 25 states in the study. In third-grade math, Colorado ranks 24 out of 26, meaning that 23 states' cut scores were higher, while only two were lower. Colorado either places last or second-to-last in half the categories.

Part 2: Changes in Cut Scores over Time

In order to measure their consistency over time, Colorado's proficiency cut scores were mapped to their equivalent scores on NWEA's MAP assessment for the 2002 and 2005 school years. Cut score estimates for both years were available for grades 3-8 for reading, and grades 5-8 for mathematics.

States may periodically re-adjust the cut scores they use to define proficiency in reading and mathematics, or update the tests used to evaluate 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.

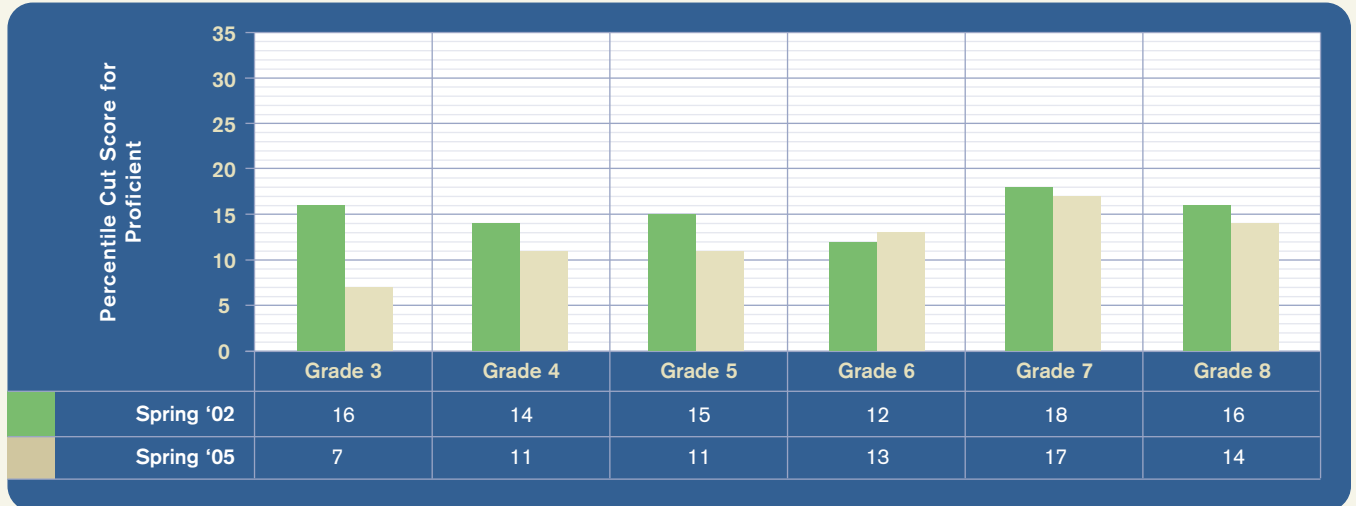
Is it possible, then, to compare the proficiency scores between the earlier era of Colorado's tests and today's? Yes. Assume once again that we're judging a group of fourth graders on their high-jump ability and that we measure this by finding how many in that group can successfully clear a three-foot bar. Now assume that we change the measure and set a new height. Perhaps students must now clear a bar set at 1 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 1 meter than 3 feet, because we know the relationship between the measures. The same principle applies here. CSAP in 2002 and in 2005 can both be linked to the MAP, which has remained consistent over time. Just as one can convert three feet to a meter [see comments in CA write up] 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 CSAP in 2002 and 2005 on the MAP scale and ascertain whether the test may have changed in difficulty.

Colorado's **reading** results indicate a decline in estimated proficiency cut scores in grades three, four, and five over this three-year period (see Figure 3). Consequently, one would expect the third grade students' reading proficiency rates in 2005 to be 9 percent higher than in 2002, even if actual pupil student performance remained the same. One would expect similar increases in the reading proficiency rates for fourth and fifth grades of 3 and 4 percent, respectively, if actual student performance remained the same.

Colorado's **mathematics** results indicate a decrease in estimated proficiency cut scores in grades 5, 7, and 8 (Figure 4). These changes would likely yield increased math proficiency rates in these grades of 4, 5, and 6 percent, respectively, even if pupil performance remained the same.

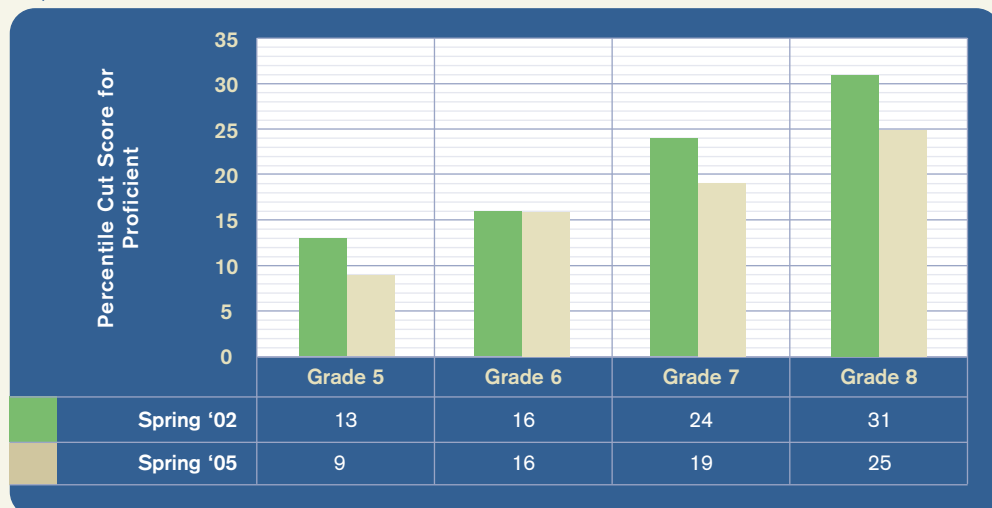
Thus, one could fairly say that Colorado's fifth grade tests in both reading and mathematics were easier to pass in 2005 than in 2002. Similarly, the reading tests for third and fourth graders were easier, as were the mathematics tests for seventh and eighth graders. As a result, some apparent improvements in Colorado students' proficiency rates during this period may not be entirely a product of improved achievement.

Figure 3 – Estimated Differences in Colorado's Proficiency Cut Scores in Reading, 2002-2005 (Expressed in MAP Percentile Ranks).



Note: This graphic shows how the difficulty of achieving proficiency in reading has changed. For example, third grade students in 2002 had to score at the 16th percentile in order to be considered proficient, while in 2005 third graders had only to score at the 7th percentile.

Figure 4 – Estimated Differences in Colorado's Proficiency Cut Scores in Mathematics, 2002-2005 (Expressed in MAP Percentile Ranks).



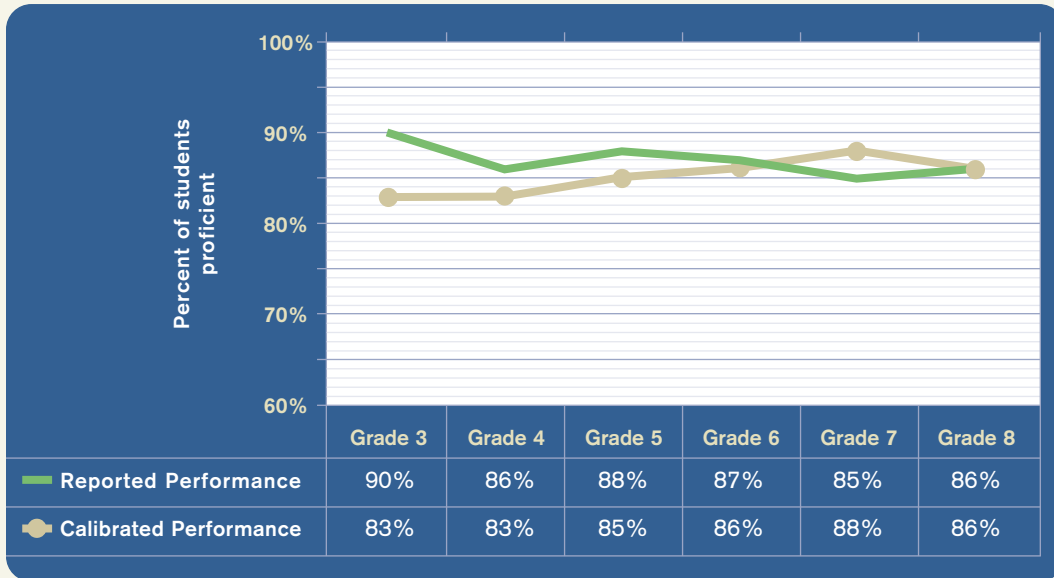
Note: This graphic shows how the difficulty of achieving proficiency in math has changed. For example, fifth grade students in 2002 had to score at the 13th percentile in order to be considered proficient, while by 2005 fifth graders only had to score at the 9th percentile.

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.

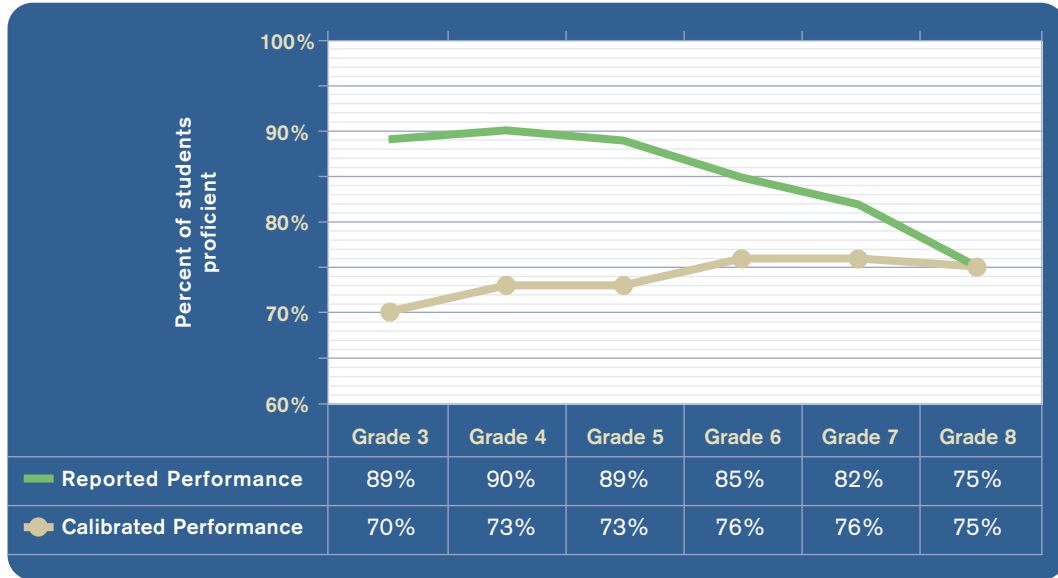
Examining Colorado’s cut scores, we find that they are not well calibrated across grades. Figures 1 and 2 showed that Colorado’s upper-grade cut scores in reading and mathematics in 2005 were more challenging than in the lower grades. The two figures that follow show Colorado’s reported performance on its state test in reading (Figure 5) and mathematics (Figure 6) compared with the rates of proficiency that would be achieved if the cut scores were calibrated to grade 8. When differences in grade-to-grade difficulty of the cut scores are removed, student performance is more consistent at all grades, particularly in mathematics. This would lead to the conclusion that the higher rates of mathematics proficiency that the state has reported for younger students are somewhat misleading.

Figure 5 – Colorado Reading Performance Relative to a Calibrated Standard, 2005



Note: This graphic shows, for example, that if Colorado’s grade 3 reading standard were as difficult as its grade 8 standard, 83 percent of third graders would achieve the proficient level, rather than 90 percent, as was reported by the state.

Figure 6 – Colorado Mathematics Performance Relative to a Calibrated Standard, 2005



Note: This graphic shows, for example, that if Colorado's grade 3 mathematics standard were set at the same level of difficulty as its grade 8 standard, 70 percent of third graders would achieve the proficient level, rather than 89 percent, as was reported by the state.

Policy Implications

When setting its cut scores for what constitutes student proficiency in reading and mathematics for NCLB purposes, Colorado aimed low, at least compared to 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 Colorado's standards to be toward the bottom of the distribution of all states studied.) Colorado's low cut scores have declined even further in recent years in several grades.

As a result, Colorado's expectations are not calibrated across all grades; students who are proficient in third grade are not necessarily on track to be proficient by the eighth grade. In addition to better calibrating the state's cut scores, Colorado policymakers might consider raising those scores across the board so that parents and educators can be assured that scoring at the NCLB proficient level means that students are truly prepared for success later in their educational careers.