

Arizona



Introduction

This study linked data from the 2002 and 2005 administrations of Arizona’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 Arizona’s definitions of “proficiency” in reading and mathematics are relatively consistent with the standards set by the other 25 states in this study. In other words, Arizona’s tests are about average in terms of difficulty.

Yet the level of difficulty of Arizona’s tests generally declined from 2002 to 2005—the No Child Left Behind era—quite significantly in some grades. This is not a surprise, as the Arizona State Board of Education adopted a new scale for both the reading and math tests for the 2004-05 academic school year, and publicly reported lowering the cut scores on those tests.

Not well known, however, is that the state’s proficiency cut scores are now relatively lower for third-grade students than for eighth-grade pupils (taking into account the obvious differences in subject content and children’s development). Plus, as is true for the majority of states studied, Arizona’s cut scores for reading are lower than those for mathematics. Arizona policymakers might consider adjusting their 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 teacher and student performance across these domains.

What We Studied: Arizona’s Instrument to Measure Standards (AIMS)

Arizona currently uses a spring assessment called the Arizona Instrument to Measure Standards – Dual Purposes Assessment (AIMS – DPA) as part of its state assessment program. This tests reading, writing, and mathematics in elementary and middle school students in grades 3 through 8. Students in grade 10 take the AIMS HS (High School) and may continue to take that test twice per year during grades 11 and 12 until they have met or exceeded the standards for proficiency in writing, reading, and mathematics. The current study

analyzed reading and math results from a group of elementary and middle schools in which almost all students took both the state’s assessment and MAP, using the spring 2002 and spring 2005 administrations of the two tests. (The methodology section of this report explains how performance on these two tests was compared.) These linked results were then used to estimate the scores on NWEA’s scale that would be equivalent to the proficiency cut scores for each grade and subject on the Arizona State Assessment. (A “proficiency cut score” is the score a student must achieve in order to be considered “proficient.”)

Part 1: How Difficult are Arizona’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 jump 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 jump 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.

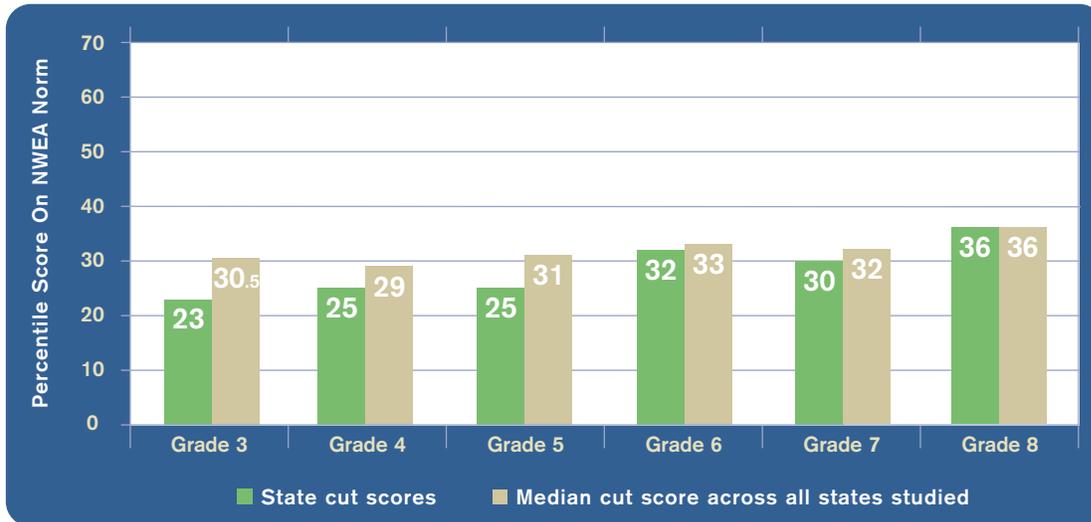
Applying that approach to this task, we evaluated the difficulty of Arizona’s proficiency standards by estimating the proportion of students in NWEA’s norm group who would perform above the Arizona standard on a test of equivalent difficulty. The following two figures show the difficulty of Arizona’s 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 proficiency cut scores for **reading** in Arizona ranged between the 23rd and 36th percentiles for the norm group, with the eighth-grade cut score being most challenging. In **mathematics**, the proficiency cut scores ranged between the 28th and 42nd percentiles with eighth grade again being most challenging.

For most grade levels, Arizona’s cut scores in both reading and mathematics are slightly below average in difficulty among the states studied. Exceptions include eighth-grade reading and sixth-grade math, which are at the median proficiency cut scores among the states examined.

Note, too, that Arizona’s cut scores for reading are lower than those 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, Arizona 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.

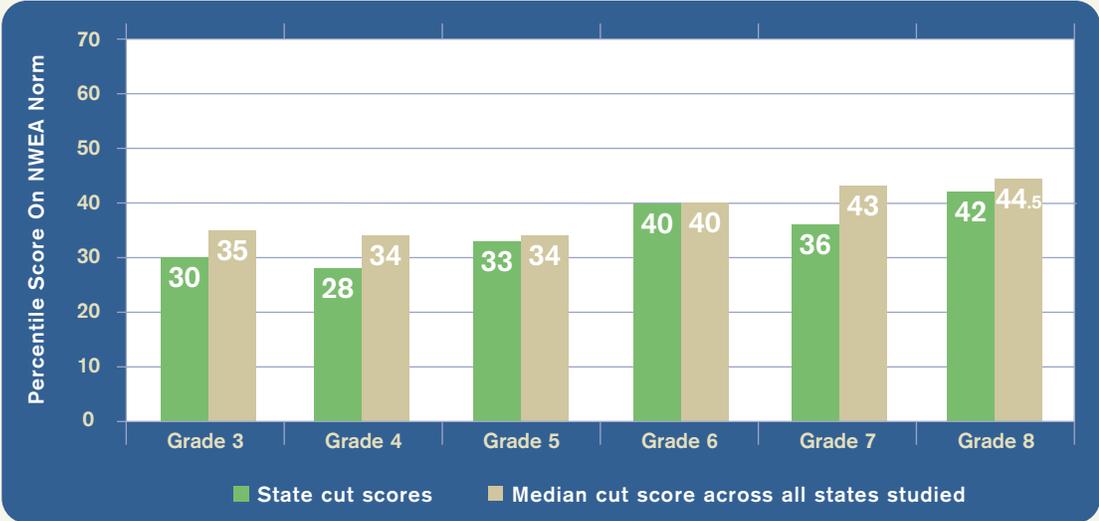
Another way of assessing difficulty is to evaluate how Arizona’s proficiency cut scores rank relative to other states. Table 1 shows that the Arizona cut scores generally rank in the mid- or bottom third among the 26 states studied for this report. Arizona’s third- and fifth-grade reading cut scores are particularly low, besting those of only seven other states in the study. On the other hand, Arizona ranks relatively high in eighth-grade math and reading and in third- and sixth-grade math.

Figure 1 – Arizona Reading Cut Scores in Relation to All 26 States Studied, 2005
(Expressed in 2005 MAP Percentiles)



Note: This figure compares reading test cut scores (“proficiency passing scores”) as percentiles of the 2005 NWEA norm. These percentiles are compared with the median cut scores of other states reviewed in this study. Only in eighth grade does Arizona’s cut score reach the median. Grades 3-7 scores are 1 to 7.5 percentile points below the median.

Figure 2 – Arizona Mathematics Cut Scores in Relation to All 26 States Studied, 2005
(Expressed in MAP Percentiles).



Note: Arizona’s math test cut scores are shown as percentiles of the 2005 NWEA norm and compared with the median cut scores of other states reviewed in this study. Only in sixth grade does Arizona’s cut score reach the median; in third grade, it lagged by 5 percentile points and in seventh grade by 7 points.

Table 1 – Arizona Rank for Proficiency Cut Scores Among 26 States in Reading and Mathematics, 2005 or 2006

		Ranking (Out of 26 States)					
		Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Reading		19	17	19	14	18	9
Mathematics		14	19	16	12	18	12

Note: This table ranks Arizona’s cut scores relative to the cut scores of the other 25 states in the study, where 1 is highest and 26 is lowest.

Part 2: Changes in Cut Scores over Time

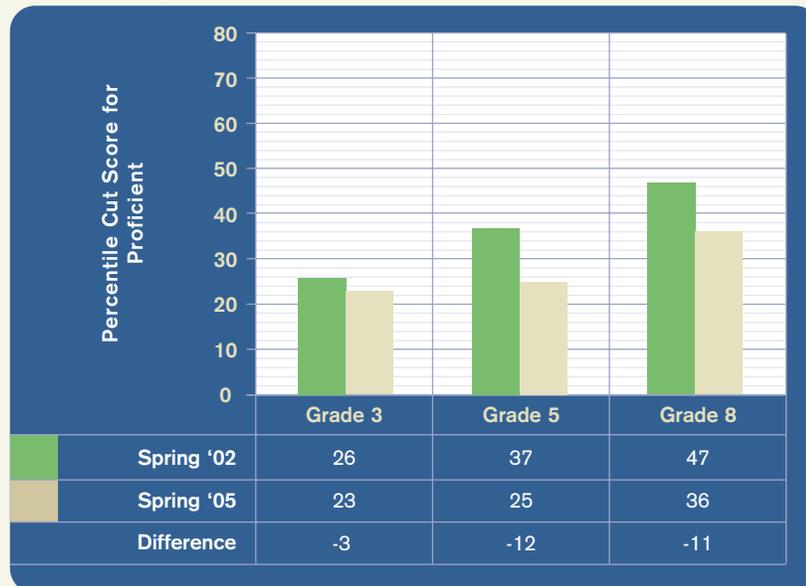
In order to measure their consistency, Arizona’s proficiency cut scores were mapped to their equivalent scores on NWEA’s MAP assessment for the 2002 and 2005 school years. Cut score estimates for both years were available for grades 3, 5, and 8.

States may periodically re-adjust the cut scores they use to define proficiency in reading and math, or may update the tests used to test 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 occurred in Arizona, in the 2004-05 academic year, when the State Board of Education adopted new scales and publicly lowered cut scores both for the reading and math tests.

Is it possible, then, to compare the proficiency scores between earlier administrations of Arizona’s tests and today’s? Yes.

Assume that we’re judging a group of fourth graders on their high-jump prowess. We can 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 to judge proficiency. 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 measure or scale used by the AIMS in 2002 and in 2005 can both be linked to the scale that was used to report MAP, which has remained consistent over time. Just as one can compare one meter to three feet 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 AIMS in 2002 and 2005 on the MAP scale and ascertain whether the test may have changed in difficulty—and whether those changes are consistent with what the state reported to the public.

Figure 3 – Estimated Change in Arizona’s Proficiency Cut Scores in Reading, 2002-2005 (Expressed in MAP Percentiles).



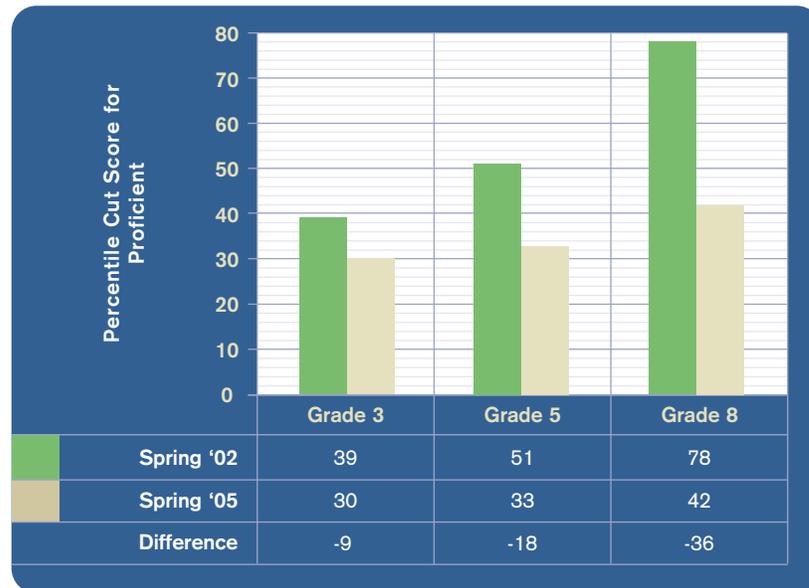
Note: This graphic shows how the difficulty of achieving proficiency in reading has changed. For example, fifth-grade students in 2002 had to score at the 37th percentile of the NWEA norm group in order to be considered proficient, while in 2005 fifth graders only had to score at the 25th percentile of the NWEA norm group to achieve proficiency. The change in grade 3 was within the margin of error (in other words, it is too small to be considered substantive).

Arizona's estimated **reading** cut scores decreased in grades 5 and 8 over this three-year period, though no substantive change was found in grade 3 (see Figure 3). Consequently, even though student performance on MAP did not change, one would expect the fifth- and eighth-grade reading proficiency rates in 2005 to be 12 percent and 11 percent higher than in 2002, respectively. (Arizona reported a 12-point gain for fifth graders and an 11-point gain for eighth graders over this period.)

Thus, one could fairly say that Arizona's third-grade reading test was about as difficult to pass in 2005 as in 2002, while the other tests were easier to pass for the other grades examined. As a result, some apparent improvements in the Arizona students' proficiency rates during this time may not be entirely a product of improved achievement.

Arizona's estimated **mathematics** cut scores indicate a dramatic decrease in proficiency cut scores in grades 3, 5, and 8 over this three-year period (see Figure 4). Consequently, even if student performance stayed the same on an equivalent test like NWEA's MAP assessment, the changes in grades 3, 5, and 8 would likely yield increased math proficiency rates of 9, 18, and 36 percent, respectively. Arizona reported a 15-point gain for third graders, a 25-point gain for fifth graders, and a 42-point gain for eighth graders over this period.)

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



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 51st percentile of the NWEA norm group in order to be considered proficient, while in 2005 fifth graders only had to score at the 33rd percentile of the NWEA norm group to achieve proficiency.

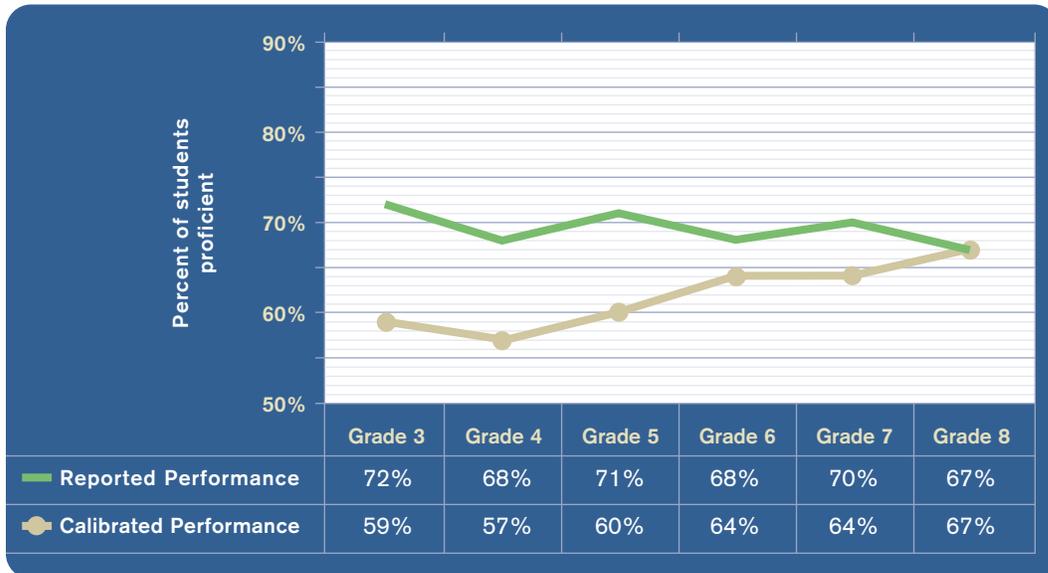
Part 3: Calibration across Grades

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.

Examining Arizona’s cut scores, we find that they are not well calibrated across grades. Figures 1 and 2 showed that Arizona’s upper grade cut scores in reading and mathematics in 2005

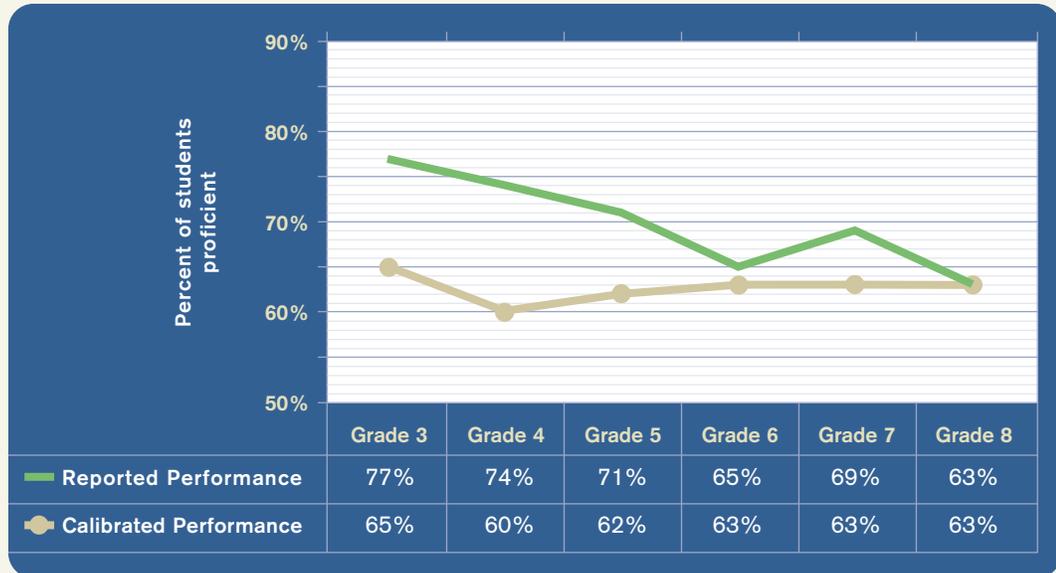
were more challenging than the cut scores in the lower grades. The two figures that follow show Arizona’s reported performance on its state test in reading (Figure 5) and mathematics (Figure 6) compared with the rates of proficiency that would be achieved if the cut scores were all calibrated to the grade 8 standard. When differences in grade-to-grade difficulty of the cut scores are removed, student performance in mathematics is more consistent at all grades. This would lead to the conclusion that the higher rates of mathematics proficiency that the state has reported for elementary school students are somewhat misleading. It also becomes clear that actual reading performance is lower at the elementary level than in middle school—while the state’s published passing rates appear to indicate relatively consistent performance from grades 3 to 8.

Figure 5 – Arizona Reading Performance as Reported and as Calibrated to the Grade 8 Standard, 2005



Note: This graphic shows, for example, that if Arizona’s grade-3 reading standard were as difficult as its grade-8 standard, 59 percent of third graders would achieve the proficient level, rather than 72 percent, as reported by the state.

Figure 6 – Arizona Mathematics Performance as Reported and as Calibrated to the Grade-8 Standard, 2005



Note: This graphic shows, for example, that if Arizona's grade-3 mathematics cut score were as difficult as its grade-8 standard, 65 percent of third graders would achieve the proficient level, rather than 77 percent, as was reported by the state.

Policy Implications

Arizona proficiency cut scores stand in the middle to bottom third of the pack when 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 Arizona's standards to be in the bottom half to the bottom third of the distribution of all states studied. Arizona's cut scores, which weren't particularly difficult in most grades in 2002, have over the past several years been adjusted—making them generally less challenging (and, in some grades,

significantly less challenging). Arizona's expectations are not well calibrated across grades, particularly for mathematics. Students who are proficient in third grade are not necessarily on track to be proficient by the eighth grade. Arizona policymakers might consider adjusting their proficiency cut scores across grades so that parents and schools can be assured that young students scoring at the proficient level are truly prepared for success later in their educational careers.