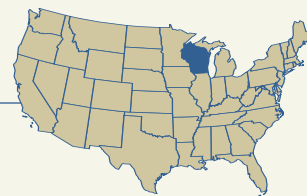


Wisconsin



Introduction

This study linked data from the 2003 and 2005 administrations of Wisconsin’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 Wisconsin’s definitions of proficiency in reading and mathematics are relatively less difficult than the cut scores set by other states. In other words, Wisconsin’s tests are below average in terms of difficulty.

The level of difficulty of these cut scores decreased in some grades from 2003 to 2005—the No Child Left Behind era. For example, Wisconsin’s eighth-grade tests for reading and mathematics were easier in 2005 than in 2003.

Wisconsin’s cut scores in mathematics are now more difficult in the lower grades than in the higher grades (taking into account the obvious differences in subject content and children’s development). Consequently, the proportion of younger students who are on track to meet the cut scores at the later grades may be underestimated. Wisconsin policy-makers 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.

What We Studied: Wisconsin Knowledge and Concepts Examinations - Criterion Referenced Test (WKCE-CRT)

Wisconsin currently uses a fall assessment called the Wisconsin Knowledge and Concepts Examinations - Criterion Referenced Test (WKCE-CRT), which tests reading, language applications, mathematics, science, and social studies in students in grades 3 through 8 and 10, as expected by NCLB. Fall 2005 was the first year the criterion-referenced test was used. It replaced the Wisconsin Knowledge and Concepts Examinations (WKCE), an augmented version of the nationally-normed Terra Nova test, first used in fall 2002 to test reading, language arts, mathematics, science, and social studies in grades 4, 8, and 10. The current study linked reading and math data from fall 2003 WKCE administrations and fall 2005 WKCE-CRT administrations to a common scale also administered in the 2003-4 and 2005-6 school years.

To determine the difficulty of Wisconsin’s proficiency cut scores, we linked data from state tests 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 Wisconsin’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 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.

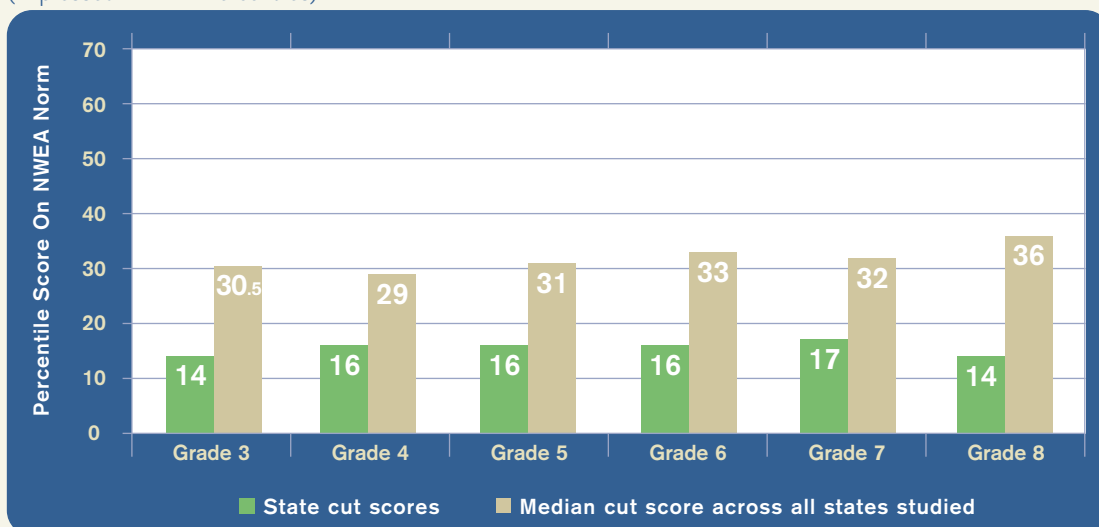
Applying that approach to this assignment, we evaluated the difficulty of Wisconsin’s proficiency cut scores by estimating the proportion of students in NWEA’s norm group who would perform above the Wisconsin cut score on a test of equivalent difficulty. The following two figures show the difficulty of Wisconsin’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 Wisconsin ranged between the 14th and 17th percentiles for the norm group, with the seventh-grade cut score being most challenging. In **mathematics**, the proficiency cut scores ranged between the 21st and 29th percentiles with the third and fourth grade cut scores being most challenging.

For all grade levels, Wisconsin’s cut scores in both reading and mathematics are lower than the median cut scores of the other states in the study, and far below the capabilities of the average student of that grade within the NWEA norm group.

Note, too, that Wisconsin’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, Wisconsin students may 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.

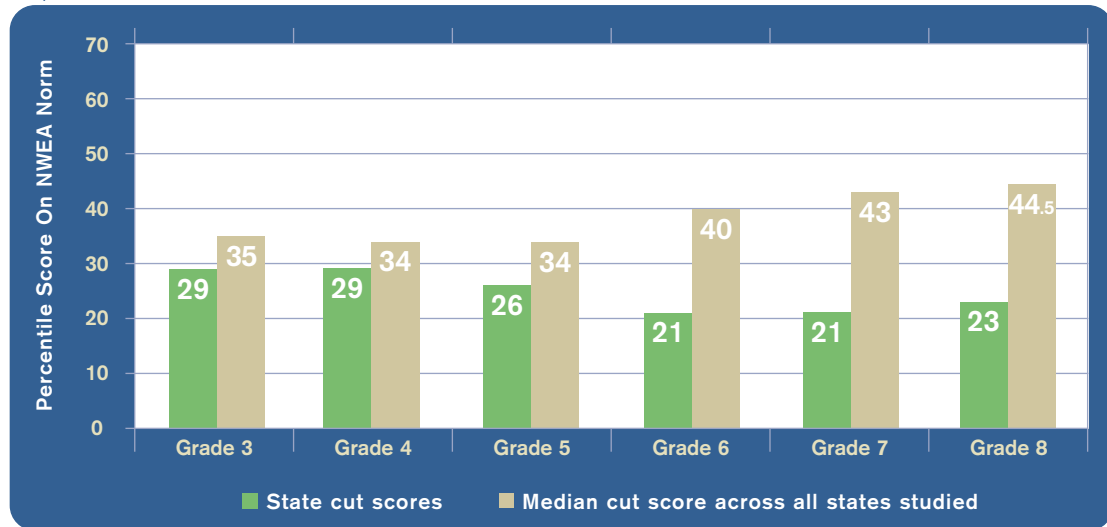
Another way of assessing difficulty is to observe how Wisconsin’s proficiency cut scores rank relative to other states. Table 1 shows that the state’s cut scores generally rank among the lowest of the 26 states studied for this report, in terms of difficulty.

Figure 1 – Wisconsin Reading Cut Scores in Relation to All 26 States Studied, 2005 (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. Wisconsin’s scores range from 13 to 22 percentile points behind the median.

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



Note: This figure compares reading test cut scores as percentiles of the NWEA norm. These percentiles are compared with the median cut score of all 26 states reviewed in this study. Wisconsin's scores range from 5 to 22 percentile points behind the median.

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

Ranking (Out of 26 States)						
	Grade 3	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Reading	23	24	23	24	25	23
Mathematics	19	18	18	23	23	21

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

Part 2: Changes in Cut Scores over Time

In order to measure their consistency, Wisconsin's proficiency cut scores were mapped to their equivalent scores on NWEA's MAP assessment for the 2003-4 and 2005-6 school years during the same season. Cut score estimates for reading and mathematics were available for both years in grades 4 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 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 the case for Wisconsin which, as explained above, adopted a new test for 2005.

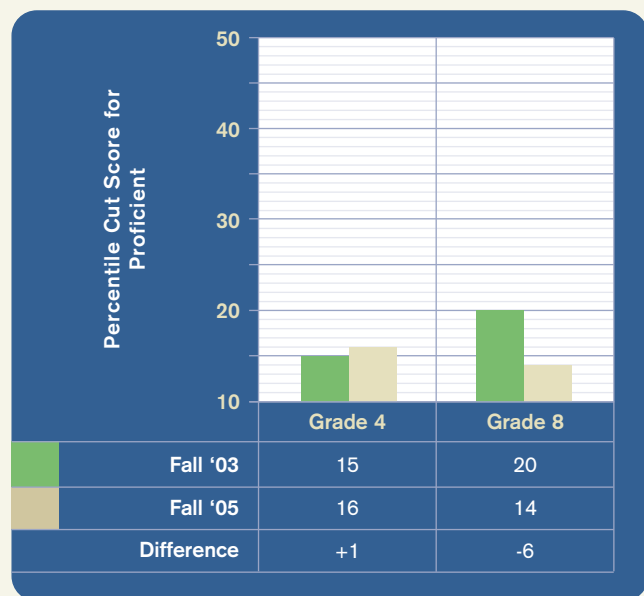
Is it possible, then, to compare the proficiency scores between the earlier and later administrations of Wisconsin tests? Yes. Assume that we're judging a group of fifth graders on their high-jump prowess and that we gauge this by finding how many in that group can successfully clear a three-foot bar. 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 WKCE in 2003 and the WKCE-CRT in 2005 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 use the MAP scale to estimate whether the WKCE-CRT in 2005 is easier or more difficult than the prior test and proficiency cut scores that were in place.

In **reading**, Wisconsin showed a moderate decrease in the estimated eighth-grade reading cut score estimate over this two-year period, but essentially no change in the fourth-grade reading cut score (see Figure 3). Consequently, even if student performance stayed the same on an equivalent test like NWEA's MAP assessment, one would expect the eighth-grade reading proficiency rate in 2005 to be 6 percent higher than in 2003. (In fact, Wisconsin reported a 6-point gain for eighth graders over this period.)

Wisconsin's **mathematics** results show the same pattern, with a moderate decrease in the estimated eighth-grade cut score and essentially no change in the fourth-grade cut score. Consequently, even if student performance stayed the same on an equivalent test like NWEA's MAP assessment, one would expect the eighth-grade math proficiency rate in 2005 to be about 11 percent higher than in 2003, even if actual student performance remained the same. (Wisconsin Wisconsin reported a 9-point gain for eighth graders over this period.)

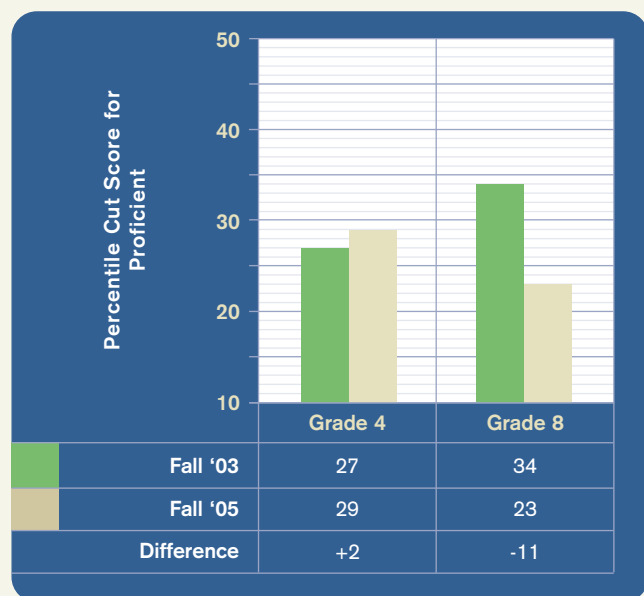
Thus, one could fairly say that Wisconsin's fourth-grade tests in both reading and mathematics stayed about the same from 2003 to 2005, while the eighth-grade tests became easier to pass. As a result, improvements in state-reported proficiency rates during this period may not be entirely a product of improved achievement.

Figure 3 – Estimated Differences in Wisconsin's Proficiency Cut Scores in Reading, 2003-2005 (Expressed in MAP Percentile Ranks)



Note: This graphic shows how the difficulty of achieving proficiency in reading has changed. For example, eighth-grade students in 2003 had to score at the 20th percentile nationally in order to be considered proficient, while by 2005 eighth graders had to score at the 14th percentile to achieve proficiency. The change in grade 4 was within the margin of error (in other words, too small to be considered substantive)

Figure 4 – Estimated Differences in Wisconsin's Proficiency Cut Scores in Mathematics, 2003-2005 (Expressed in MAP Percentiles)



Note: This graphic shows how the difficulty of achieving proficiency in math has changed. For example, eighth-grade students in 2003 had to score at the 34th percentile nationally in order to be considered proficient, while in 2005 eighth graders only had to score at the 23rd percentile of the NWEA norm group to achieve proficiency. The change in grade 4 was within the margin of error (in other words, too small to be considered substantive).

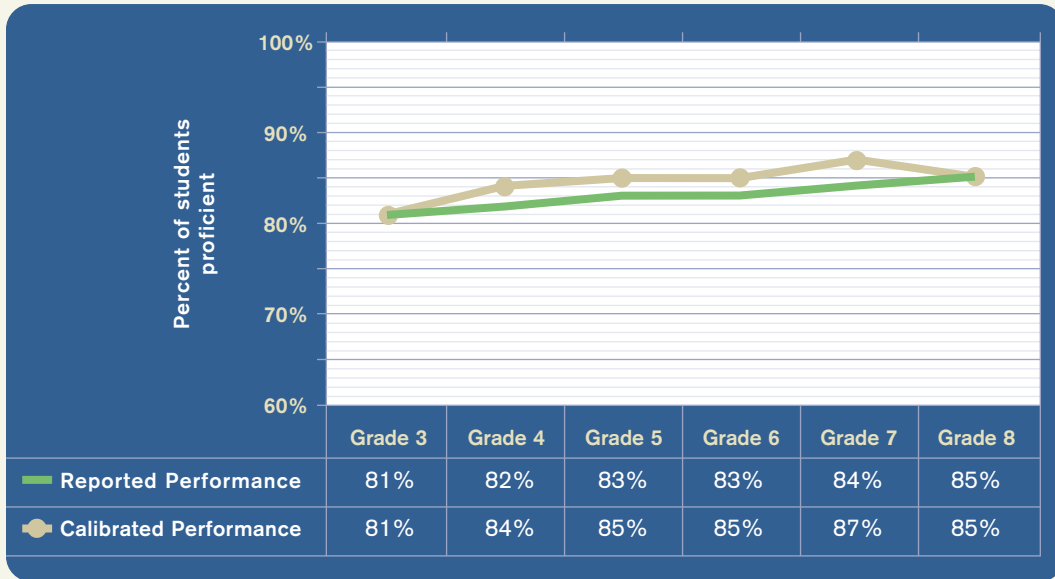
Part 3: Calibration across Grades

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 Wisconsin’s cut scores, we see in Figures 1 and 2 showed that the state’s reading cut scores across grades 2

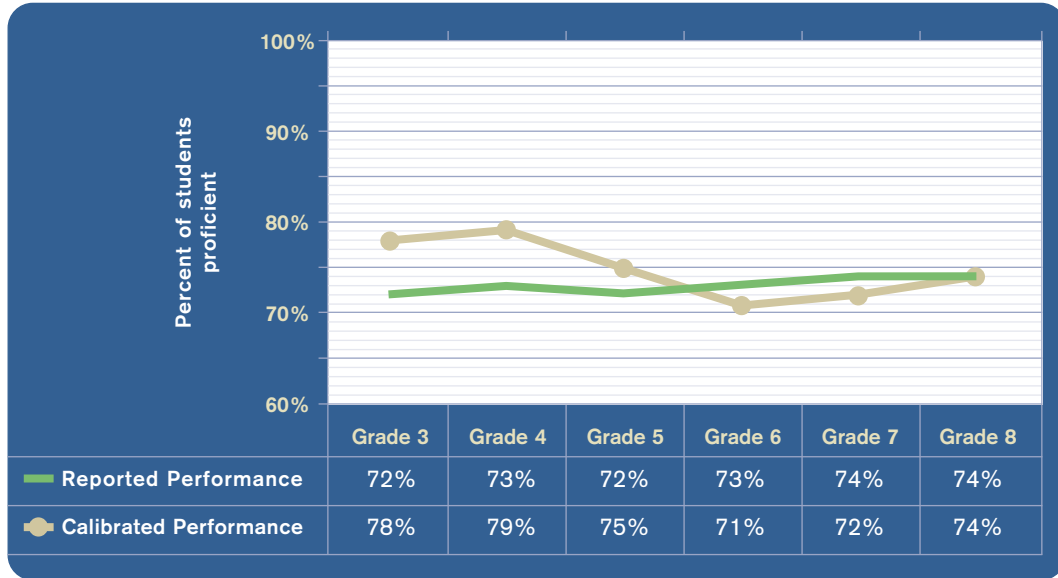
through 8 were fairly well calibrated, while the math cut scores in the lower grades were slightly more difficult than in the upper grades. These are reflected in Figures 5 and 6, which show how Wisconsin’s reported performance on the state test in reading (Figure 5) and mathematics (Figure 6) compared with the rate of proficiency that would be achieved if the cut scores were all calibrated to the eighth-grade standard. In Figure 5, the differences between the observed proficiency rates and those that would be expected with calibrated cut scores are quite small. In Figure 6, however, we see that the uncalibrated standards at the earlier grades slightly underestimate the proportions of third and fourth graders who are on track to eventually demonstrate proficiency at the later grades.

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



Note: This graphic shows that, for example, that if Wisconsin’s grade-5 reading standard was at the same difficulty level as its grade-8 standard, 85 percent of fifth graders would achieve the proficient level, rather than 83 percent, as was reported by the state.

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



Note: This graphic shows, for example, that if Wisconsin's grade-3 mathematics cut score was set at the same difficulty level as its grade-8 cut score, 78 percent of third graders would achieve the proficient level, rather than 72 percent, as was reported by the state.

Policy Implications

When setting its cut scores for what students must know and be able to do to be considered proficient in reading and math, Wisconsin is low, 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 Wisconsin to have some of the lowest standards of all states, at least in reading. In the past several years, the difficulty of the grade-8 cut scores has declined somewhat. As a result,

Wisconsin's expectations for mathematics are not smoothly calibrated across grades, so Wisconsin currently underestimates the proportion of students in the younger grades who are on track to meet the (low) eighth-grade mathematics cut scores. Wisconsin policymakers might consider adjusting their cut scores across grades so that proficiency at the earlier grades more accurately predicts proficiency at the later grades.