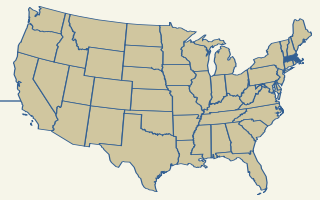


# Massachusetts



## Introduction

This study linked data from the 2006 administration of Massachusetts'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 Massachusetts's definitions of proficiency in reading and math are relatively high compared with the standards set by the other 25 states in the study. In other words, Massachusetts's tests are well above average in terms of difficulty.

However, unlike most of the states in this study, Massachusetts's proficiency cut scores for reading and English/language arts are less difficult in the later grades than in the earlier grades. Therefore, reported results for younger students may underestimate the number who are on track to be proficient in eighth-grade reading. Massachusetts policy-makers might consider adjusting their reading 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: Massachusetts Comprehensive Assessment System (MCAS)

Massachusetts currently uses the Massachusetts Comprehensive Assessment System (MCAS), which tests mathematics and reading/ELA in grades 3 to 8 and grade 10, and high school science and technology in grades 9 and 10. The current study linked reading and math data from spring 2006 MCAS administrations to a common scale also administered in the 2006 school year.

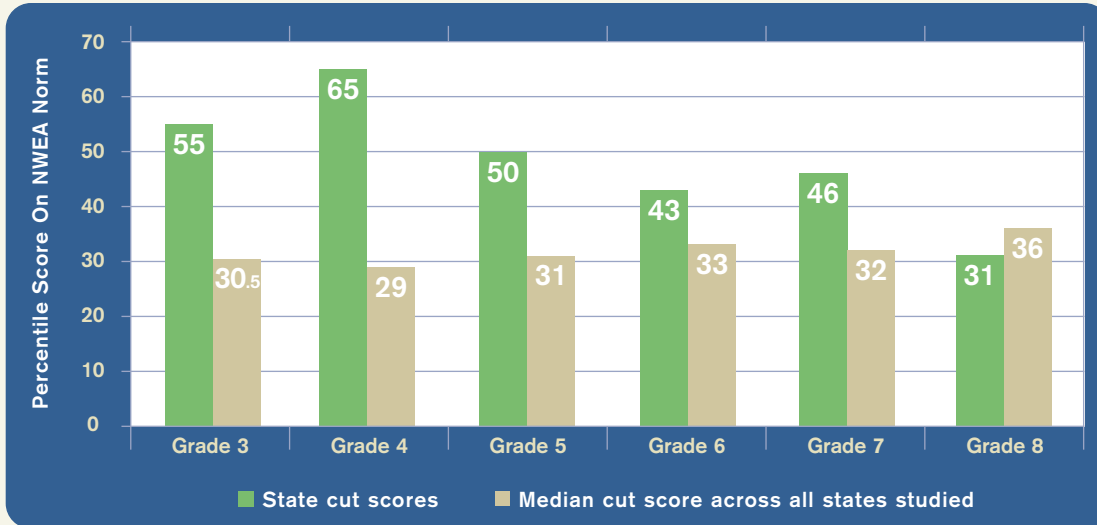
To determine the difficulty of Massachusetts's proficiency cut scores, we linked data from Massachusetts's tests to the NWEA assessment. (A "proficiency cut score" is the score a student must achieve in order to be considered proficient.) This was done by analyzing a group of elementary and middle schools in which almost all students took both the state's assessment and the NWEA test. (For more details on how this was done, please see the methodology section of this report.)

### Part 1: How Difficult are Massachusetts'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 jump bar is easy to leap? 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? We know because only one (or perhaps none) of those same 100 individuals would successfully meet that level of 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.

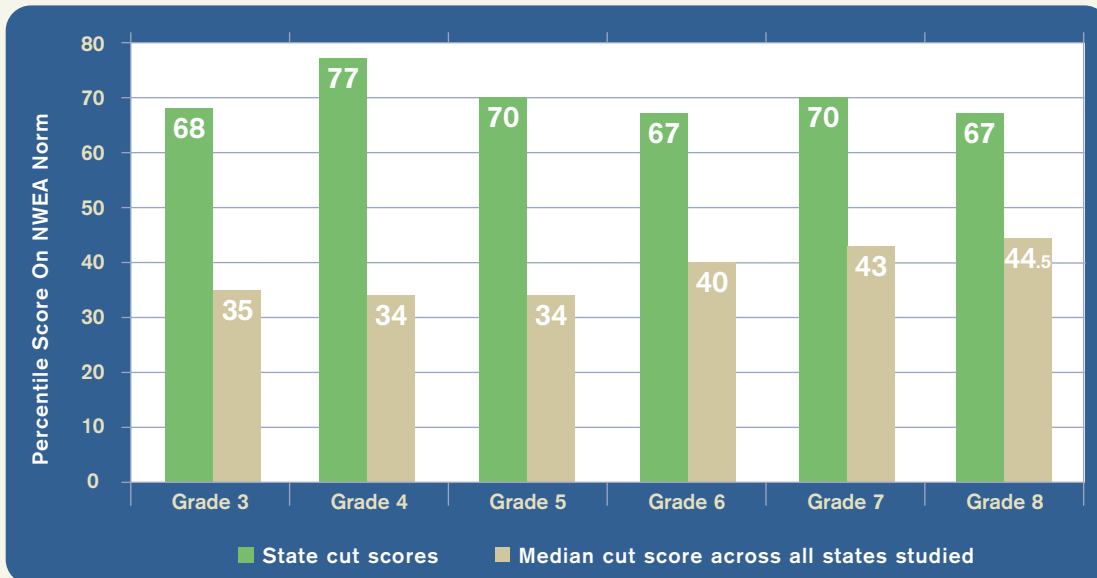
Applying the concept to this task, we evaluated the difficulty of the Massachusetts proficiency cut scores by estimating the proportion of students in NWEA's norm group who would perform above the cut score on a test of equivalent difficulty. The following two figures show the difficulty of Massachusetts's proficiency cut scores for reading (Figure 1) and mathematics (Figure 2) in 2006 in relation to the other states in the study, and compared with the NWEA norm group. The proficiency cut scores for **reading** in Massachusetts ranged between the 31st and 65th percentiles in the norm group, with the fourth-grade cut score being most challenging. In **mathematics**, the cut scores ranged between the 67th and 77th percentiles with fourth grade again being most challenging.

Figure 1 – Massachusetts 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 scores of all 26 states reviewed in this study. Massachusetts is consistently above average—as much as 36 percentile points above the median in fourth grade—except for eighth grade, when it falls 5 percentiles below the median.

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



**Note:** Massachusetts math test cut scores are shown as percentiles of the NWEA norm and compared with the median cut scores of all 26 states reviewed in this study. The math cut scores are consistently 22.5 to 43 percentile points above the median.

Massachusetts's reading cut scores are consistently above the median difficulty of the 26 states that we examined, except in grade 8. Massachusetts's mathematics cut scores are above the median in every grade. Note, too, that the reading cut scores are consistently less difficult than the corresponding mathematics cut scores. Thus, reported differences in achievement on the MCAS between reading and mathematics might be more a product of differences in cut scores than in actual student achievement. In other words, Massachusetts students

may be performing worse in reading 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 Massachusetts's proficiency cut scores rank relative to other states. Table 1 shows that the Massachusetts cut scores rank at the very top in difficulty among the 26 states in this study, except in eighth grade reading.

Table 1 – Massachusetts Reading and Mathematics Cut Scores for Proficient Performance, 2006

| Ranking (Out of 26 States) |         |         |         |         |         |         |
|----------------------------|---------|---------|---------|---------|---------|---------|
|                            | Grade 3 | Grade 4 | Grade 5 | Grade 6 | Grade 7 | Grade 8 |
| Reading                    | 2       | 1       | 4       | 4       | 4       | 18      |
| Mathematics                | 2       | 1       | 2       | 1       | 1       | 2       |

**Note:** This table ranks Massachusetts'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: Calibration across Grades\*

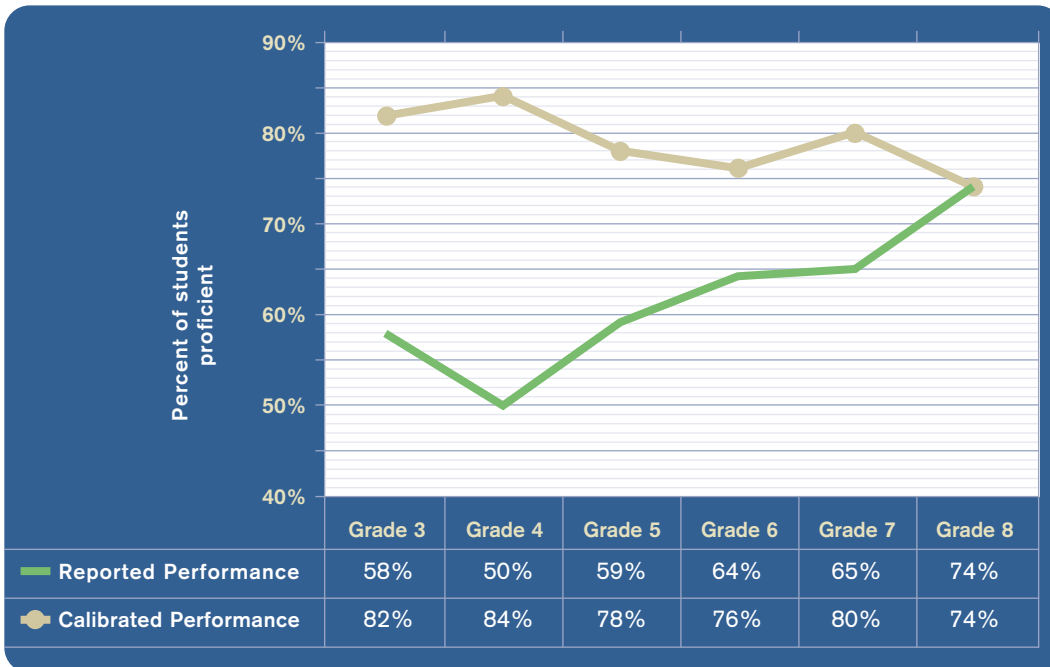
Calibrated proficiency cut scores are those that are relatively equal in difficulty across all grades. Thus, an eighth-grade cut score would be no more or less difficult for eighth graders to achieve than a third-grade cut score is for third graders. When cut scores are so calibrated, parents and educators 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 Massachusetts's cut scores, we find that they are not well calibrated across grades. Figures 1 and 2 illustrated that Massachusetts's reading and mathematics proficiency cut scores differed across grades in terms of their relative difficulty. These figures showed that the reading cut scores at the earlier grades were somewhat more difficult than the cut scores at the later grades. (The opposite is true in most states studied.) The

mathematics cut scores, however, were fairly consistent across grades. These differing patterns are reflected in Figures 3 and 4, which show Massachusetts's reported performance in reading and mathematics on the state tests, and how those proficiency rates would look if the cut scores were all calibrated to the grade-8 standard. In Figure 3, we see that the state-reported proficiency rates underestimate the proportion of students who are on track to eventually meet the easier eighth-grade reading requirements. In Figure 4, we see less difference between the calibrated and actual reported proficiency rates, since the math cut scores themselves are much more consistent across grades.

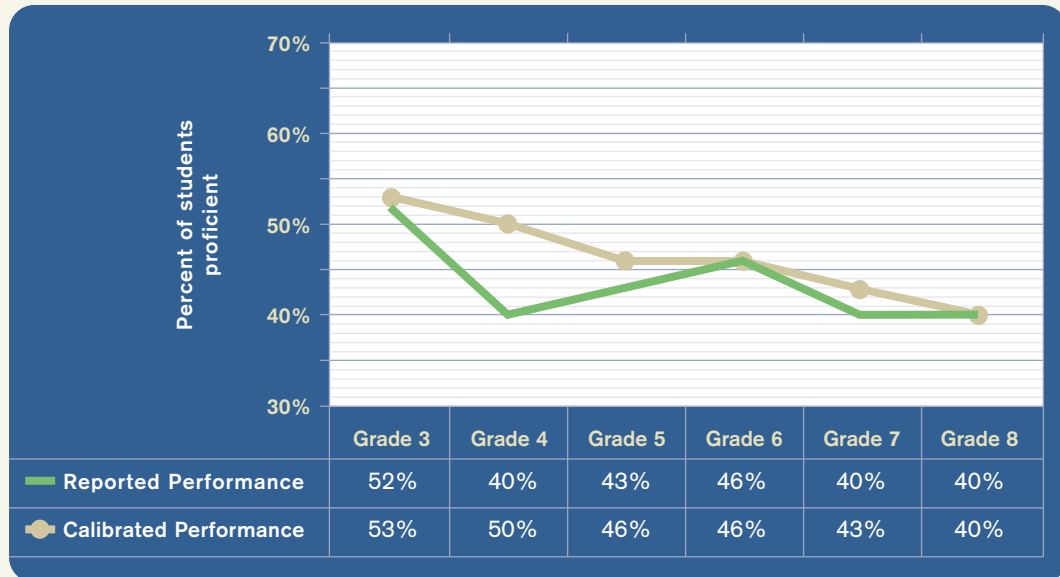
\* Massachusetts was one of seven states in this study for which cut score estimates could be determined only for one year. Therefore, it was not possible to examine whether its cut scores have changed over time.

Figure 3 – Massachusetts Reading Performance as Reported and as Calibrated to the Grade-8 Standard, 2006



**Note:** This graphic means that, for example, if Massachusetts's grade-3 reading cut score were set at the same level of difficulty as its grade-8 cut score, 82 percent of third graders would achieve the proficient level, rather than 58 percent, as was reported by the state.

Figure 4 – Massachusetts Mathematics Performance as Reported and as Calibrated to the Grade-8 Standard, 2006



**Note:** This graphic shows, for example, that if Massachusetts's grade-4 mathematics cut score were set at the same level of difficulty as its grade-8 cut score, 50 percent of fourth graders would achieve the proficient level, rather than 40 percent, as was reported by the state.

### Policy Implications

When setting its cut scores for what it takes for a student to be considered proficient in reading and math, Massachusetts is relatively high, 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 Massachusetts's standards to be in the top third among all states studied. However, Massachusetts's grade-8 reading cut score is significantly less difficult than in earlier grades. State

policymakers might consider adjusting their reading standards 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.