PIONEER
EDUCATION

Post-MCAS Assessments and Accountability in Massachusetts

By Dr. Richard Phelps

PIONEER REPORT
No. 290 | January 2026

PIONEER
INSTITUTE

Mission

Pioneer Institute develops and communicates dynamic ideas that advance prosperity and a vibrant civic life in Massachusetts and beyond.

Vision

Success for Pioneer is when the citizens of our state and nation prosper and our society thrives because we enjoy world-class options in education, healthcare, transportation and economic opportunity, and when our government is limited, accountable and transparent.

Values

Pioneer believes that America is at its best when our citizenry is well-educated, committed to liberty, personal responsibility, and free enterprise, and both willing and able to test their beliefs based on facts and the free exchange of ideas.

Pioneer Institute is a tax-exempt 501(c)3 organization funded through the donations of individuals, foundations and businesses committed to the principles Pioneer espouses. To ensure its independence, Pioneer does not accept government grants.

EDUCATION

This paper is a publication of Pioneer Education, which seeks to increase the education options available to parents and students, drive system-wide reform, and ensure accountability in public education. The Center's work builds on Pioneer's legacy as a recognized leader in the charter public school movement, and as a champion of greater academic rigor in Massachusetts' elementary and secondary schools. Current initiatives promote choice and competition, school-based management, and enhanced academic performance in public schools.

OPPORTUNITY

Pioneer Opportunity seeks to keep Massachusetts competitive by promoting a healthy business climate, transparent regulation, small business creation in urban areas and sound environmental and development policy. Current initiatives promote market reforms to increase the supply of affordable housing, reduce the cost of doing business, and revitalize urban areas.

HEALTH

Pioneer Health seeks to refocus the Massachusetts conversation about health care costs away from government-imposed interventions, toward market-based reforms. Current initiatives include driving public discourse on Medicaid; presenting a strong consumer perspective as the state considers a dramatic overhaul of the health care payment process; and supporting thoughtful tort reforms.

TRANSPORTATION

Pioneer Transportation seeks reforms that allow commuters to take full advantage of the coming mobility revolution—with access to a range of affordable and on-demand public and private transportation options, as well as transit-friendly real estate development.

Table of Contents

Executive Summary

In late 2024, the citizens of Massachusetts voted in favor of Question 2, which prohibits the continuation of the MCAS exam graduation requirement. Technically, however, it changes only one paragraph of state law, leaving other sections in apparent contradiction. The Massachusetts legislature must resolve the ambiguity.

The Massachusetts Comprehensive Assessment System (MCAS) is one product of the 1993 Massachusetts Education Reform Act (MERA). Other components of the holistic reform included extra funding for public schools, more school choice, and an independent statewide school inspectorate. In the years following, Massachusetts students would rise to the top on national and international assessments.

The Massachusetts Teachers Association (MTA), in both person power and funding, embodied the main force behind Question 2. It not only opposed the MCAS exam as a graduation requirement, but any other external exam with direct consequences for students or teachers. It also employed some long-debunked research tropes to argue against testing.

The Voices for Academic Equity Coalition organized the most prominent entities opposed to Question 2. Included were some of Massachusetts' leading political, business, and non-profit organizations.

The governor appointed the K–12 Statewide Graduation Council after Question 2's passage, which included a wide variety of stakeholders. They met throughout 2025 to devise a plan to replace the MCAS examination as a graduation requirement. They proposed a new program of statewide end-of-course tests, along with locally designed portfolios and capstone projects.

Question 2 prohibits only the retention of the MCAS exam graduation requirement. This may have the beneficial effect of encouraging the Commonwealth to consider quality control over the entirety of the K–12 program, rather than just at its end point. It is recommended that such a quality control framework comprise four elements:

1. Consultations with students (and sometimes parents) at key decision points for a student's progress as to appropriate next steps in student's career
2. Choice of curricular pathways in high school with some state-required courses, plus micro-credentials
3. Standardized state examinations with consequences for students:
   - End-of-course examinations in high school
   - Mid-8th grade full battery exam used to aid high school pathway decisions
   - 2nd or 3rd grade exams for reading, and 4th or 5th grade exams for math
4. Re-instatement of the Office of Education Quality and Accountability (EQA) in its original, independent form

It is important that key quality control measures be independent of and external to the entities being evaluated. It is just as important that the education system offer multiple curricular options with multiple targets for success. Students vary in ambition, comportment, organization, and aptitude, and those differences widen over time during their school years. An efficient education system offers meaningful, challenging options of interest to all.

Introduction

In November 2024, a strong majority of Massachusetts voters supported Question 2, which repealed the requirement that students achieve a specified competency level on the Massachusetts Comprehensive Assessment System (MCAS) exam in order to graduate from high school.

Question 2 appeared on the ballot as a Citizens' Petition sponsored by the Massachusetts Teachers Association and allied groups. Proponents began work in 2023 on the petition for an "indirect initiated state statute" that, after the requisite number of signatures were gathered, would be placed on the ballot as Question 2.

The MTA claims 117,000 members, many with spouses, other family, and friends who also vote in Massachusetts.[1] The petition process required two rounds and approximately 87,000 signatures. An organization like the MTA, with a large, organized membership spread throughout the state is far better positioned to gather signatures than an ordinary citizen promoting a cause.

According to Ballotpedia, Question 2 benefited from $16.4 million in campaign spending, with $15.2 million coming from the Massachusetts Teachers Association. Other supporters included the Massachusetts AFL-CIO and some progressive elected officials.[2]

Opponents included several business and education reform organizations, such as the Greater Boston Chamber of Commerce and Democrats for Education Reform, along with the governor, current and former secretaries of education, the Massachusetts Association of School Superintendents, and Pioneer Institute. Former New York City Mayor Michael Bloomberg provided the largest donation—$2.5 million—to the opposition campaign.

Question 2 passed by 59 percent to 41 percent. Perhaps ironically, the dozen or so cities and towns that voted against (i.e., in favor of retaining the MCAS graduation requirement) tend to have more highly educated and wealthier populations (e.g. Brookline, Newton, Wellesley, Weston, Lexington, Concord).

Question 2's Contradictions

Question 2's passage swapped out one part of a single sentence in the Commonwealth's General Laws—in Part I/Title XII/Chapter 69/Section 1D(i). The new wording seems to bestow the power to "determine competency" and therefore diploma eligibility, to school districts, rather than the Commonwealth. The new wording states:

> " ... by satisfactorily completing coursework that has been certified by the student's district as showing mastery of the skills, competencies, and knowledge contained in the state academic standards and curriculum frameworks in the areas measured by the MCAS high school tests described in section one I administered in 2023, and in any additional areas determined by the board."

Notice the reference to "coursework." The old wording referred to an "assessment." Specifically, the old, now discarded text read:

> " ... as measured by the assessment instruments described in section one I."

Yet, other statutory language referring to the old regime remains confusingly on the books. For example, just two sentences later one finds,

> "If the particular student's assessment results for the tenth grade do not demonstrate the required level of competency, the student shall have the right to participate in the assessment program the following year or years."

What assessment program? As the Commonwealth plans to retain the 10th grade MCAS to meet its federal school accountability requirement, one assumes the law's language here refers to the MCAS. But if a student's "competency" is now determined by coursework alone, what does the MCAS have to do with a student's "required level of competency"?

It will be up to the Massachusetts Legislature and, perhaps, the Courts to resolve such questions. As asserted by the Voices for Academic Equity Coalition:[3]

> "Locally chosen and administered examinations cannot replace the MCAS unless they are explicitly aligned with state standards and performance expectations; this alignment is necessary for the validity, reliability, and comparability of the results, as well as to ensure basic fairness for students. The only way to meet the state's constitutional requirements of uniform educational quality is to require that all students are assessed in a manner aligned with state standards."

The Graduation Council

At her third State of the Commonwealth Address (January 2025) Governor Maura Healy revealed her executive order creating the K–12 Statewide Graduation Council.[4]

The Council was tasked with determining what students should learn and how to measure that learning. More specifically, the Council was to focus on how to award Competency Determinations (CDs) for high school diplomas, emphasizing alignment with state standards. Their recommendations were to ensure a baseline of educational quality and meet constitutional requirements. The Council met monthly throughout the calendar year 2025.

In response to the indirect initiative, the Commonwealth's leaders chose to be careful and deliberate, navigating between popular sentiment and an enduring desire to

maintain Massachusetts' reputation and effective standards. Indeed, the Massachusetts Constitution (Article CXXI) specifically mentions quality, in requiring the provision of "resources for quality public education." Continuing (Chapter V, Section II):

> "Wisdom, and knowledge, as well as virtue, diffused generally among the body of the people, being necessary for the preservation of their rights and liberties; and as these depend on spreading the opportunities and advantages of education."

The Council's 30+ appointed members included: five students, three teachers, one parent, three teachers union officials, two district superintendents, two higher education administrators, three education administrator associations, four legislators, three business association executives, three social service nonprofits, and the Commonwealth's top executives for K–12 and higher education.

All these voices would be assembled in one room, so to speak, until they hashed out a plan, so individuals or groups could not say in the end that they were blindsided.

MERA and MCAS: Some Background

The Massachusetts Comprehensive Assessment System (MCAS) covers a dozen subject fields assessed across eight grade levels. Students first encountered the state's high school exit examination in the 10th grade. Passage of a low-threshold score was required to receive a standard diploma. Passage of a higher threshold score could qualify a student for a scholarship to a state college or university.

In 2003, the tenth anniversary of the Massachusetts Education Reform Act (MERA), the University of Massachusetts School of Education published a "Ten Years After ..." report with twelve authors, some of whom were involved in MERA's gestation.[5] The report's summary begins with:

> In June 1993, Governor William Weld signed into law the Massachusetts Education Reform Act (MERA). MERA greatly increased the state role both in funding public education and in guiding the local educational process. The state's role changed to incorporate setting curriculum frameworks and holding schools accountable for student performance. Because MERA was designed to be a systemic reform of education, all the various state activities and policies needed to fit together into a coherent whole based on state educational standards. This comprehensive set of reforms included the following elements:
>
> - Increased State Funding for Public Education.
> - A "Foundation Budget" for All Districts.
> - Learning Standards.
> - Student Assessment.
> - An Accountability System for School and District Performance.
> - Changes in Local Education Governance and Management.
> - Enhancing Educator Quality.
> - Ensuring Readiness to Learn Through Early Childhood Education Programs.
> - Implementing Choice and Charter Schools.

Ten years in, in 2003, the report authors claimed several areas of success.

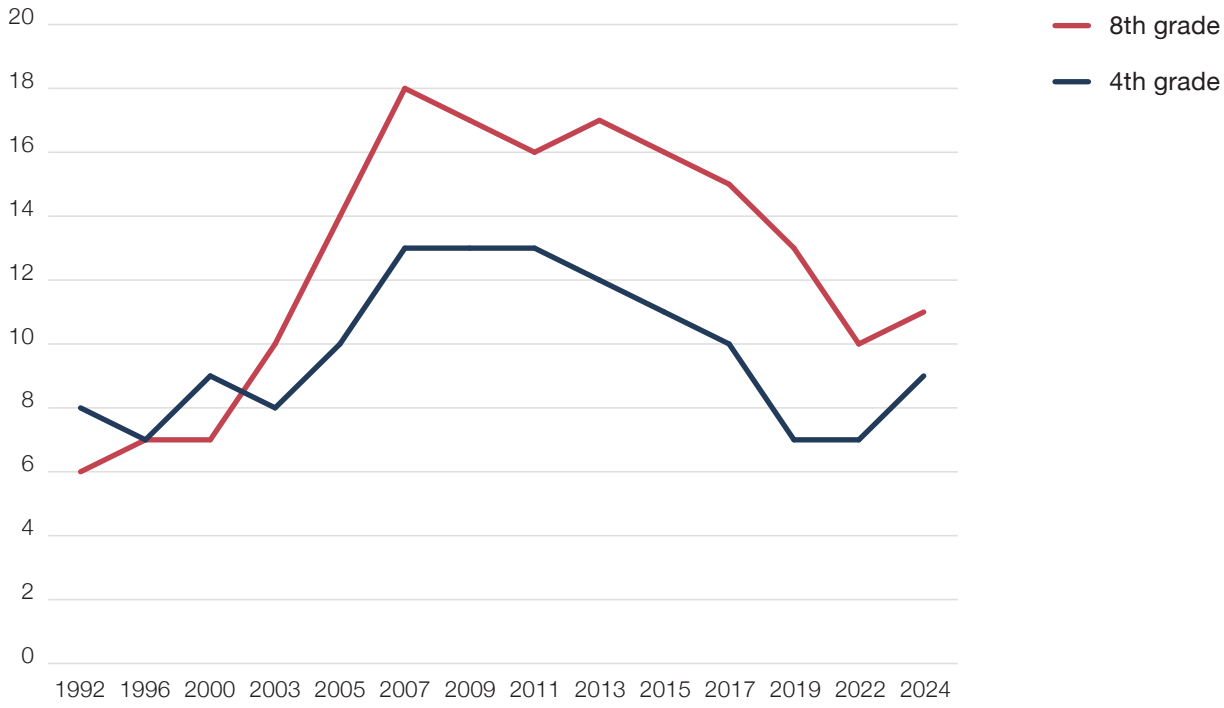
> "1. The legislature has brought virtually all districts to a common level of foundation funding.
>
> 2. The Department of Education (DOE) has developed high-quality, well-aligned standards and assessments.

3. The state is paying more attention to the neediest students.
4. Districts are paying more attention to curriculum and instruction.
5. Students, overall, are improving their performance."

Indeed , Massachusetts was rising to the top spot among U.S. states and on international academic assessments was becoming competitive with high-performing nations. More than just a source of pride, the Commonwealth's position atop the National Assessment of Educational Progress (NAEP) state rankings was attracting entrepreneurs and established businesses to Massachusetts.

Massachusetts once-comfortable lead atop the states, however, peaked starting at 2007 and then steadily declined after 2012. Figure 1 shows the difference between Massachusetts students' NAEP scale scores and U.S. states' public-school average in mathematics over a 32-year period. NAEP reading score differences on the same measure trend more volatile but otherwise look similar.

Figure 1. Difference between Massachusetts NAEP scale score in mathematics and the national public-school average, by year and grade level



Specifically, what happened in state education policy from 1992 up to 2007? The Massachusetts Education Reform Act (MERA) was passed in 1993. The Commonwealth issued MCAS Curriculum Frameworks (i.e., content standards) in 1996 for the Arts, Health, Foreign Language, Mathematics, and Science and Technology/Engineering; in 1997 for English Language Arts and Literacy and History and Social Science; in 2003 for English Language Proficiency; and in 2003 for Kindergarten. In 2000, the Commonwealth established an independent Office of Educational Quality and Accountability (EQA)—its version of a school inspection service.

After 2012? The independence of the EQA had been substantially curtailed. Mathematics and English language arts curriculum frameworks had been adapted to align with the Common Core Standards in 2011. Passing thresholds drifted lower for the MCAS diploma exam. Thus, Question 2 arrived after the MCAS had already weakened.

The Single Target Problem

Some education policy cynics might have predicted the MCAS diploma exam would inevitably fall.

Critics often deride consequential statewide tests as "one-size-fits-all." This criticism is not unique to such tests, however. Most statewide student requirements—such as attendance or course completion—are also uniform by design. Still, it is true that the MCAS tests and their associated graduation requirements applied equally to all students. Most advanced countries don't do it this way. Rather, they manage multi-level, multi-target student assessment systems. An ideal student assessment system would both make sense as an integrated whole, but also offer opportunities and incentives to all students, and students are not all the same.[6]

Multi-level means administering high-stakes tests at more than just one educational level (i.e., primary, intermediate, lower and upper secondary). European and East Asian students typically face high-stakes tests at the beginning or the end (or both) of two educational levels, and, in a few cases, for more than two educational levels (e.g., the end of primary school, the beginning and end of lower- and upper-secondary school, the beginning of postsecondary education, etc.).

Multi-target means that every student, regardless of achievement or choice of curriculum, faces a high-stakes test that, ideally, offers a challenging but attainable goal. In some systems, tests are set at multiple levels of difficulty with multiple levels of certification (e.g., a "regular" diploma and an "honors" diploma). In other systems, different tests cover different subject matter.

In the United States, high-stakes student tests have traditionally been uncommon at any but the upper-secondary level (i.e., high school). Moreover, with few exceptions, they have been single-target tests—meaning that every student, no matter what level of achievement or ability, course selection, or curricular preference, must meet only one common standard of performance.

Ironically, largely social-democratic Europe, with its relatively smaller socioeconomic (and academic achievement) disparities, acknowledges children's differences by offering a range of academic options and multiple achievement targets. The more libertarian United States, with its relatively larger socioeconomic (and academic achievement) disparities, has often provided all children the same curriculum—typically the "college track"—and set a single academic achievement target.

Supporters of the one-size-fits-all traditional U.S. system may label the European and East Asian systems as "elitist" and our system as a more "democratic," "second-chance" system. That contrast may have been valid seventy years ago, but no longer. It is easier to enter upper academic levels in the current European systems, and most countries now offer bridge programs for, say, a dissatisfied vocational-track graduate to enter university or advanced technical programs. Typically, bridge programs are free of charge.

If the U.S. system is neither less elitist nor more conducive to "second chances," in what sense is it superior? It is not, really. In a typical European or East Asian system, multiple programs and tracks create myriad pathways to high achievement, allowing students to pursue their own interests and aptitudes. A Swiss, German, Austrian, Korean, or Japanese student who enters a vocational-technical track at the lower-secondary level and completes an industry-guild certification as a machinist joins an elite of the world's most skilled (and best-paid) craftspeople.[7] Moreover, retention within a grade carries far less stigma in Europe, in part because it is more common than in the United States. That is possible when the perceived goal is to master the subject matter rather than simply avoiding embarrassment.

A single academic-achievement target must by necessity be low: otherwise, politically unacceptable numbers of students will fail to reach the targets. School systems that set low targets typically focus on bringing the lowest-achieving students up to that target. Unfortunately, they may also neglect average- and higher-achieving students.

Everyone acknowledges that students differ in their interests, motivation, comportment, and home environment. In the United States, students also begin with large academic achievement disparities already present in first grade. Absent educator efforts to reduce those achievement disparities, they naturally increase over time. The achievement distribution naturally widens over time as the fastest learning students stretch the top end while the slowest students brace the bottom end.

By 10th grade, when the MCAS diploma exam has been first administered, the wide range in student achievement (and interests, motivation, etc.) from ten years earlier has only stretched wider. That's late in the game to begin seriously evaluating a student's preferred or optimal academic pathway.

According to psychologist Russell Warne:

> "The differences among students' educational achievement start early and increase as children grow. By 5th grade, the average American classroom has children whose achievement in mathematics and reading ranges from the 2nd grade level to the 8th grade level or higher. It is simply impossible for a single teacher to prepare lessons in every subject that allow every student to learn new information."[8]

Yet, that is precisely what "personalized learning"—a popular concept in American education schools—demands. Education schools have actively discouraged most solutions to wide student variation—grade-skipping, ability grouping, grade retention, or curricular differentiation. All in the name of equity.

It remains a political conundrum: how to get teachers, and their unions, to separate themselves from education schools' myopic focus on equity that ultimately harms all students, including those they most wish to support. There simply exists no reasonable method in a constitutional republic to "close" an achievement gap seven grade-levels wide.

Erroneous ideas about how students learn, steeped in Rousseauian theories popular in education schools, make teachers' jobs more difficult. At the same time, education school curricula too often ignore a wealth of practical evidence from more than a century's worth of psychological experiments and rigorous program evaluations—evidence that could make teachers' jobs easier, more satisfying and more productive.[9]

Given the almost impossible task of passing such a wide variety of students on a common rigorous 10th grade exam, Commonwealth educators managed as well as could be expected. They kept the passing thresholds low and provided failing students abundant second chances and supports for passing. Students with certain disabilities were given alternative options for completing their diplomas.

Lower standards exact a price. Over a decade ago, in 2013, a task force convened by Higher Education Commissioner Richard M. Freeland reported[10]

> "... that of the 11,000 community college students who took remedial math in fall 2010, 9,000 have yet to pass a credit-bearing math course. In Massachusetts, 60% of community college students, 22% of state university students, and 10% of UMass students take at least one remedial course."[11]

The MCAS was easy to pass when compared to international standards. And since the Freeland report, passing thresholds for the MCAS exam have drifted lower.

In the early days of the MCAS, the organization Mass Insight studied the nature of student failure on the MCAS exam.[12] They surveyed students who failed on their first attempt

as they negotiated test retakes, remedial classes, and tutoring sessions. The primary characteristic of the students who ultimately remained without diplomas was, by their own admission, failure to show up for the help offered to them for free. Even then, adult education courses leading to a General Educational Development (GED) degree—again, typically free of charge—remained available to them any time they felt ready to finish high school.

Recent versions of the MCAS exam have been less difficult to pass, with only about one percent unable to obtain their diploma solely due to exam failure. Most non-graduates also failed to meet other, local district diploma requirements (e.g., course completions, attendance).[13]

Concern over this one percent has anchored the MCAS exam. In most other advanced countries, a few or even several grade levels earlier those students would have chosen alternative pathways, where they would have faced different challenges and requirements more meaningful to them.

U.S. school systems sometimes offer alternative pathways for the top students, such as gifted programs, state math and science academies, magnet schools, International Baccalaureate programs, and Advanced Placement courses and tests.

The equity focus on the lowest few percent in regular public-school programs, however, can leave the middle two-thirds of students left adrift, neither sufficiently challenged nor learning as much as they could.

If standards for a diploma are to be relatively low despite the wide range in aptitude and achievement, motivating and challenging the middle two-thirds requires offering something extra—whether choices, honors, certificates, badges, or the like.[14]

The No Child Left Behind/Every Student Succeeds/ Common Core Hangover

The federal No Child Left Behind (NCLB) Act of 2001/2002 required all U.S. states to add a new testing program across seven grades. Some states were already administering their own comprehensive testing and accountability regimes. When state budgets tightened at the end of the decade during the Great Recession, most states phased out their own testing programs, as they were still required to administer the federal tests. Ironically, in many states this meant that students, who had for years faced consequential exams, no longer did. NCLB tests had restricted the consequences for student test performance to schools (and in a few locations to teachers).

The U.S. Congress updated the relevant federal legislation with 2015's Every Student Succeeds Act (ESSA), modifying but retaining the federal testing requirement.[15]

Judging schools by student performance creates adverse incentives; for example, schools may increase their average scores by retaining high-achieving students with their age-level peers rather than letting them advance a grade or they may make them take courses in basic subject matter already mastered.

Massachusetts kept its MCAS testing program, but by the mid-2010s the state had aligned it to the academically weaker Common Core Standards offered by a national consortium. The Commonwealth would edit the Core at the margins, but it remained essentially a Common Core test administered within the old No Child Left Behind test administration structure.

Despite mounting evidence of ineffectiveness, Common Core remains in place. Most states have done little more than tinker at the margins of their Common Core-aligned tests. Much now depends on guidance from the Trump administration regarding annual reporting requirements. For two decades, the federal government has required states to

report annually on their compliance with federal accountability regulations. Were states to be relieved of that burden, Common Core would likely wither away over time.

It may be no coincidence that the southern states' recent reading literacy reforms—discussed below—were implemented in the early grades, before federally required testing begins in grade 3. Although these states may still retain Common Core Standards on the books for grades PreK–2, they no longer follow them, at least not in reading instruction.[16] Instead, they built their own reading curricula, with classroom teachers involved in both their development and ongoing evolution. Students in Mississippi and Louisiana read real books with engaging narratives, rather than the excerpts of "informational text" favored by Common Core. Similarly, the high school level enjoys more freedom from Common Core's shackles, as annual federally required testing ends in grade 8. The Commonwealth should therefore feel free to design its own new quality-control system for high school completion.

Testing must primarily focus on students—and specifically on using tests to accelerate learning. As the inimitable Will Fitzhugh—founder of the Concord Review[17] (headquartered in Sudbury, Massachusetts) and outspoken Common Core opponent—has remarked, "It's the students, stupid!" Three aspects of testing accelerate student learning: the act of testing itself (as opposed to simple restudy), consequences for the student (i.e., stakes), and feedback to the student based on test performance.[18]

The federally required No Child Left Behind/Common Core tests made use of only the first of the three, and arguably not even that as most states had already developed sensible accountability programs before the federal government intervened. Just as the subsidiarity principle posits, jurisdictions closer to the classroom, such as states and school districts, should be better able to employ all three of these known learning accelerants.[19]

Contrasting Stakeholder Positions

1. The Massachusetts Teachers Association

Perhaps it was inevitable that teachers' organizations would eventually turn against testing. For decades, most teachers favored standardized testing, including testing with consequences or stakes for the students.[20] Under that earlier regime, testing reinforced teachers' professional authority and gave students a clear reason to take assessments seriously. The NCLB Act of 2001/2002 altered this arrangement by initiating a testing regime with consequences for schools—and in some jurisdictions for teachers—but none for students.

For tests with student stakes, teachers may adopt the role of a coach, helping students succeed on a test imposed by outside forces—as math teacher Jaime Escalante does in the movie "Stand and Deliver" (1988). Under NCLB, however teachers grudgingly assume the role of mendicant, imploring student effort on a test that matters little to them. It should have surprised no one, then, that teachers and their unions recoiled at their degraded status. The Massachusetts Teachers Association (MTA) directed that resentment where it had leverage—at the state level—even though Congress and the federal executive branch were most responsible for creating the incentive structure that produced it. The response was predictable; the attribution of blame was not.

More disappointing, MTA leadership framed its campaign against high school exit exams around tired anti-testing tropes of decades past: that tests make students stressed and anxious, and that absent testing, students are otherwise happy and excited to be in school. In reality, American students place among the least stressed in the world academically. Some modicum of academic stress may even be helpful if it

diverts students' attention away from social media and other sources of teenage anxiety. MTA publications also resurrected the long-discredited concept of individual student "learning styles."[21]

The largest of the teachers' unions in Massachusetts, the MTA is closely linked with the largest national teachers' union, the National Education Association (NEA). More so than its chief competitor, the American Federation of Teachers, the NEA has long opposed external testing with consequences attached to student performance.[22] This opposition reflects not just disagreements over pedagogy, but a deeper resistance to externally imposed accountability.

"External" testing means that school educators do not control most aspects of test development, administration, and analysis. Instead, another authority—a higher level of government or an independent third party—maintains control, typically with strict test security. Educators can influence student readiness for an external test, but they cannot know the specific content in advance, assist students during test administration, or manipulate test materials or results. With externally controlled standardized testing, educators necessarily lack control over both test administration and results. In place of consequential tests, the MTA proposes pseudo-accountability methods that its members can more easily control.

Ironically, in so doing, their anti-testing arguments often imply that standardized tests should somehow be responsible for all educational outcomes. The MTA's Question 2 promotion materials, for example, criticized standardized testing for ignoring "important personal and interpersonal skills" and for failing to address "underinvestment in the schools and other essential services." Of course they do. Standardized tests are designed to evaluate academic knowledge and skills, not social development or public finance. The fact that standardized tests cannot validly measure *everything* important in education is not a serious argument against using them for what they *can* validly measure.

The MTA further asserts that standardized tests do not even cover all academic standards, instead, "focus[ing] intensely on the limited set of standards that are likely to be on the test." Yet, any responsibly developed standards-based test is designed to cover all content standards for that grade level and subject field. Test items are written and field-tested for each standard, not a select few.

Those content standards, in turn, are developed painstakingly over many months by representative committees of teachers, professors, and testing experts—including many MTA members. Which, then, are the non-tested content standards to which the MTA refers? They can only be standards that were never taught during the school year and thus shouldn't be tested in the first place.

In place of externally controlled, objective tests, the MTA favors internally controlled performance-based assessments (PBAs), which grant educators far greater discretion over assessment design, administration, and scoring. This preference is not new.

In 2007, the NEA lobbied the federal government to expand the suite of measures permitted for calculating school progress under NCLB's. Included in the suggested expansion were performance testing in general and portfolio testing in particular.[23]

Intrigued by the NEA's interest in portfolio assessment, the late journalist Robert Holland reviewed the scholarly literature evaluating two statewide experiments with portfolio testing in the 1990s and concluded:[24]

> "In the 1990s, Vermont and Kentucky implemented portfolio assessment as an integral part of education reform plans. Separate studies by nationally respected

> researchers showed that as a school accountability tool, portfolio assessment was a huge flop in both states, yielding results that were wildly unreliable and very expensive to obtain."

The studies Holland summarized identified multiple, recurring failures:

- A failure to yield reliable, comparative data.
- Large differences in the way teachers implemented portfolios.
- Major differences in the opportunities students were given to revise their work, resulting in misleading data when students' collected work was compared.
- Great differences in the degree of difficulty of assignments, rendering comparisons among students or groups of students highly misleading.
- A high price exacted in money, time, and stress on staff.
- A lack of control factors, such as teachers' initial instructions to students.
- Variations in the degree of assistance students amassing their work portfolios receive from peers, parents, teachers, and other sources.
- Opportunities for cheating by importing work not one's own.

Holland asked pointedly: "The question is why anyone sincerely interested in holding schools accountable for results would want to revive such a failed method of assessment."

At the same time, Holland lauded the use of portfolios by classroom teachers to gauge student progress over time, facilitate parent-teacher discussions, and showcase student artwork. For large-scale, high-stakes accountability, however, he concluded that objective selected-response tests offered:

> "... the best value in terms of reliability, accuracy, ability to generalize the results, ease of scoring, and costs. Obviously, education is about much more than test-taking. But to test once a year to ensure that kids are learning to read and compute up to an acceptable standard does not seem to be an unreasonable requirement."

An expanded discussion of the pitfalls of PBAs can be found in the Appendix.

2. Voices for Academic Equity Coalition

Given the "equity" in its title, one might assume the group embodies the most progressive policy interests. In fact, it is a consortium of groups that now represent the political mainstream. These are the groups with most of the money these days, gathered from governments, wealthy foundations, and school choice and edtech investors.

Most importantly, they are supportive of tests that carry consequences, though they may quibble about test formats and where exactly those consequences should fall. Similar consortia in the past supported Common Core testing, with its strong embrace of more easily assessable types of performance testing (e.g., open-ended responses, on demand essays, diagrammatic test items, and interactive test items).

Their report, *Charting a New Path to Success: A Framework on Assessments for the Massachusetts K–12 Graduation Council*, delivers on its title, outlining for policymakers a range of options for planning post-MCAS state assessments. Straightforwardly, the report provides a framework for the Massachusetts K–12 Graduation Council to establish new graduation requirements, focusing on awarding Competency Determinations (CDs) for high school diplomas and emphasizing alignment with state standards. *Charting a New Path* outlines four categories of examinations for consideration, stressing the need for valid, reliable, and comparable assessments. Key considerations include equity impacts, flexibility for diverse learners, and cost implications.

The Coalition deserves kudos for its realistic judgments about both high school exit exams and PBAs. On the former, the report states that a high school exit exam...

> " ... is the gold standard for reliability, validity, and comparability. DESE can guarantee reliability and validity in the testing instruments, and all students in a given cohort will take exit exams in the same testing window." (p. 16)

But that, of course, describes the now forbidden MCAS exams. As for performance-based assessments, the report is notably candid:

> "Among all models considered in this report, PBAs present the greatest challenges for validity, reliability, and consistency across classrooms, schools, and districts—all of which DESE has a state Constitutional duty to preserve. This is because no two projects will be alike."[25]

The authors also declare, however, that "DESE certainly can protect these assessment qualities with rigorous oversight."[26] Experience might cast their confidence as wishful thinking.

3. K–12 Statewide Graduation Council Interim Report

After almost a year of work, the Graduation Council published its Interim Report in November,[27] with a final report planned for June 2026. As the report notes,

> "The recommendations will continue to be built out in the months ahead and additional stakeholder engagement activities will be conducted before drafting and submitting a final report in June 2026 that provides a detailed outline for how to implement the recommendations." (p. 5)

To support its work, the Council hired a contractor, Public Consulting Group, LLC., which appears to have carried out much of the background research, survey administration, and drafting. The resulting Interim Report represents a serious and substantial effort—thorough, reasonable, and intelligent—and it carries considerable "face validity."

The Council drew on a broad base of original data and document review. All Massachusetts residents were free to participate in web-based surveys—when they were aware of them—and 6,615 individuals responded. Represented were cross-sections across stakeholder roles (e.g., students, parents, staff, employers), racial/ethnic groups, and other pertinent demographics, including students with disabilities, English language learners, and parents of both. In addition, 120 of the Commonwealth's more than 300 school districts responded to provide details about their course offerings and graduation requirements.

The Council also reviewed documents from all U.S. states and from many Massachusetts school districts to map the current "landscape" of course requirements, assessments, and student supports.

Though any Massachusetts resident was welcome to respond to the surveys, those already directly involved in schooling—staff, students, and parents—comprise the bulk of respondents—over 100 percent out of a total 119 percent due to multiple selections (i.e., some respondents identified with more than one category).[28] Only about 15 percent out of the 119 percent of respondents were not currently involved in schooling. Employers, for example, accounted for just one percent of the respondent pool (N=69).

These participation patterns help explain both the strengths and limitations of the Interim Report. The report excels at original data collection—especially the surveys—and at summarizing primary sources, such as official documents from other states and Massachusetts' school districts. By contrast, its use of secondary data sources (e.g., literature review of scholarly research) can be hit-or-miss, tending toward readily accessible, first

page-of-search-results material. In the case of performance-based assessments, for example, the scholarly source is the country's most conspicuously funded and enthusiastic proselytizer of that approach. The survey results themselves offer vivid insight into the preferences of the respondent population. The most popular knowledge areas and skills identified among respondents' curricular preferences are those not typically taught in the classroom. That finding is surprising given that school and district staff make up just under half of survey respondents and currently teach something quite different.

Moreover, some of the most preferred topics—such as "communicate effectively," "critical thinking," "interact respectfully with those from diverse backgrounds," and "collaborate effectively"—are ambiguous, difficult to measure, or skills that were once widely regarded as the responsibility of families rather than schools. Among topics that are currently taught, at least in some schools, personal finance and civics jump out as desired courses of instruction, with support near or above 90 percent across all respondent groups. Preferences for assessment instruments similarly reflect respondent composition. Portfolios (over 80 percent) and capstone projects (70 percent or more) are very popular.[29] Standardized tests? Not so much. They garner around 50 percent approval. Notably, employers were 20 percentage points more likely than all other groups to select "standardized testing" in response to the prompt "How Should Students Demonstrate their Learning and College, Career, and Civic Readiness?"[30]

Given these survey results, the Council's recommendations are unsurprising. As the Interim Report explains,

> "These seven elements (program of study, end-of-course assessments, capstones or portfolios, MyCAP, FAFSA/MASFA, financial literacy, and seals of distinction) are intentionally designed to form a cohesive and comprehensive framework that supports student success in college, career, and civic life."[31]

Or, rendered in complete sentences, the Council recommends:

> "Students will ...
>
> 1. ... complete a rigorous program of study that aligns with the admissions requirements for higher education.
> 2. ... participate in end-of-course assessments that are connected to select courses and that are designed, administered, and scored by the state.
> 3. ... complete a capstone or portfolio that is defined by the state, and designed, administered, and scored locally.
> 4. ... develop and maintain an individual career and academic plan (MyCAP).
> 5. ... complete the FAFSA, with an option to complete the MASFA instead, or to opt out of the requirement.
> 6. ... develop knowledge and skills in financial literacy.
> 7. ... have the option to earn state-designated seals of distinction."

More discussion of the Interim Report's policy recommendations follow at the end of the next section.

Quo Vadis Massachusetts?

The MCAS testing program has been lionized—and rightly so—for clarifying statewide performance standards and raising the achievement of many students, with a minimum of negative effects on disadvantaged students. As a quality-control decision point in a Commonwealth student's academic career, however, the high school diploma exam

was lonely, low, and late. That is, it functioned as a single measure, posed a serious challenge to only some students, and came late in students' academic experience.

The Commonwealth's rejection of a statewide graduation examination requirement, however, could nonetheless open new opportunities. Rather than relying on a single exam for all students—regardless of achievement level, interests, and aptitude—the Commonwealth could offer more choice. Properly designed, greater choice can help more students succeed.

Importantly, Question 2 applies only to high school diploma requirements. It does not ban statewide grade-promotion examinations, such as a third-grade reading literacy exam, a fourth-grade mathematics exam, or comprehensive assessments at the end of middle school for all core subjects.

Nor does Question 2 ban non-examination requirements, such as the completion of specific courses or a required number of courses selected from a menu. As the *Boston Globe* has reported, Massachusetts stands out among U.S. states for its relative scarcity of statewide course-completion requirements. In Massachusetts, such requirements are set by local school districts and they vary.[32]

Nor does Question 2 prevent local school districts from requiring districtwide diploma examinations. As many observers have noted, however, with district-level discretion comes district variation. Some Commonwealth leaders remain optimistic that districts will voluntarily choose to adopt common examinations.

Finally, Question 2 does not stop the Commonwealth from awarding certificates distinct from the regular high school diploma.

There are many strategies for quality control in K–12 education. Question 2 hinders Massachusetts from employing only one of them.

A Recommended Quality Control Framework for Massachusetts

Making recommendations—even to an individual friend—carries risk. Offering policy suggestions to a centuries-old Commonwealth of seven million citizens seems even more perilous. The situation certainly requires one disclaimer. Massachusetts' elected officials are best positioned to judge the political prospects of any recommendation. A policy recommendation may make perfect sense theoretically even while making none politically.

One could support, for example, more grade retention, ability grouping, and grade skipping. But state officials must bear the blowback from parents and teachers who object to such policies. Their task is not only to identify what serves the common good, but also to judge what is politically practical.

With that caveat in mind, optimal quality control should cover the entire length and breadth of the system—from PreK to Higher Ed—and operate consistently throughout the state:[33]

> The more quality control checks, the better. Some may be consequential, others merely advisory. Some decision points in a good, comprehensive quality control system consist simply of "checking in" occasionally.
>
> To be meaningful, however, quality control checks must be external and independent of the entity (e.g., district, school, classroom) being evaluated.
>
> Involving teachers as much as practicable in the system that affects them—developing curricula, assessments, and micro-credentials, or inspecting schools—is respectful and encourages agency. After all, teachers manage the system's core competency.

Recommended: a quality control framework comprising four elements

1. Consultations with students (and sometimes parents) at key decision points for a student's progress as to appropriate next steps in student's career.
2. Choice of curricular pathways in high school with some state-required courses, plus micro-credentials.
3. Standardized state examinations with consequences for students:
   - End-of-course examinations in high school
   - 8th grade full battery (probably scheduled mid-year)
   - 2nd or 3rd grade for reading
   - 4th or 5th grade for math
4. Re-instate the original, independent Office of Education Quality and Accountability (EQA)

This may seem like an overly ambitious wish list. But Massachusetts has already begun implementation of much of this agenda.

1. Consultations with students (and perhaps parents) at the end of each level of education—primary, middle, and high—as to appropriate next steps in a student's career

U.S. critics of the widespread curricular tracking found in many European and East Asian school systems often paint a picture of forced-choice where families are told by the state where their child will be enrolled the following year—both in terms of school and curricular track.

In practice, these consultations embody more give-and-take and consider student and parent wishes. Compromises are common. Still, when a student is clearly not qualified to manage a particular curriculum at a given level, families are encouraged to pursue a different pathway that better matches the student's demonstrated readiness.

North Dakota has received much attention in recent years for its efforts to emulate this European-style consultation process while offering students a broad array of choices. Its system is helpfully illustrated in the single chart on the next page.[34]

One may notice among the chart's details a "9-week Career Education Course/Individual Counseling" requirement, as well as a "Career Exploration Experience." The North Dakota Choice Ready framework also *requires* some activities that, in an earlier era, would have been labeled "life skills," such as financial literacy, computer science/cybersecurity, and civics.

The framework offers many choices, but those options are also concrete and demanding, making them difficult for students to ignore or defer.

Above all, the system forces students to consider options and plans that they might otherwise postpone. In the process, it compels them to assume agency over their own education.

Massachusetts has developed its own guidance and support systems and made them available to school at all grade levels.[35] In some respects, these systems are just as detailed—and, arguably more comprehensive—than North Dakota's, starting with an "Early Warning System" in the primary grades. The Commonwealth's "My Career and Academic Plan" (MyCAP) initiative includes an online tool available for each student to use in the furtherance of academic and career planning, though some argue that it is not nearly as widely known or used as it could be.[36]

NDDPI
NORTH DAKOTA
DEPARTMENT OF
PUBLIC INSTRUCTION

NORTH DAKOTA CHOICE READY

The North Dakota ***CHOICE READY*** framework is a tool to assist educators to ensure all students successfully depart high school possessing the **ESSENTIAL SKILLS** necessary to be ready for life. The journey begins by ensuring students leave having the **ESSENTIAL SKILLS** to be successful for whichever path they choose. Students shall then strive to be **POST-SECONDARY READY, WORKFORCE READY,** and/or **MILITARY READY.**

North Dakota
Choice Ready

ESSENTIAL SKILLS

Earn a **North Dakota high school diploma**

Complete a **9-week Career Education Course/Individual Counseling** (15.1-21-18), **Financial Literacy** (15.1-21-21), and pass **ND Civics Test** (15.1-21-27), **Computer Science/Cybersecurity Requirement** (15.1-21-02.2), and **four or more additional indicators:**

- 25 hours of Community Service
- Two or More Years in Organized Extra-Curricular Activities
- 95% Attendance (not counting school-related absences)
- Successfully Complete a Capstone Project
- Career Exploration Experience
- Successfully Complete an Online Learning Course
- Two or More Years in Organized Co-Curricular Activities
- Demonstrate Competency in 21st Century Skills
- Multicultural Awareness

Students shall then complete two or more of the *CHOICE READY* components below.

POST-SECONDARY READY

Complete a **Four-Year Rolling Career Plan,** and earn a **2.8 GPA or greater,** and complete one academic indicator set below:

ACT / SAT minimum or subsections scores or equivalent ND A+ score:

ACT English – 18*

ACT Reading – 22* SAT Reading/Writing – 480*

ACT Math – 21* SAT Math – 530*

*ND A+ equivalent score will be added once available

or

Two or more additional indicators:

- Advanced Placement Course (A, B, or C) or (4, 3, or 2)
- Dual Credit Course (A, B, or C) or (4, 3, or 2)
- Algebra II (A, B, or C) or (4, 3, or 2)
- Advanced Placement Exam (3+)
- International Baccalaureate Exam (4+)
- 3.0 GPA in core course requirement for NDUS admission
- CREAM (Eng./Math) Course (70% or greater)
- Complete three Fine Arts Courses (A, B, or C) or (4, 3, or 2)

WORKFORCE READY

Complete a **Four-Year Rolling Career Plan,** and complete two or more additional indicators:

- Complete three CTE courses or more (A, B, or C) or (4, 3, or 2)
- Complete Career Ready Practices (3.0)
- Dual Credit Course (A, B, or C) or (4, 3, or 2)
- WorkKeys (Gold or Silver)
- Technical Assessment/Industry Credential
- Workplace Learning Experience (40 hrs)
- Work-based Learning Experience (Perkins V) (40 hrs)
- ND A+ (Reading/Math) Mid-point Level 2 or higher (Math 1000) (ELA 975)
- Complete three World Language Courses (A, B, or C) or (4, 3, or 2)
- Complete 2 credits of teaching-related coursework or 2 semesters of Educators Rising

MILITARY READY

Complete a **Four-Year Rolling Career Plan** and obtain an **ASVAB score of 31 or greater** (as determined by branch), or acceptance into the military.

Quality Citizenship (No Expulsions/Suspensions)

Physically Fit (Successfully complete required PE courses (A, B, or C) or (4, 3, or 2))

and

Complete **two or more** additional indicators from the **Post-Secondary** or **Workforce** options

or

Complete two credits of JROTC or Civil Air Patrol

or

Complete **two credits** in the Military Pathway Program, selecting from the following 1/2 credit courses:

- Intro to Military Careers
- Military Health & Fitness
- ASVAB Essentials
- Career Foundations in Military Leadership
- Military Leadership Capstone

Revised October 2025

The key difference between the two states' systems lies in structure and implementation. North Dakota's system is more circumscribed, concentrated on a single, consequential decision point between grades 8 and 9, and is more state-managed. Massachusetts, by contrast, offers students and schools a rich array of resources starting in the early grades, but largely leaves their implementation and enforcement to local districts, schools, and students themselves.

2. Choice of curricular pathways in high school, plus micro-credentials

"Pathway" has become a popular term for curricular tracking. In Massachusetts today, a full flowering—the more choices the better—is available to students. They can choose from a smorgasbord of pathways, especially in career vocational-technical education (CVTE). As mentioned earlier, demand for CVTE seats exceeds supply, and the Commonwealth could—and should—fund additional capacity.[37] Beyond that bottleneck, however, most students can name their own pathway.

Career-tech education costs more than traditional education, but most of its graduates immediately enter the workforce as taxpaying citizens. By contrast, many traditional education graduates remain out of the workforce for several more years while enrolled in higher education, some of them receiving subsidies via government grants or subsidized loans. Notably, CVTE students perform just as well as traditional students on the MCAS diploma exam.[38]

As policymakers and educators have come to recognize the shortsightedness of the previously popular "college-for-all" maxim, many public, private, and non-profit organizations have turned to creating "micro-pathways" or "micro-credentials," sometimes called badges. These offerings typically involve shorter course lengths targeted on more discrete skills and knowledge sets. The country's unfortunate low rate of college completion has led to the unfortunate production of a large population of highly indebted and impoverished college dropouts, often with no degree to show for sometimes years' worth of educational investment.[39]

One type of pathway with room to grow in Massachusetts is for the most advanced students—state advanced academic academies. High-achieving Commonwealth high schoolers can satisfy their academic ambitions through early college, dual credit, and Advanced-Placement offerings. Some schools also offer International Baccalaureate (IB) programs.[40]

But none of those is quite the same as mingling with one's best and brightest colleagues daily in one's chosen field of study under the supervision of university professors. The Massachusetts Academy of Math & Science on the campus of Worcester Polytechnic stands out for its junior-senior years program, but it offers no boarding option for students from other parts of the Commonwealth. With an abundance of higher education institutions and private college preparatory boarding schools, one can imagine more possible arrangements to provide stimulating programs for the state's most academically ambitious students, rich or poor.

Examinations developed by entities other than the state form a crucial component of curricular-choice diploma and credential systems. Many are available. College Board's dozens of Advanced Placement (AP) exams may be the best known in the United States. Great Britain's University of Cambridge offers an array of examinations worldwide tied to its curricula. IB offers its own battery of core course exams as well, but full IB implementation is expensive.

One can find hundreds of different career-tech programs that offer credentials in centers distributed across Massachusetts. Yet, external exams from ACT WorkKeys, the National Occupational Competency Testing Institute, Junior ROTC, iScience, SkillsUSA, or others can sometimes substitute. A wide variety of industry groups and individual firms offer courses and exams leading to certificates in their specialized fields (e.g., Microsoft or Google in cybersecurity).

Passing external exams can reward a student with a nationally recognized credential and standard of achievement.

When all students take the same test, it can be administered in their classrooms. With a potpourri of testing dates, types, and vendors, tests inevitably will be administered in dedicated testing centers. If no such center exists on campus, students will need to travel to one. In dedicated centers, tests are administered by computer under tight security. But the goal is a standardized testing environment that is the same for everyone. Some career-tech testing might be exceptional, where they must be administered on site with the requisite equipment.

Relying on tests controlled by outside entities is not without risk. Testing organizations may change the test in unfavorable ways or drop the tests altogether. New York State, for example, offered the SAT II Subject Tests as graduation test alternatives.[41] In 2021, however, the College Board ended their development of and support for the exams. Copies of older SAT II tests remain widely available online, accessible to students and the general public alike.

3. Standardized state examinations with consequences for students

High School End-of-Course Examinations

Massachusetts leaders have declared their preference for a program of high school end-of-course (EOC) exams.

It is a wise, tried-and-true choice. EOC exams enjoy a long and successful history. Well-known EOCs include the hundreds of Advanced-Placement (AP) exams and the over-a-century-old New York State Regents' exams. Several U.S. states replaced older, minimum-competency examinations with a battery of EOCs in the 1990s and early 2000s. EOCs remain widely used today.

John Bishop's foundational study showed that states using EOCs experienced greater subsequent gains on NAEP scores—and other desirable outcomes—than states relying on other policies such as school rewards, school sanctions, or minimum-competency tests.[42]

End-of-course exams can be structured in different ways. They may be uniform for all students, with nearly all graduates—excluding only those students with the most severe disabilities—required to pass the same core courses, such as language arts, algebra, biology, and history. Alternatively, students could be offered some choice.

For example, students could be required to pass EOC exams in algebra *or* a foreign language, rather than in algebra *and* a foreign language. Such flexibility makes sense under a philosophy that Massachusetts expects its high school graduates to demonstrate mastery of a discrete body of knowledge that prepares them for life after high school—but not necessarily the same body of knowledge for every student. Consider, for example, an unusual student who cannot pass an Algebra course, but has mastered three foreign languages. Conversely, one might imagine a student who cannot pass a foreign language class but aces the more difficult of the two Advanced Placement calculus exams.

Another approach would require that students pass *a certain number* of EOC exams from a menu. For example, the Commonwealth could offer a dozen subjects—such as biology, chemistry, physics, computer science, algebra, geometry, U.S. history, world history, geography, art, performing arts, foreign languages, and various technology fields (e.g., mechanical, chemical, electronic, etc.). Districts might agree to require passage of a subset of these exams—perhaps half a dozen, including at least one in science, one in social studies, etc.). Under such a system, a student could fail one or more EOC exams and still attain a diploma by passing others.

A combination of, or compromise among, these approaches might require a basic core of courses that all students must pass, with choices allowed among a number of more advanced or more specialized courses.

To those skeptics suspicious that such choices represent the first step to watering down graduation standards, consider that most countries offer such choices, and have for decades, either with different diplomas or the same diploma with different curricular foci. Think of the old British O levels and A levels (i.e., different diplomas) or the old French *series*, or tracks, for the humanities and languages, for natural sciences, for math, physics and chemistry, for technology, etc.[43]

8th grade full-battery examination

In light of Question 2's passage, Massachusetts education leaders may feel constrained—at least for now—in their assessment planning at the high school level. Question 2 does not, however, preclude examination requirements at other grade levels.

In K–12 education systems like Massachusetts' and North Dakota's, an 8th-grade full-battery examination makes a lot of sense. Upon entering 9th grade, students

suddenly face a cornucopia of curricular and career choices. Information from a standard achievement test spanning all relevant subject fields could help students make those decisions. To know where one wants to go, it helps to know where one is.

The Commonwealth already administers 8th grade MCAS tests in math, science and technology, civics, and English language arts. All that's missing is writing, the social sciences (other than civics), and non-core subjects like arts and music.

Currently, the 8th grade MCAS carries no consequences for students. The state could decide whether to attach consequences or not. Either way, the exam's primary purpose would be to inform the planning for and assignment to a pathway through high school

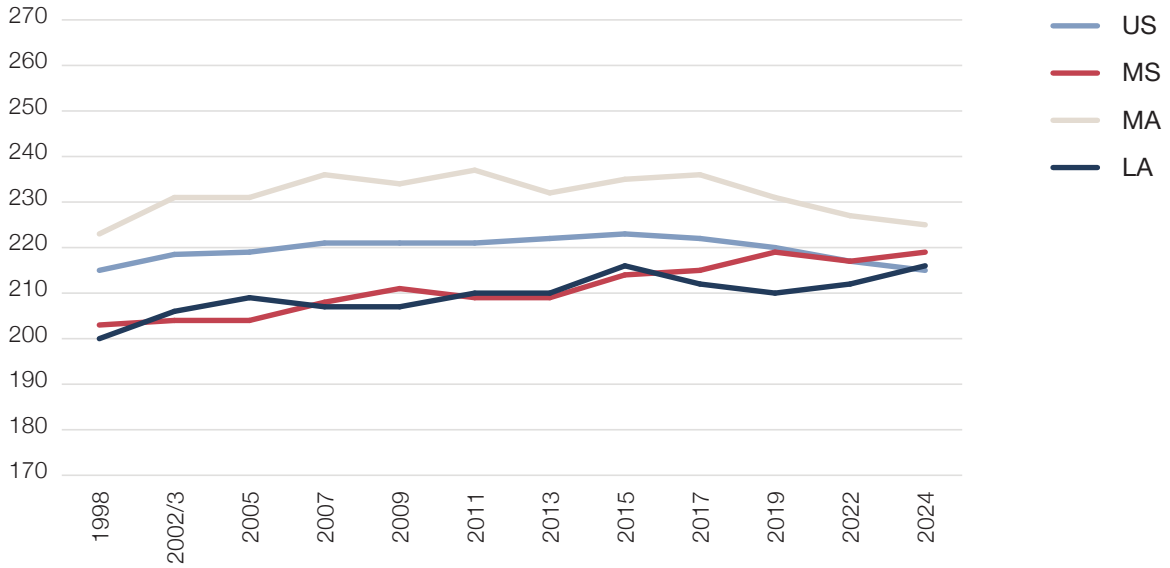
Early-grade reading exam: The "Miracle" moved to Mississippi

The Massachusetts Education Reform Act (MERA) "miracle" was once the darling of education policy observers. But those days—before Common Core subsumed the MCAS—are long gone. Today, Mississippi and Louisiana attract attention for successfully adopting holistic reforms of early-grade reading instruction. By "holistic," advocates mean that all aspects of children's learning were redone at once: curriculum, teacher training, assessment, professional development workshops, and student and parent expectations.

Sound familiar? Massachusetts' turn-of-the-millennium MERA reforms were similarly holistic. Many observers attributed the "Massachusetts Miracle" primarily to the MCAS assessment. But Sandra Stotsky, one of the miracle's architects, has argued that teacher training and higher teacher standards were just as important, if not more so.[44] MERA also gave birth to EQA, expanded school-choice options, and introduced significant changes in education financing and local governance.

Some have urged Massachusetts education officials to copy the southern states' playbook.[45] Readers encountering such exhortations from state organizations or in the media may have seen a chart resembling Figure 2 below. The U.S. Education Department administers the National Assessment of Educational Progress (NAEP) directly—along with its own hired contractors—under secure, consistent conditions. NAEP scores have been reported for decades and remain the country's most reliable data for tracking long-term trends in educational achievement.

Figure 2. Average 4th Grade NAEP Reading Score by jurisdiction, 1998–2024

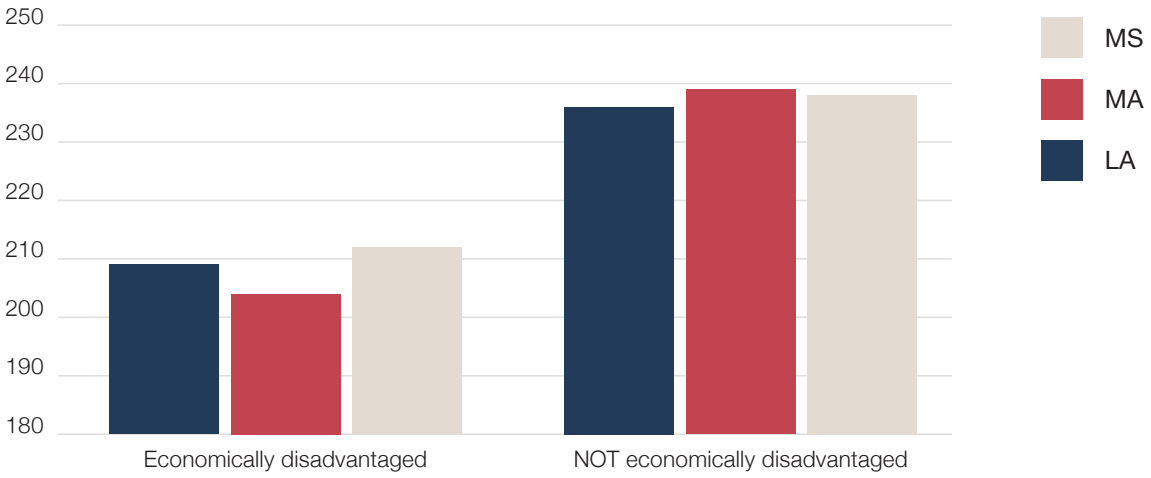


Source: U.S. Education Department, National Assessment of Educational Progress
Note: Time series (horizontal axis) is not truly linear: there was no state NAEP in 2000 or 2021, and the series switched from even years to odd in 2003 and from odd to even in 2022.

Whereas Massachusetts' 4th grade readers improved somewhat up to around 2010, they have fallen back since, notably after the Common Core Standards were implemented.[46] Meanwhile, first Mississippi and then Louisiana applied their holistic early grades reading literacy reforms and have seen steady progress throughout the entire period.

One may be tempted to respond that Massachusetts' 4th grade readers still score noticeably higher than their Southern counterparts, so why make changes? But the line graph above includes only overall average scores. Adjust the NAEP reading scores by "economic disadvantage" and the comparisons look quite different. The Southern states have already caught up with and, among poorer students, surpassed Massachusetts (see Figure 3).[47,48]

Figure 3. 2024 NAEP 4th grade reading scores, by state and economic status



Source: www.nationsreportcard.gov/reports/reading/2024/gr4_8/

Massachusetts' not-economically-disadvantaged test scores remains higher than Mississippi's, but the slight difference falls within the statistical margin of error.

During the 2000s, many high-standards reformers recommended that other states adopt Massachusetts' MERA playbook. A few states made a start at that, until Common Core disrupted this trajectory. Today, the Commonwealth and other states may benefit from studying the progress made more recently by several southern states.

Some critics have objected to Mississippi's program because it has increased grade retention: students who fail the early-grade reading exam may be held back. Mississippi educators report, however, that the number of students retained has been kept much lower than was initially feared, largely because the policy was paired with a wide array of targeted supports for the struggling students.[49]

It is worth mentioning that we would not know of the promising news out of Mississippi and Louisiana were it not for standardized testing. Without consistent external assessments, their progress would be largely anecdotal—not demonstrable.

At the same time, it is important to recognize that these early-grade literacy gains did not occur overnight. Mississippi's and Louisiana's successes grew out of several years' worth of effort that included extensive curricular redesign and teacher training.[50]

Massachusetts legislators and educators have been aware of the southern events for some time. Senate Bill 338, "An Act promoting high-quality comprehensive literacy instruction in all Massachusetts schools" is currently winding its way through the state General Court.[51]

Early grades mathematics exam

If Massachusetts can manage to implement an early-grade reading program, it might consider a similar program for mathematics, perhaps in 4th or 5th grade. Reading remains the most important school subject, but math ranks not far behind.

There may be no state miracle template, as there is for reading. Here, the Commonwealth could be first.

4. Reinstate the original Office of Educational Quality and Accountability

Why? For two reasons. First, it worked. Second, as greater authority devolves to local districts, the Commonwealth must correspondingly reinforce its mechanisms for district accountability.

A school inspection bureau is the old-fashioned quality control method for maintaining some semblance of consistency across jurisdictions. School inspectors are either analysts or school visitors who monitor progress and adherence to standards, somewhat like school accreditation, but operating all the time rather than just periodically. England & Wales have famously relied on their "inspectorate" for over a century.

An effective school inspection must operate independently of a jurisdiction's education bureaucracy to be effective (and honest), however. Completely internal assessments or program evaluations suffer from an asymmetry—pressure for leniency from students, parents, and administrators on one side pushing against a void for rigor on the other side. And even those inside organizations advocating for more rigor may not notice it slipping, as internal evaluations lack external benchmarks for comparison.

From 2002 to 2007, Massachusetts profited from its own inspectorate, the revelatory school district auditor, the EQA, and "Massachusetts witnessed a steady rise in MCAS performance and claimed first-place in all four National Assessment of Educational Progress exams."[52] Responding to opposition from the institutional opponents of accountability (teachers unions and some administrator groups) EQA's replacement lacks much of its predecessor's independence.[53] Reinstating a truly independent EQA would help to blunt the deleterious effects of the MCAS demise.

Ironically, opponents of consequential educational testing often propose school inspection services as a corrective alternative. The type of inspection they have in mind, however, tends to be the internal, just-between-us, no-publicity (i.e., largely ineffective) type.[54]

Comparing Sets of Recommendations

Two overarching differences distinguish the two sets of recommendations. One takes a systemwide view of K–12 education and makes specific recommendations for the early grades and for 8th grade. The Graduation Council, by contrast, focused on the high school years, as its charge required. The governor, however, has also emphasized that Massachusetts should remain the nation's top-performing state academically. Achieving that goal will ultimately require examining the entire system, from PreK through higher education.

A second difference concerns performance-based assessments. One set of recommendations discourages their use—especially portfolios—for external accountability purposes (see Appendix). The Graduation Council, however, recommends their inclusion. There is little dispute about their popularity, at least among those currently directly involved in schooling. The Council's compromise is to have portfolios or capstones managed locally by districts and schools, where they are more workable, while having them "defined" by the state. That's certainly better than attempting to grade portfolios

or capstones on a reliable statewide scale, which numerous previous attempts have shown to be cumbersome and impractical. Effective standardized measures require more standardized student products.

In other areas, the recommendations largely align. A rigorous program of study—check. Standardized end-of-course assessments—check. Individual career plans via MyCAP—check. Coursework in financial literacy—a good (and popular) idea. Statewide seals of distinction (i.e., honors certificates or diplomas)—double check.

Questions remain, however. With respect to completing the FAFSA, what about students with no desire to go to college? At least the recommendation includes an opt-out provision. Similarly, why should students with no intention of enrolling in college complete a rigorous program of study that aligns with higher education admissions requirements?[55] If "higher education" for this purpose is defined to include community college, this recommendation is probably fine. Not all high school graduates want to attend UMass, but the doors of community colleges should remain open to all.

It is disappointing that the Interim Report addresses neither the capacity constraints in career vocational-technical education nor the question of a school inspection office. Here's hoping the final report will address both.

Postscript

Critics often contrast standardized educational assessment against utopian perfection rather than what is likely to take place in its absence. It is true that neither assessments nor the way they are administered will ever be perfect, but the consequences of abandoning assessments are far from perfect, too.

Absent any external assessment, schools may rely solely on teacher grading and testing, which are far more likely to prove idiosyncratic and nongeneralizable than any standardized test. Individual teachers can "narrow the curriculum" to what they personally prefer. Grades are increasingly susceptible to "inflation" as students learn teachers' idiosyncrasies and how to manipulate their opinions.[56] According to research on the topic, many teachers, when assigning grades, tend to consider noncognitive outcomes, including student class participation, perceived effort, progress over the period of the course and comportment. Actual subject-matter mastery is just one among many factors. Moreover, given most teachers' relatively brief training in testing and measurement, it is not clear that their testing and grading practices would be superior even if they focused only on subject-matter mastery.

If the curriculum is not tested, it becomes difficult to know whether any of it works. Without standardized tests, reliably gauging student progress becomes problematic for anyone outside the classroom. One must accept whatever each teacher says, and without external assessment, points of comparison for different classrooms become progressively rarer.

Without either common standards or high-stakes standardized tests, there may be no effective way to monitor systemwide performance at all. Some Commonwealth teachers may do a wonderful job in their totally customized classes, but some may do an unsatisfactory job. How is one to know or tell which? One must hope that teachers will overcome their own natural inclinations—as well as pressures from students, parents, and administrators—to avoid accountability and, instead, hold themselves and their students to high standards. One must also hope that teachers will know how to do so.

Experience elsewhere suggests that such hopes are often disappointed. In jurisdictions that have eliminated external quality-control measures, teachers brave enough to assign

failing grades have sometimes seen those marks erased or changed by administrators in order to avoid controversy and allow the failing students to graduate.[57] In schools where some students pass courses and graduate despite doing little work, other students—and their parents—will assume that they, too, can pressure school administrators for easy credentials. Behind-the-scenes prerogatives then become the implicit academic standards.

Human behavior in environments with ambiguous incentives and no external quality-control checks tends, over time, toward lower standards and poorer outcomes.

Appendix

Performance-Based Assessment is the wave of the future, and always will be

Absolutely, performance-based assessments (PBAs)—such as projects, demonstrations, and portfolios—have their place. One wouldn't wish to be a passenger in an airplane with a pilot who had taken only pencil and paper tests and never flown before, no matter how well the pilot had performed on the paper tests. But there are very real limits to authentic testing's practicality. We cannot afford to have each of our students construct a galactic black hole of their own to test its effects. But we still might want them to know what they can about black holes, even if through disembodied, abstract information delivered by voice or book.

In its excellent pros-and-cons overview of (PBAs) the Voices for Academic Equity Coalition refers to a New Hampshire program as an "Early PBA pilot."[58] New Hampshire's experimentation with PBAs ran from 2014 to 2022. Some of the activity—under the acronym PACE—was federally funded through ESSA's Innovative Assessment Demonstration Authority (IADA).

In March 2022, NH Commissioner of Education Frank Edelblut informed the U.S. Education Department that it had "decided to fully withdraw from participation in the IADA." Why? Among other reasons,

> "Unfortunately, PACE assessment task development, implementation, and scoring placed significant administrative burden on educators and school leaders that detracted from instructional time and was costly to implement. PACE also fell short of expectations for comparability and assessment quality."

George Santayana famously said, "Those who cannot remember the past are condemned to repeat it." Apparently unbeknownst to the Voices for Academic Equity Coalition and New Hampshire government officials, New Hampshire's time-consuming frustration with PBAs simply repeated a cycle experienced in other states and localities going back at least four decades. New Hampshire's PBA program was neither "early" nor "innovative."

The cycle repeats when PBA advocates sell a rosy but incomplete story, neglecting to mention the string of past failures.

PBA systems were implemented disastrously from the late 1980s to the early 2000s in the states of Maryland, California, and Kentucky. Their performance assessments featured presentations of group projects, "hands-on" demonstrations (as in a science laboratory experiment), long written assignments, and open-ended test questions (i.e., no more selected-response questions). They emphasized process over content. Naturally, none of the test results could be machine-scored, so costs ballooned.

The programs ultimately failed because they were unpopular due to: unreliable scores; volatile test-score trends; secrecy in the non-release of items and forms; an absence

of individual student scores in some cases; individual students being judged on group work; large expenditures of time; inconsistent (and some improper) test preparation procedures from school to school; long time delays between administration and release of scores; little feedback for students; and little evidence after several years that education had improved.[59]

With its focus on recent evidence, the Voices for Academic Equity report claims a lack of information on PBA costs. A Government Accountability Office project some years ago collected complete cost data for all systemwide testing in the United States (i.e., all 48 states with testing programs and a representative sample of over 600 school districts). Included was a Maryland oversample that included all its school districts. Maryland attracted interest because, at that time, they tested with PBAs.[60]

At that time, PBAs cost four times as much as mostly selected-response tests.[61]

That was a different era, though, before computer-delivered tests. These days, the cost ratio is likely even greater than four-to-one. Mostly selected-response tests are now both scored and administered by computer, with paper handling, mailing, and teacher proctoring no longer needed (though there's a cost to additional computers when purchased). Meanwhile, PBAs still require natural, authentic, human performance and scoring. In fact, to account for inter-rater reliability, a minimum of two humans is needed to grade each and every exam.

This represents a case study in what economists call "the cost disease." Methods that can substitute cost-saving technology for human labor become relatively less expensive over time while the costs of human inputs inexorably rise.[62] In the larger economy, services tend to rise in price relative to manufactured goods.

While the price difference of performance tests over traditional standardized tests has risen over the years, so has the disparity in test security risk. Traditional, mostly selected-response tests have become more secure when administered by computer. Meanwhile, the World Wide Web and improvements in artificial intelligence render the student-made materials inserted into portfolios increasingly suspect.

Endnotes

1 https://massteacher.org/about-the-mta/who-we-are

2 https://ballotpedia.org/Massachusetts_Question_2,_Repeal_Competency_Assessment_Requirement_for_High_School_Graduation_Initiative_(2024)

3 Voices for Academic Equity Coalition. (2025). *Charting a New Path to Success: A Framework on Assessments for the Massachusetts K-12 Graduation Council*, p. 5. https://voicesforacademicequity.org/wp-content/uploads/2025/10/Fall2025_DataBrief_10.10.25.pdf
Continuing on, "As even ballot question proponents acknowledged throughout the 2024 campaign, the state cannot—legally or in the spirit of educational equity—abandon standards-aligned assessment in determining its graduation requirements moving forward. There must be a statewide process for awarding Competency Determinations that maintains a valid, reliable, and consistent baseline of educational quality associated with a high school diploma."

4 https://www.mass.gov/executive-orders/no-639-establishing-a-k-12-statewide-graduation-council

5 University of Massachusetts Amherst School of Education (2003). Education REFORM: Ten Years after the Massachusetts Education Reform Act of 1993. *Education Connection 2003*. https://nonpartisaneducation.org/Review/Resources/EdConnection.htm

6 See, for example, Figure 1 (p. 12) comparing the systems of the United States, Japan, and Germany, in Mark McQuillen, Richard P. Phelps, & Sandra Stotsky (2015, October). How PARCC's False Rigor Stunts the Academic Growth of All Students, *White Paper No. 135*, Pioneer Institute.

7 Those countries whose students perform best on international assessments employ more "decision points" throughout a student's time in school as a means of quality control. The "points" may be tests or consultations and the "decisions" assign students (or not) to the next grade level, a program, or an ability or curricular group or pathway. See, Richard P Phelps. (2001, August) Benchmarking to the World's Best in Mathematics: Quality Control in Curriculum and Instruction Among the Top Performers in the TIMSS, *Evaluation Review, 25*(4), pp. 391-439.

8 Russell T. Warne. (June 16, 2025). Education's Elephant in the Room, *Quillette*.

9 See, for example, Jay Shalin. (2019, February). *The Politicization of University Schools of Education*. The James G. Martin Center for Academic Renewal. https://www.jamesgmartin.center/wp-content/uploads/2019/02/The-Politicization-of-University-Schools-of-Education.pdf

10 Massachusetts Department of Higher Education. (2016). Key Outcome #2 College Completion, Vision Project

11 https://www.mass.edu/visionproject/completion.asp

12 Mass Insight Education. (2002). *Taking Charge: Urban High School Students Speak Out about MCAS, Academics and Extra-help Programs,* Author.

13 Kinga Borondy. (October 24, 2024). How many students don't graduate because they fail MCAS? A Question 2 Q&A, *Worcester Telegram & Gazette*. https://www.telegram.com/story/news/state/2024/10/24/question-2-how-many-students-dont-graduate-because-they-fail-mcas/75795172007/

14 A good source for learning about micro-credentials is the Education Design Lab, to be found here: https://eddesignlab.org/

15 States were no longer required to use "consortium" tests (i.e., PARCC or SBAC) or align completely to the Common Core Standards. They were still, however, required to adhere to standards closely resembling the Common Core, and that adherence is verified through state's annual reporting requirement.

16 Whereas the Common Core Standards may be best known for their weakness in the upper grades, they may be age-inappropriately difficult in grades K–2, as well as ambiguous, confusing, and tedious.

17 "*The Concord Review* ... [publishes] ... exemplary history essays by high school students ... [with] ... more than 1,500 research papers ... published from authors in 46 states and 46 other countries. *The Concord Review* remains the only quarterly journal in the world to publish the academic history papers of secondary students."

18 See Richard P. Phelps (2019, August 6). Test Frequency, Stakes and Feedback in Student Achievement: A Meta-Analysis, *Evaluation Review.* https://journals.sagepub.com/doi/abs/10.1177/0193841X19865628#abstract

19 Subsidiarity is a key principle in European Union Law. It is also embedded in the U.S. Constitution's tenth amendment, though not always respected by federal lawmakers.

20 See, for example, Richard P. Phelps (2005). "Persistently Positive: Forty Years of Public Opinion on Standardized Testing," Chapter 1 in R. P. Phelps, Ed., *Defending Standardized Testing*, Psychology Press.

21 See, for example, Paul A. Kirschner, John Sweller, & Richard E. Clark. (2006) Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching, *Educational Psychologist, 41*(2), 75–86. DOI: 10.1207/s15326985ep4102_1 and P. A. Kirschner. (2017). Stop propagating the learning styles myth, *Computers & Education*. doi: 10.1016/j.compedu.2016.12.006.

22 Confirming that the two teachers unions continue to differ on testing, American Federation of Teachers of Massachusetts President Jessica Tang joined Governor Healey in announcing the state's end-of-course testing proposal, while the Massachusetts Teachers Association declined to endorse, arguing that the proposal "defies the will of voters." Though, granted, the AFT lobbied in favor of Question 2 along with the MTA. See Sam Drysdale. (December 1, 2025). "Council recommends new tests as high school graduation requirement in Mass." State House News Services. https://www.wbur.org/news/2025/12/01/massachuesetts-graduation-high-school-tests-mcas

23 National Education Association (March 21, 2007) "NEA's Top Legislative Priorities for ESEA."

24 Robert Holland, (2007, September) *Portfolios: A Backward Step in School Accountability*, Lexington Institute. https://www.lexingtoninstitute.org/wp-content/uploads/2013/11/holland_portfolio_assessment_8_29_07.pdf

25 Voices for Academic Equity Coalition. (2025). *Charting a New Path to Success: A Framework on Assessments for the Massachusetts K-12 Graduation Council*, p. 20. https://voicesforacademicequity.org/wp-content/uploads/2025/10/Fall2025_DataBrief_10.10.25.pdf

26 Voices for Academic Equity Coalition, p. 20.

27 Public Consulting Group, LLC. (December 2025). *Reimagining High School: Reimagining Readiness: Interim Report of the Statewide Graduation Council.* K-12 Statewide Graduation Council, Commonwealth of Massachusetts. https://www.mass.gov/doc/reimagining-high-school-reimagining-readiness/download

28 Public Consulting Group, LLC., p. 65.

29 By offering portfolios and capstone projects as available responses alongside standardized tests, the survey itself may have implied that portfolios and capstone projects are equivalent to standardized tests in reliability, validity, and practicality.

30 Public Consulting Group, LLC., p. 78.

31 Public Consulting Group, LLC., p. 6

32 Christopher Huffaker. (October 21, 2024). "Most states have extensive graduation requirements. In Massachusetts, it's just the MCAS." *Boston Globe*. https://www.bostonglobe.com/2024/10/21/metro/mcas-ballot-measure-national-comparison-exit-exams/

33 For more on quality control management systems, see, for example, the American Society for Quality (ASQ) at https://asq.org/ ; the Northeast Advanced Manufacturing Consortium (NAMC) at https://namcnetwork.com/about/ ; the National Institute of Standards and Technology (NIST) at https://www.nist.gov/mml/csd/products-and-services/historical-programs ; the International Standards Organisation (ISO) at https://www.iso.org/quality-management/what-is-qms ; Gartner at https://www.gartner.com/

34 Choice Ready Chart: https://www.nd.gov/dpi/sites/www/files/documents/Division%20of%20SS%26I/ESSA/Choice%20Ready/ChoiceReadyChart.pdf ; Choice Ready Video: https://youtu.be/RUk7GI6KkhI

35 See https://www.doe.mass.edu/ccte/secondary-design.html

36 See https://www.doe.mass.edu/ccte/sec-design/mycap/default.html

37 Ken Ardon. (September 2025). Expanding Access to Career Vocational Technical Education, *White Paper #228*, Pioneer Institute. See also, Roshan Fernandez. (December 27, 2025). The Hottest High Schools in Massachusetts are Trade Schools, *Wall Street Journal*. https://www.wsj.com/us-news/education/massachusetts-trade-high-school-popularity-effc9513

38 Ken Ardon, p. 1.

39 See, for example, Education Design Lab, News & Events at: https://eddesignlab.org/news-events/

40 Massachusetts high schools with IB diploma programs include Abby Kelley Foster Charter Public School; Brockton High School; Mystic Valley Regional Charter School; Nauset Regional High School; Pioneer Valley Chinese Immersion Charter School; Snowden International School at Copley; Sturgis Charter Public School; and Wareham High School. See https://www.ibo.org/programmes/find-an-ib-school/

41 See https://www.nysed.gov/standards-instruction/multiple-pathways

42 John H. Bishop, Ferran Mane, Michael Bishop, and Joan Moriarty (2001). "The Role of End-of-Course Exams and Minimum Competency Exams in Standards-Based Reforms," in Diane Ravitch, Ed. *Brookings Papers on Education Policy*.

43 See Richard P. Phelps (2006, Fall). Characteristics of an Effective Student Testing System, *Educational Horizons.*

44 See, for example, Sandra Stotsky. (2006). Why American Students Do Not Learn to Read Very Well: The Unintended Consequences of Title II and Teacher Testing, *Nonpartisan Education Review/Articles, 2*(1). https://nonpartisaneducation.org/Review/Articles/v2n1.pdf

45 See, for example, Pioneer Institute. (2025, July 16). *On Literacy, Time to Learn from Louisiana & Mississippi.* https://pioneerinstitute.org/news/education-literacy-louisianna-mississippi/ ; Chrisopher Huffaker (October 1, 2025). New England schools are failing—and 'nobody seems to care.' *Boston Globe*. https://www.bostonglobe.com/2025/10/01/magazine/northern-nosedive-reading-math/ ; and the Voices for Academic Equity report.

46 See, for example, Theodore Rebarber (2020). *The Common Core Debacle: Results from 2019 NAEP and Other Sources*, Pioneer Institute. https://pioneerinstitute.org/download/the-common-core-debacle-results-from-2019-naep-and-other-sources/

47 One may wonder how Mississippi's scores can equal or exceed Massachusetts' in both groups while scoring lower on average. This is a case of Simpson's Paradox which, despite its name, is a perfectly natural statistical outcome that occurs when the entities compared have very different subgroup proportion ratios. In this case, Massachusetts' proportions are 39% disadvantaged/60% not disadvantaged, whereas Mississippi's are 75% disadvantaged/25% not disadvantaged.

48 Though not a state, the Department of Defense Education Activity (DoDEA) also participates in the NAEP and has, within the past decade, surpassed all states to rank highest among all participating states, districts, and territories in 4th and 8th grade math and reading. Its scale scores in 2024 exceed Massachusetts' by 4, 8, 9, and 14 points.

49 See Christopher Huffaker. (October 1, 2025). New England schools are failing—and 'nobody seems to care, *Globe Magazine*. https://www.bostonglobe.com/2025/10/01/magazine/northern-nosedive-reading-math/

50 Notably, the southern states banned the inferior "balanced literacy" curricula in favor of research and evidence-based instruction tied to the five pillars of systematic reading instruction—phonemic awareness, phonics, fluency, vocabulary & comprehension.

51 See https://malegislature.gov/Bills/194/S338

52 See, for example, Educational Management Audit Council, Office of Educational Quality and Accountability, *2007 Annual Report*. https://www.mass.gov/doc/2007-eqa-annual-report/download ; Jamie Gass and Grant Wynn (November 1, 2006). *Education Reform in Massachusetts: Using Student Data to Improve District Performance*, Pioneer Institute; Jim Stergios. (July 16, 2010). *Day 4: Recreate a Strong Accountability System*, Pioneer Institute.

53 Sam Davis (2025, August 20). *How Massachusetts Let School Accountability Slip—and Student Achievement with It,* Pioneer Institute. https://pioneerinstitute.org/academic-standards/how-massachusetts-let-school-accountability-slip-and-student-achievement-with-it/

54 See for example, Linda Darling-Hammond, et al. (2016, April 20). *Pathways to New Accountability Through the Every Student Succeeds Act*, Learning Policy Institute. https://learningpolicyinstitute.org/product/pathways-new-accountability-through-every-student-succeeds-act ; and REL Mid-Atlantic (2025, November). Strategies and Considerations for Effective School Inspection Systems, U.S. Department of Education. https://ies.ed.gov/sites/default/files/rel-mid-atlantic/document/2025/11/RELMA-SchoolInspectionsFactsheet.pdf

55 To be sure, there exists a tension between two goals—providing the most efficient path for students to maximize their potential and not wanting to preclude opportunities for them. Some students may not realize that they want to pursue a certain path until they are exposed to the pathway and learn that they can and wish to succeed in it. Indeed, this is one of the arguments for providing college admission test administrations to all high school juniors. Some who had not considered going to college discover that they easily could.

56 According to Mary Tamer of MassPotential "It is far more than this—as one teacher on the grad council stated at one meeting, 'I need the ability to fail students again.' Teachers are being put in a position to pass everyone, regardless of true performance. This allows school and district leaders to tout increasing graduation rates while we know achievement continues to decrease."

57 See, for example, Erich Martel (2017) Documents Related to DCPS Altered Student Records, *Nonpartisan Education Review/Resources*. https://nonpartisaneducation.org/Review/Resources/DCPSdocuments.html

58 Voices for Academic Equity Coalition. (2025). *Charting a New Path to Success: A Framework on Assessments for the Massachusetts K-12 Graduation Council*, p. 21. https://voicesforacademicequity.org/wp-content/uploads/2025/10/Fall2025_DataBrief_10.10.25.pdf

59 See, for example, ACT. (1993). *A study of core course-taking patterns for Kentucky ACT-tested graduates of 1991–1993 and an investigation of the relationship between Kentucky's performance-based assessment results and ACT-tested Kentucky graduates of 1992.* Iowa City, IA: Author; Hamp, P. H., & Summers, C. B. (2002, Fall). Education. In P. H. Hamp & C. B. Summers (Eds.), *A guide to the issues 2002–2003* (pp. 59–69). Rockville, MD: Maryland Public Policy Institute. https://www.mdpolicy.org/docLib/20051030_Education.pdf ; HumRRO. (1998). Linking teacher practice with statewide assessment of education. Alexandria, VA: Author. https://www.humrro.org/corpsite/page/linking-teacherpractice-statewide-assessment-education ; Innes, R. (2003). Education research from a parent's point of view. Louisville, KY: Author. https://www.eddatafrominnes.com/index.html; KERA Update. (1999, January). *Misinformed, misled, flawed: The legacy of KIRIS, Kentucky's first experiment with OBE testing*. Retrieved from https://www.eddatafrominnes.com/pdfs/KERAUP32.pdf ; Montgomery County Public Schools. (2002, February 11). Joint teachers/principals letter questions MSPAP. Public Announcement. Rockville, MD. Retrieved from https://www.montgomeryschoolsmd.org/press/index.aspx?pagetype=-showrelease&id=644 ; Schulte, B. (2002, February 4). MSPAP grading shocked teachers. *Washington Post*. https://www.donaldsauter.com/mspgrad.htm ; Strong, S., & Sexton, L. C. (2002, September). A validity study of the Kentucky performance-based assessment system with national merit scholar and national merit commended. *Journal of Instructional Psychology, 29*. https://www.thefreelibrary.com/A+Validity+-Study+of+the+Kentucky%27s+Performance+Based+Assessment...-a066355142

60 See Richard P. Phelps. (2000, Winter). Estimating the cost of systemwide student testing in the United States. *Journal of Education Finance, 25*(3) 343–380. https://www.jstor.org/discover/10.2307/40704103?uid=3739896&uid=2134&uid=2&uid=70&uid=4&uid=3739256&sid=21106063737141; and U.S. General Accounting Office. (1993, January). *Student testing: Current extent and expenditures, with cost estimates for a national examination.* GAO/PEMD-93-8. Washington, DC: Author.

61 Maryland's test cost $64 per student in 1991 for a battery of core subjects. With a consumer-price-index adjustment to 2025, it would now cost $153. An average mostly multiple-choice full-battery test cost $16 per student in 1991. Today, it would cost about $38. Thus, a full performance-based test cost then about four times as much as a mostly selected-response, machine-scored test. These estimates included all relevant costs—purchase costs and personnel time, in test development, planning, administration, analysis, and reporting.

62 William J. Baumol, & William J.; Bowen. (1965). "On the Performing Arts: The Anatomy of Their Economic Problems". *The American Economic Review. 55* (1/2): 495–502.

About the author

Richard P. Phelps wrote The Malfunction of US Education Policy (Bloomsbury/Rowman & Littlefield, 2023) and edited and co-wrote Correcting Fallacies about Educational and Psychological Testing (American Psychological Association, 2008/2009).

Mission

Pioneer Institute develops and communicates dynamic ideas that advance prosperity and a vibrant civic life in Massachusetts and beyond.

Vision

Success for Pioneer is when the citizens of our state and nation prosper and our society thrives because we enjoy world-class options in education, healthcare, transportation and economic opportunity, and when our government is limited, accountable and transparent.

Values

Pioneer believes that America is at its best when our citizenry is well-educated, committed to liberty, personal responsibility, and free enterprise, and both willing and able to test their beliefs based on facts and the free exchange of ideas.

PIONEER
INSTITUTE

185 Devonshire Street, Suite 1101 Boston MA 02110 617.723.2277
www.pioneerinstitute.org Facebook.com/PioneerInstitute Twitter.com/PioneerBoston