



MAINE
STATE
CHAMBER
of
COMMERCE

mainespark^M
CONNECT TO YOUR FUTURE

60%
by 2025



TESTING 1-2-3. HOW MANY TESTS? WHAT KIND OF TESTS?
THE MAINE SCHOOL TESTING DEBATE | 2020

INTRODUCTION

We test ourselves all the time. We keep track of our daily steps and calories. We make note of when others get raises at work and we don't—vice versa. We keep track of the percentage of times our sports heroes get a hit or throw a touchdown pass.

Usually such measurements are useful. Sometimes they are not. It is possible to have 500 social media friends, and still be lonely; to take 10,000 steps a day, and still live an unhealthy lifestyle; to make a good salary, and still be dissatisfied with one's job.

The fact that the tests are not as useful as they might be is not a reason to stop keeping track. It is a reason to think them through more carefully.

Maybe I didn't take enough steps this week, but am I improving or going backwards? Maybe I'm taking enough steps, but am not eating well. To measure one's personal health, it is useful to have multiple measures, of different types, and to look at them over time.

This is the kind of discussion that is happening today around educational testing in Maine. Many Maine educators are concerned that there are too many school tests, and that the tests measure the wrong things. On the other hand, education is an enterprise on which Maine taxpayers spend \$2.5 billion a year, and to which Maine parents entrust the future of their children, so it is inevitable that people will want to measure how things are going.

Educational testing is the subject of this brief: what are the purposes of educational tests, what is the content, what are the avenues for improvement?

This is one of a series of policy briefs produced by *Educate Maine*. *Educate Maine* is a non-profit organization that works closely with educators, businesses, and government officials. We believe that investment in education creates lifelong learners, opens pathways to promising careers, and grows the economy. It also produces civically engaged citizens who build vibrant communities.

The series is brought to you through a partnership between *Educate Maine* and the Maine State Chamber of Commerce, and through the generous support of the John T. Gorman and Nellie Mae Education Foundations. The production of this report was made possible by the contributions of Lisa Plimpton (research), Frank O'Hara (writing), and Pica (design and layout).

We want to particularly thank those who have shared their experiences and insights with the writers of this report, including: Duke Albanese, Senior Policy Advisor to the Great Schools Partnership and former Maine Commissioner of Education; David Silvernail, former Director of the University of Southern Maine's Center for Education Policy, Applied Research, and Evaluation (CEPARE); Jon Doty, Director of Curriculum, Instruction and Assessment for RSU 34 in Old Town; Tom Gray, Camden Hills Regional High School Teacher and 2019 Knox County Teacher of the Year; Claire Sullivan, Assistant Vice Chancellor for Innovation in Digital Badges and Micro-Credentials, University of Maine System; Debra Allen, Assistant Provost for Institutional Research and Assessment at the University of Maine; and Mandy Barrington, Assessment Coordinator at the University of Maine.

In addition, while many reports were consulted in preparing this brief, one was particularly helpful: *A Review of Standardized Testing Practices and Perceptions in Maine* (April 2018), prepared by the Maine Education Policy Research Institute.



482 CONGRESS ST., STE. 303
PORTLAND, ME 04101
207-347-8638
EDUCATEMAINE.ORG



128 STATE ST. #101
AUGUSTA, ME 04330
207-623-4568
MAINECHAMBER.ORG

WHAT ARE THE PURPOSES OF EDUCATIONAL TESTS?

David Silvernail, a longtime and highly regarded Maine educational analyst, distinguishes between assessments *of* learning and assessments *for* learning.

Assessments *OF* learning (also called *summative assessments*) are provided by the standardized tests most are familiar with, like the Maine Educational Assessment (MEA) or the Scholastic Aptitude Test (SAT). Their purpose is to evaluate how the individual, the school, the state, compare to others with regards to achievement of the expected knowledge and skills of a certain grade level.

These tests are infrequent (once a year); slow to be processed; and often unrelated to a student's current curriculum. For these reasons, they are of little use to teachers in planning classroom instruction. But they are useful to regional and state curriculum planners and school boards, to legislators and taxpayers, and to admissions officers at institutes of higher education.

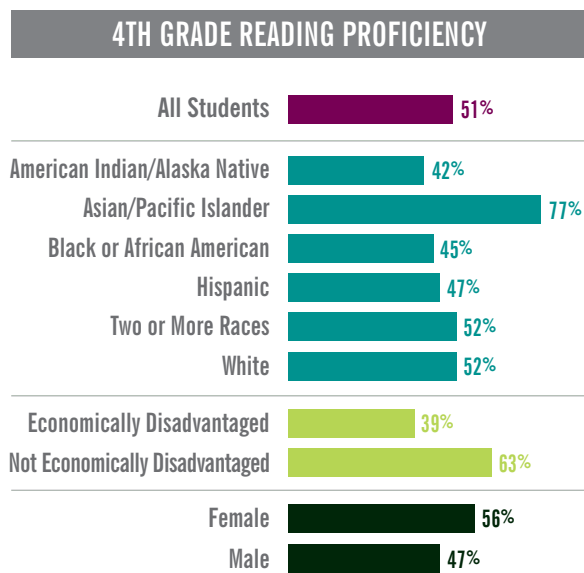
Assessments *FOR* learning (also called *formative assessments*) are conducted primarily by classroom teachers. They may be comprised of essays, presentations, participation in team projects, art work, as well as conventional tests of knowledge and skills.

These tests are frequent and used by teachers to check in on student skills and progress. They help the student, teachers, and parents to know where the student is and what the student needs to work on in the current school year.

At the present time, with current technologies, it is difficult to design an educational test or sequence of tests that meet both purposes at once. This brief is largely concerned with the first type of test, the assessment of learning. However, developments in the second type of testing, which include performance assessments and new innovations like digital badges, are also discussed.

Understanding Student Proficiency Gaps

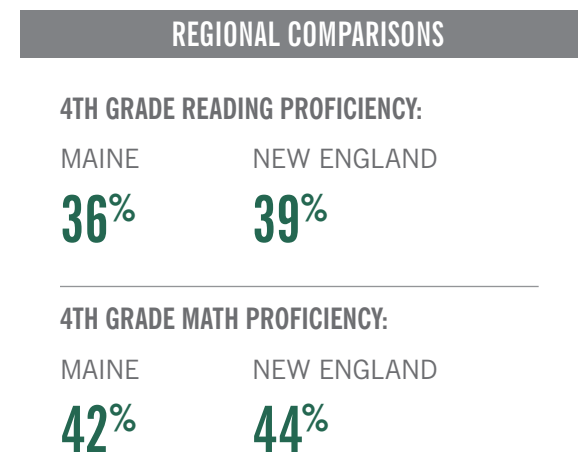
Standardized state tests like the MEA have many uses. One is to allow comparison of how different groups of students are doing. For example, this chart from *Education Indicators for Maine 2019* shows 4th grade reading proficiency by race/ethnicity, economic status, and gender. These assessments show there are important gaps that need to be addressed.



SOURCE: Maine Assessment and Accountability Reporting System (MAARS). Students with English Learner status excluded from race/ethnicity figures.

National and Regional Comparisons

Maine also participates in the National Assessment of Educational Progress (NAEP), which uses standardized tests of random samples of students to track academic outcomes across states. As reported in *Indicators 2019*, Maine 4th graders trail their New England counterparts by several points in math and reading proficiency.



For a display of 8th and 11th grade proficiency, see the *MaineSpark* website at: <https://mainespark.me/future-success-dashboard/>



482 CONGRESS ST., STE. 303
PORTLAND, ME 04101
207-347-8638
EDUCATEMAINE.ORG



128 STATE ST. #101
AUGUSTA, ME 04330
207-623-4568
MAINECHAMBER.ORG

WHAT IS MAINE'S TESTING REGIMEN?

Maine created the **Maine Educational Assessment** in 1984 under Governor Joe Brennan. It was the time of the national debate around the *A Nation at Risk* report, and the tests were part of comprehensive reform package that included tougher graduation requirements, teacher raises, increased state funding, teacher training grants, and mandatory kindergarten. The tests, known as the MEAs, covered reading, writing, and math every year, and science and social studies every other year. They were administered in the 4th, 8th, and 11th grades.

Assessments *OF* learning

In 1997, under Governor Angus King, Maine revised the testing program through the adoption of the **Maine Learning Results**. The Learning Results identified the skills expected of students in each grade in 8 content areas:

- Career & Education Development
- English Language Arts
- Health Education & Physical Education
- Mathematics
- Science & Technology
- Social Studies
- Visual & Performing Arts
- World Languages

In its Guiding Principles, the Learning Results state that every student should leave high school as:

- 1) A clear and effective communicator
- 2) A self-directed and lifelong learner
- 3) A creative and practical problem solver
- 4) A responsible and involved citizen
- 5) An integrative and informed thinker

The MEA tests were revised to incorporate tests of student proficiency in 6 of the 8 content areas (not including career and health education).

In 2001, the Federal Government passed the **No Child Left Behind Act**. The goal of the Act was to equalize educational opportunities for students of all backgrounds. The Act required tests in math and reading every year in elementary school and once in grades 10-12. States were required to give “report cards” to every school. Those schools which failed to make “adequate yearly progress” on bringing students to a proficient level in the tests were open to penalties such as the firing of the principal.

In 2015, the Federal Government replaced No Child Left Behind with the **Every Child Succeeds Act (ESSA)**. The new Act provides flexibility with regard to testing, and eliminates the extreme penalties. Although ESSA allows new options, Maine has opted to essentially keep the same state testing program as before.

Today Maine contracts with a private company named Measured Progress, based in New Hampshire, to administer assessments called **eMPowerME** in mathematics and literacy in grades 3-8, and a science assessment for students in grades 5, 8, and 11. Maine also uses the **Scholastic Aptitude Test (SAT)** results from the College Board to test 11th graders.

The Maine Department of Education works with Measured Progress to determine the “state expectation” cut scores each year, and the expectations have changed over time. A comparison of Maine with other states found that Maine’s proficiency standards are of average difficulty for reading, but well above average difficulty for math.¹

Assessment *FOR* learning

In addition to state assessments, a majority of public school systems in Maine contract with private vendors for testing and assessment work. A 2018 survey of District Assessment Coordinators in Maine² found that 99% of districts were using at least one type of commercially-developed assessment. The most popular products are tests from the **Northwest Assessment Association (NWEA)** that enable teachers and school districts to do detailed analyses of where individual students and classrooms stand. Many used more than one product. Most districts (70%) also have developed their own tests to measure student progress.

Time and Money

The cost of state standardized tests in Maine is about \$6 million per year, or \$26/student.³ Jon Doty, the RSU 34 Curriculum Director, thinks that these numbers underestimate the true cost of the state standardized tests, and that a complete accounting of cost should include the teacher training time (that could be dedicated to other classroom related learning), the device and network preparation, the cost of the lost instructional time, and the cost of instruction disrupted by make-up testing. Doty’s district estimates the cost of instructional time at \$11.32 per student per hour for 2019-2020. A survey of Maine teachers for the MEPRI study⁴ found that standardized test-taking and preparation took 12-21 hours of classroom time per year for grades 3-8. Counting the lost instructional time, an estimate of the true cost of the state testing may well be in the range of \$100-\$200 per student.



482 CONGRESS ST., STE. 303
PORTLAND, ME 04101
207-347-8638
EDUCATEMAINE.ORG

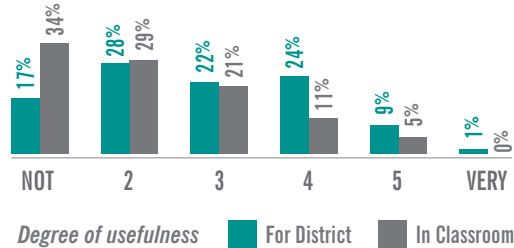


128 STATE ST. #101
AUGUSTA, ME 04330
207-623-4568
MAINECHAMBER.ORG

WHAT DO MAINE EDUCATORS THINK ABOUT THE STATE-MANDATED TESTS?

Maine educators, looking at the state tests through the lens of how useful the results are at the district and classroom level, generally do not think highly of them. Most think they are not at all useful or only marginally useful (see chart).⁵ To the extent that they are considered useful, it is at the district, rather than the classroom level.

DISTRICT ASSESSMENT COORDINATOR VIEW OF STATE TESTS USEFULNESS



One of the characteristics of an effective testing system is **consistency over time**. Consistency in assessment standards allows educators to measure progress. Maine has adopted new standards and new testing protocols every 5 or 10 years over the past few decades. A quarter of the written comments on the coordinator survey concerned this issue. One comment was typical:

*Because our state test assessment has changed so many times over the past few years, the data is virtually useless in making any district-level decisions.*⁶

Other coordinators complained that the data is **too slow in coming** to be useful in making student placements, and is unable to be aggregated to see the performance of subgroups of students:

*State-mandated testing is very disconnected from the data we use in our schools and classrooms to inform programming and instruction. First, we receive the data far too late (6 months or so after the tests are taken) to impact current student programming. Second, at this point the data is at such a global level we cannot drill down far enough to look at trends and achievement on specific items for individual students.*⁷

Teachers echoed the comments of the District Assessment Coordinators:

*We don't get access to our test results until the following year. So teachers can't plan too well based on that, and parents are getting feedback about how the previous grade went after it's already in the rear-view mirror.*⁸

*It does not factor in where the student was at the beginning of the year in terms of background knowledge and how far they have come since then. Instead, these scores keep deflating students who have difficulty learning and may never meet the benchmarks their peers do. Special education students are an example. No matter how much they grow throughout the year, their standardized test score will always show that they never meet "the standards" set for students who are their age.*⁹

Teachers have a much higher view of the commercial testing products that their schools purchase for assessment purposes (assessment *for learning*). The state tests ranked below the commercial tests in terms of feedback to parents, helping the teacher in the classroom, or helping the school district adjust curricula.

TEACHER PERCEPTIONS OF USEFULNESS OF STATE-MANDATED VS. COMMERCIAL TESTS

	STATE	COMMERCIAL
Feedback to students & parents	2.59	3.84
Guiding instructional practices	2.37	3.71
Helping school & district make changes to curriculum	2.93	3.60

(mean rating, 1 = not useful, 6 = very useful)

HOW ARE STANDARDIZED TESTS VIEWED AROUND THE NATION?

Maine educators are not unique in their concerns. There is a national conversation underway about how to improve student testing. Different national organizations recommend different strategies.

Excel in Ed recommends streamlining and improving the existing tests. The nonprofit organization believes that standardized tests are important:

Parents deserve to know how their kids' schools are doing—plain and simple. Only standardized test scores can provide parents with clear, consistent and objective information about how their kids are doing compared to other students across the state and across the nation. This information empowers parents to hold schools accountable and improve the quality of education in their schools.

Our kids deserve the opportunity to be prepared for good jobs. To get those jobs, they need the right skills. Test results arm parents with the information to ensure their kids don't fall behind—and better-informed parents mean better-prepared kids.¹⁰

But the group thinks that such tests can be improved. The organization recommends that state education directors “evaluate the quality of all state tests and remove/replace any state tests deemed duplicative, unnecessary, or low-quality.”

The group **FairTest: The National Center for Fair and Open Testing** takes a more radical approach. The Center argues that standardized testing should be thrown out entirely, because such tests:

... reward quick answers to superficial questions. They do not measure the ability to think deeply or creatively in any field. Their use encourages a narrowed curriculum, outdated methods of instruction, and harmful practices such as grade retention and tracking.¹¹

The Center wishes to replace standardized tests with performance evaluations that use information from a variety of sources and are designed and administered at the local level. The organization's research cites the State of Wyoming and the Scandinavian countries of Finland and Sweden as providing models for such locally-driven performance evaluations.

The **Center for the Improvement of Educational Assessment** steers a middle course between tinkering with the existing system and throwing it out. This group argues for what they call a “balanced assessment” system.¹²

A balanced assessment system typically includes at least three levels of assessment measures: summative measures (such as annual state tests), interim assessment measures (such as common tests administered by districts or schools), and formative assessment measures (an assessment process used to direct classroom learning and teaching). These measures help provide coherent information across the levels of the educational system.

The trick is making all three levels of testing work coherently together. Balanced assessment systems require content standards that are agreed upon, and timely reporting of results.

New Hampshire runs a pilot program called the **Performance Assessment for Competency Education (PACE)** with twelve school districts.

The districts, individually and together, develop assessments that:

...support deeper learning, as well as allow students to demonstrate their competency through multiple performance assessment measures in a variety of contexts.

PACE implementing districts give the NH state standardized assessment once in elementary school, once in middle school and the SAT in high school. In all other years, the PACE districts administer carefully designed common and local performance assessments developed by the districts themselves, and validated at the state level.¹³

In short, the views from around the nation are as diverse as they are within Maine.

INNOVATIONS AT MAINE COLLEGES

One of the purposes of high school assessments is to determine whether students are prepared for postsecondary education. The 11th grade assessment in Maine is the Scholastic Aptitude Test (SAT), which is commonly used in college admissions applications.

Maine chose to use the SAT as its high school student assessment in part, according to former Education Commissioner Duke Albanese, because many high school juniors would take the test more seriously than the MEAs. Students are more likely to fully demonstrate what they know and can do on the SAT since it is tied to college admissions. In addition, giving all students the opportunity to take the SAT without cost may contribute to raising college aspirations. David Silvernail points out that the SAT score results have some correlation with college grades, but the test misses much about student character, skills and interests.

Bowdoin, Bates, and some other colleges in Maine now make SAT test results optional in their admissions processes. But in most college admissions processes, high SAT scores are advantageous. However, once a student is in college, SAT test scores are largely irrelevant.

Instead, the University of Maine and other colleges are moving towards **performance assessments**. Debra Allen and Mandy Barrington point out:

Students can get any info they want now, compared with 20 or 30 years ago. They don't need to "download content" in their discipline as much as in decades past. Providing students with skills is more important now, because content can become irrelevant rapidly. We're moving to a more skill-based way of assessing and thinking about student learning in higher ed.

The workforce changes so quickly, we need to instill hard skills and information navigating skills. Arming students to articulate that this is what I know and can do, and that's why I'm prepared for this job.

So we look at "high-impact practices," things like study abroad, internships, capstone projects, e-portfolios that are different than taking quizzes and test and writing.

Maine colleges and universities are also implementing a practice of certifying "microcredentials" and "digital badges" for cocurricular activities.

Traditionally, what an employer knows about a graduating student is that person's major, minor, and grades - which does not provide a full picture.

The University of Maine offers **Engaged Black Bear Digital Badges**¹⁴ that provides an employer with information about specific experiences and work in fields of importance. A **badge** is a certificate that a student has done important work in a particular area—such as, in the University of Maine's case, environmental stewardship, global perspectives, multicultural awareness, outdoor leadership, community service, or any of 13 other categories. Each badge requires a unique set of work experiences or products, and each is maintained online, along with associated research or videos, for any employer or graduate school to look at.¹⁵

As these practices become more common at the college level, they will begin to filter downward—first to college admissions processes, which in turn will promote changes in high school grading practices.



LOOKING AHEAD

David Silvernail cautions that the Maine Education Assessment will still be necessary in the future. Parents, educators, elected officials, and interested citizens need some kind of standardized testing to assess what students are learning and how we compare to other populations. The important thing is to be clear about what we are looking for. Silvernail adds,

I would like to see them continue to refine the MEA to make it the best snapshot assessment of learning that it can be, and that the DOE would have a cadre of people to help schools make sense of their MEA results, and to help schools develop tools for assessment FOR learning. We should be asking schools: What else are you doing that's important for learning? Let's devise a way to measure that and demonstrate those values. Broaden the definition of what it means to demonstrate learning.

While Maine teachers rate private testing products like the NWEA higher than the state tests, Silvernail responds that both are needed. Private products cannot replace the MEA because their methods are proprietary and not subject to public review, and are not designed to meet specifically Maine educational goals. But private products also provide more help to teachers in the classroom, because they provide faster results that can be related to individual students and groups.

Tom Gray, a Camden High School teacher, looks to technology to make assessments more effective:

There's an increasing role for technology, using digital learning platforms to help kids demonstrate learning and prepare for assessments. That opens up a lot of opportunities, but we need to not replace human interactions with machines. We need to be judicious about it.

Gray's point is reaffirmed by Debra Allen and Mandy Barrington at the University of Maine:

We're assessing students at various points for the purpose of improving a program, improving instruction. But when you think of the course level--with all of the changes in technology like flipped classrooms, clickers, and response apps--instructors in the future will be able to have constant assessment of students and feedback from students, and that can improve the formative assessment. I think the ways we do assessment will be different, because technology is evolving and improving so rapidly.

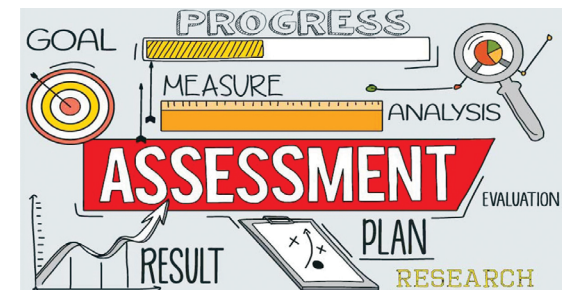
Jobs for the Future recently published a market scan of workforce assessment technology, and concluded that technology is reducing the cost and increasing the efficacy of workforce assessments.

Today's advances in assessment technology help us create accurate, affordable, and efficient systems to evaluate talent based on what they know and what they can do, as opposed to relying on inaccurate signals such as a degree, pedigree, or social and professional network connections.¹⁶

Assessment is a field that is constantly changing. However, it is also a field in need of stability and consistency, in order to be able to make measurements over time.

The last word belongs to Maine's local District Assessment Coordinators. They believe that state-mandated tests could prove more useful in the future if they could:

- maintain consistency over time,
- return results more quickly, and
- align assessments more closely with curricula.¹⁷



SOURCE: Image from <https://www.trackyourmetrics.com/index.php/2019/07/17/midline-assessment/>

RECOMMENDATIONS

Our recommendations are based on the different purposes of assessments of learning—and assessments for learning used in school districts and classrooms. By considering these assessments separately, Maine can make both more effective.

- 1) Move the timing of eMPowerME tests (now typically administered in March) closer to the end of the school year.** This would give teachers and students a full year of instruction to meet state expectations. Since it is unrealistic to quickly return the test results to schools, and the end of the school year is typically lighter on content instruction, use the last part of the school year for state testing.
- 2) Minimize time spent on testing by randomly assigning students to take only one section each of math and English Language Arts assessment.** Currently all students take two math and five Language Arts test sections every year. By acknowledging that these are assessments of learning at the district and state level—and not assessments for learning at the student level—Maine can reduce the testing burden on schools and students, and still collect robust test results.
- 3) Establish transparency about the process for setting state expectation test score levels each year, and whether and when expectation standards change.**
- 4) Establish a pilot program of interested Maine school districts for the development of assessments that use multiple measures to assess a broader set of student competencies, similar to New Hampshire's PACE pilot and Maine's former Local Assessment System.** Involve teachers in the development of assessments and scoring rubrics.



482 CONGRESS ST., STE. 303
PORTLAND, ME 04101
207-347-8638
EDUCATEMAINE.ORG



128 STATE ST. #101
AUGUSTA, ME 04330
207-623-4568
MAINECHAMBER.ORG

ENDNOTES

- 1 *Mapping State Proficiency Standards onto the NAEP Scales: Results From the 2017 NAEP Reading and Mathematics Assessments*, August 2019
- 2 Page 20, Fairman et al, *A Review of Standardized Testing Practices and Perceptions in Maine*, April 2018.
- 3 Ibid., page 65
- 4 Ibid., page 24
- 5 Ibid., page 26
- 6 Ibid., page 29
- 7 Ibid., page 29
- 8 Ibid., page 31
- 9 Ibid., page 33
- 10 See <https://www.excelined.org/fewer-better-tests/>
- 11 See <http://fairtest.org/facts/whatwron.htm>
- 12 See <https://rmcresearchcorporation.com/portsmouthnh/wp-content/uploads/sites/2/2019/01/Balanced-Assessment-Systems-GONG-002.pdf>
- 13 See <https://www.education.nh.gov/assessment-systems/pace.htm>
- 14 See <https://umaine.edu/engagedblackbear/learning-pathways>
- 15 See <https://umaine.edu/engagedblackbear/about-ebb/#what>
- 16 *Jobs for the Future, 2020 Assessment Technology*
- 17 Ibid., page 29



482 CONGRESS ST., STE. 303
PORTLAND, ME 04101
207-347-8638
EDUCATEMAINE.ORG

MAINE
S T A T E
CHAMBER
of
COMMERCE

128 STATE ST. #101
AUGUSTA, ME 04330
207-623-4568
MAINECHAMBER.ORG