A Shifting Landscape for State Testing

COVID-19 added a layer of complexity to a fragmented market and states' interest in different approaches.

Lynn Olson

Last spring's sudden shift to remote learning in response to the COVID-19 pandemic disrupted K-12 education in an unprecedented way. Upon Education Secretary Betsy DeVos's waiver of federally mandated state testing for spring 2020, it also triggered the first nationwide break in state testing in half a century.

With students returning to school for the 2020–21 school year, state boards of education will need to get a handle on exactly where students are in their learning in order to target resources and support. States also will need to restart annual assessments mandated by the federal Every Student Succeeds Act (ESSA) in order to track progress over time, particularly for vulnerable subgroups of students.

Yet even before the pandemic, two recent FutureEd reports found that state testing systems were already in transition. A national analysis by FutureEd found that between 2014 and 2019, lawmakers in 36 states passed legislation to respond to concerns about overtesting, including by reducing testing in a variety of ways.[1]

That trend is likely to continue, as opponents to standardized testing eye the coronavirus crisis as an opportunity to leverage more cuts—and even end state testing entirely. At the same time, interest is growing in innovations that could make state testing systems more useful for teachers and students by moving away from a single, end-of-year test to more frequent assessments that better reflect classroom curriculum and instruction.

As state boards think about the right questions to ask about testing in order to make smart decisions moving forward, it is important to understand the history of state summative assessment in the United States, the increasing fragmentation of the state testing landscape, and the larger political context.

A Half Century of Testing

Standardized testing has been part of the K-12 education landscape for the past 50 years, with both Democratic and Republican leaders at times pushing for higher standards and greater accountability. During this time, state tests have been viewed as an essential component of educational improvement—providing a window into school performance, helping policymakers map strategies and allocate resources, and ensuring that the needs of underserved students are being addressed.

In 1969, with federal dollars flowing to schools under the Elementary and Secondary Education Act (ESEA), Congress mandated a federally funded snapshot of student performance, the National Assessment of Educational Progress. The same year, Michigan launched the first statewide testing program.

In the 1970s, concerns about the performance of high school students in particular led to the "minimum competency" movement and expanded state testing to ensure that students graduated with the requisite basic skills.

In the 1980s and 1990s, as the nation transitioned to a knowledge-based economy, political leaders began to push for higher educational standards and national goals, as well as efforts to hold schools and districts accountable for results, particularly for the nation's most disadvantaged students. The Clinton administration's reauthorization of ESEA in 1994 required states, for the first time, to adopt state standards that would be the same for all students and to test all students' progress against those standards in at least three grades.

Not all states responded to the 1994 requirements with equal rigor. So when George W. Bush took office, he decided to place significantly more emphasis on tests

and test results in the next reauthorization of the law, with the goal of ensuring that the needs of students furthest from opportunity were being addressed. The No Child Left Behind (NCLB) Act of 2001 mandated that states test every student every year in reading and math in grades 3-8 and once in high school and in science at least once in elementary, middle, and high school. NCLB also required states, districts, and schools to publicly report test data by race and income. And it set strict timelines for schools to get every student to the proficient level on state tests or face an escalating series of supports and sanctions.

NCLB was designed to shed a bright light on educational inequities to counter what President Bush described as the "soft bigotry of low expectations." But it had unintended, negative consequences, as states, districts, and schools frequently responded to the pressure in ways that jeopardized student learning and kindled antitesting sentiment.

NCLB was designed to shed a bright light on educational inequities to counter the "soft bigotry of low expectations."

Schools emphasized instruction in tested subjects at the expense of untested subjects and stressed test-taking skills.[2] School districts piled on new benchmark tests to gauge how students would perform on end-of-year exams. Many states began to rely heavily on multiple-choice tests because they were cheaper and easier to administer in the face of tight testing timelines. And many states lowered their testing standards to get more students over the proficiency bar.

Recognizing the need for more consistent and ambitious standards, a group of states began collaborating to identify the knowledge and skills needed for college and career readiness by higher education and employers. The work laid the foundation for a 2009 agreement by the National Governors Association and the Council of Chief State School Officers (CCSSO) to jointly develop demanding voluntary standards in English language arts and math—what became the Common Core State Standards. That work coincided with the Obama administration's Race to the Top initiative, which provided billions of dollars in education funding to states to help address the 2008–09 economic crisis.

By making the competitive grants contingent on states adopting the reforms, the $4.3 billion program incentivized states to adopt tougher academic standards and more rigorous tests aligned with those standards. Most states quickly embraced the Common Core standards and joined one of two voluntary state consortia to develop Common Core–aligned tests with federal funding: the Smarter Balanced Assessment Consortium or The Partnership for Assessment of Readiness for College and Careers (PARCC).

The Backlash Begins

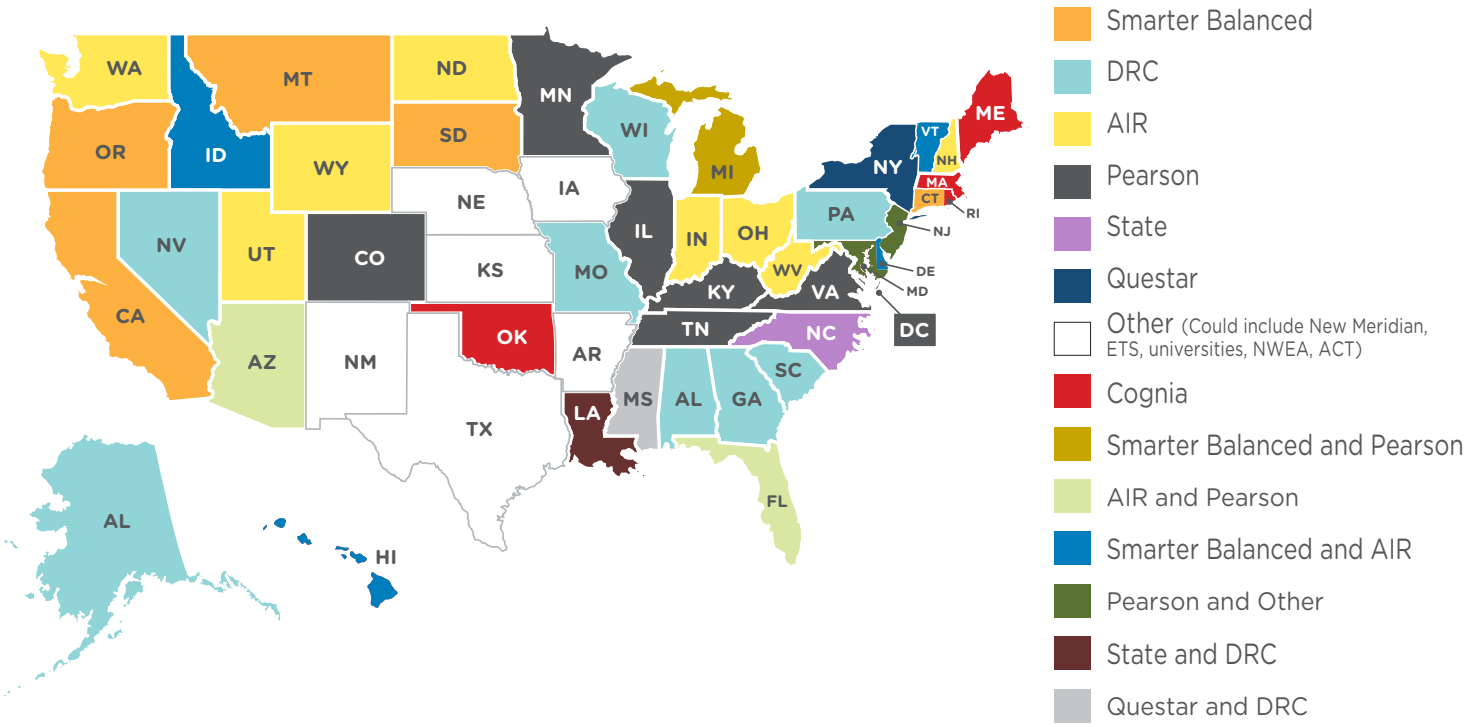
The administration's decision to leverage billions of dollars in federal funding on behalf of higher standards, harder tests, and test-based consequences for teachers brought the national teachers' unions and Tea Party conservatives to the barricades, if from opposite directions. The Tea Party and its Republican congressional allies condemned the Common Core as a federal usurpation of traditional local control in public education, even though state-facing organizations led the development of the new standards. The unions targeted the new teacher evaluation systems and the increase in teacher accountability they represented, spurred by rank-and-file members' outrage that their livelihoods suddenly depended on how well their students performed on brand new standards and tests.

The avalanche of opposition forced the Obama administration to retreat. In August 2015, U.S. Secretary of Education Arne Duncan announced his department's decision to grant states with NCLB waivers a one-year delay in incorporating student test scores into teacher evaluations. CCSSO and the Council of Great City Schools, representing the nation's large urban school districts, announced a joint project to throw "their collective weight behind an effort to reduce test-taking in public schools, while also holding fast to key annual standardized assessments."

In December of that year, the president signed ESSA, which replaced NCLB. Democratic leaders fought hard to keep the requirement for annual state testing of every student in reading and math in seven grades and results disaggregated by race, income, English-learner status, and disability status.

The new federal law gave states and districts far more power to craft their own education

Figure 1. Testing Vendors for Grades 3–8 by State, 2019



Source: Lynn Olson, "The New Testing Landscape: How State Assessments Are Changing under the Federal Every Student Succeeds Act" (Washington, DC: FutureEd, September 2019), 4.

solutions. It abandoned NCLB's requirement that states impose escalating sanctions on underperforming schools. It jettisoned Duncan's earlier push for states and school districts to use students' standardized test scores in teacher evaluations. It permitted states to use college-admissions exams—the SAT and the ACT—as substitutes for state high school standardized tests. And it allowed states to explore new ways of testing students under an Innovative Assessment Demonstration Authority.

The New Testing Landscape

To get a handle on the state of state summative assessments under ESSA, FutureEd conducted a scan of state testing programs across the 50 states and Washington, D.C., based on publicly available data and interviews.[3] We found that the marketplace for summative assessments in grades 3-8 has fragmented as an increasing number of states met ESSA's accountability requirements with tests designed to reflect their individual state's content standards (figure 1). Many states abandoned the two Common Core–aligned state assessment consortia. In 2010, 45 states planned to use tests developed by the consortia. By the 2019–20 school year, only a dozen states remained in the Smarter Balanced consortium, and only D.C. was using PARCC exams.

At the high school level, more states are embracing the ACT and the SAT as their assessments despite concerns that the tests are not fully aligned with state standards. Language permitting this option under ESSA has accelerated the trend, enabling states to offer tests that parents and students actually care about. (It will be interesting to see if this trend continues now that the University of California and other higher education institutions are no longer requiring the exams for admissions.)

While states and the federal government still prioritize the ability to compare student test results within a state, comparing results across states has faded as a dominant goal. That said, because of PARCC and Smarter Balanced, there's evidence that both the quality and rigor

Because of PARCC and Smarter Balanced, there's evidence that both the quality and rigor of state tests have risen.

of state tests have risen. Cut scores—the benchmarks that states set for student performance—appear to be holding the line in most places, but it is harder to get a handle on test quality.

In this atmosphere, churn in state testing systems is a big concern. While no one is tracking how many states have changed their tests or assessment vendors multiple times in the past five years, many of those interviewed mentioned it. Constant changes in state assessment systems make it harder to track performance over time, create problems for state and district accountability systems, and send mixed messages to educators, diminishing their morale and ability to focus instruction.

Growing Interest in Innovation

States have shown growing interest in assessment systems that better reflect and support the daily work of students and teachers.

At the same time, states have shown growing interest in designing assessment systems that better reflect and support the daily work of students and teachers in classrooms. Such systems would deliver faster turnaround of test results, as well as greater use of end-of-unit tests, performance-based tasks that ask students to apply what they know and can do, and tests that are more closely linked to the curriculum. Such efforts could provide better ongoing information about student progress, while giving teachers more guidance on how to adjust instruction. To date, five states—Georgia, Louisiana, New Hampshire, Massachusetts, and North Carolina—have been approved for the federal innovative assessment pilot (also see article, p. 40).

Louisiana, for example, is developing ELA/social studies tests to be given three times a year at the end of units that can be rolled up into an end-of-year summative score. The online assessments will enable districts to choose from a set of texts organized around key topics. Students will respond to writing tasks that require them to make meaning of texts they have already studied or of texts that are new to them but closely related to the curriculum. This dovetails with the state's efforts in the past decade to adopt higher standards and an optional text- and content-rich ELA curriculum (ELA Guidebooks 2.0) with units built around general themes, knowledge domains, and "anchor texts."

New Hampshire is taking a different approach: piloting classroom-embedded performance tasks in English language arts, math, and science to replace annual state tests with the goal of providing students a richer, more individualized learning experience. Schools and districts participating in the Granite State's innovative assessment pilot—Performance Assessment of Competency Education, or PACE—supplant much of the traditional end-of-year state testing with teacher-developed performance tasks. These include one common task in each grade and subject combination without a state test, which participating districts agree to collaboratively develop and administer.

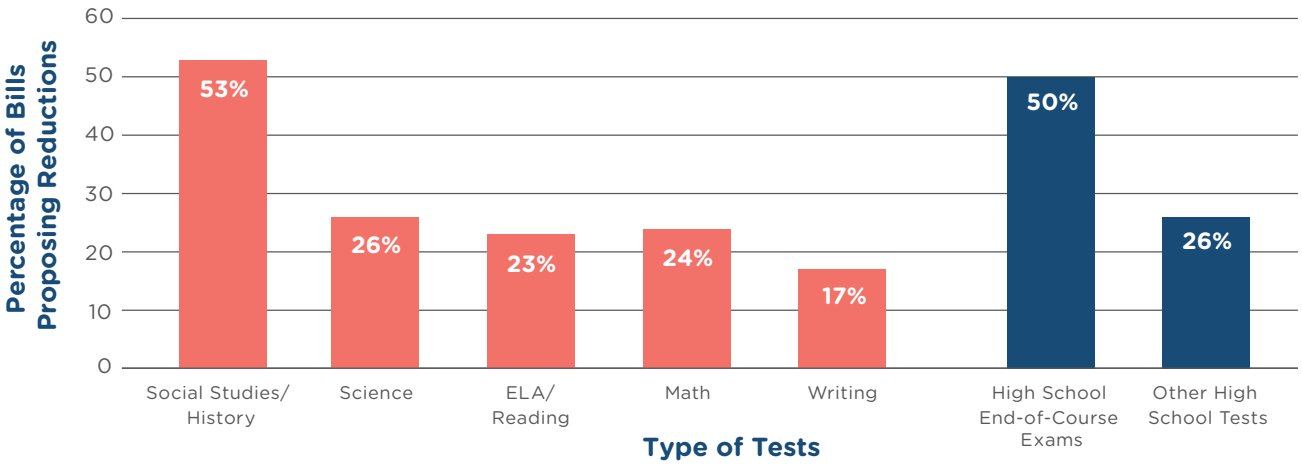
Legislative Response to Overtesting Concerns

Given the testing climate in recent years, ESSA has become a bulwark against further reductions in the measurement of school performance, even as Secretary DeVos suspended the law's requirements for 2019–20. But a close analysis of the political landscape of standardized testing makes clear that unless a new generation of tests can play a more meaningful role in classroom instruction, and unless testing proponents can again convince policymakers and the public of the value of state testing for school improvement and educational equity, annual state tests and the safeguards they provide are at risk.

From 2014 through 2019, lawmakers introduced no fewer than 426 bills and 20 resolutions in 44 states in response to critics' claims of overtesting (figure 2), and measures were adopted or enacted in 36 states.[4] There were more bills in 2019 than in 2018, an indication that antitesting sentiment remains strong in state capitals five years after the signing of ESSA. This analysis excludes dozens of parental "opt-out" bills that in most instances granted students unrestricted rights to sit out state tests. And it doesn't reflect moves to reduce testing in many states in recent years by governors, state boards of education, and state education agencies.

Lawmakers' most common legislative response was to reduce the number of state tests students must take. In other instances, they shortened the length of tests, capped standardized testing time in schools, required public reporting of testing time, or directed

Figure 2. Targets of Measures to Reduce State Testing, 2014–19



Source: Lynn Olson and Craig Jerald, "The Big Test: The Future of Statewide Standardized Assessments" (Washington, DC: FutureEd, April 2020), 10.

state agencies or local school districts to limit testing. A quarter of the 167 bills to reduce state testing demanded the discontinuation of every test not required by ESSA. Others targeted tests in grades and subjects not covered by federal law—particularly social studies. More than half the test-reduction legislation involved at least some social studies tests or high school exams in social sciences such as history. And half targeted high school end-of-course tests.

FutureEd's analysis confirmed the bipartisan nature of the legislative action against standardized testing. Sixty-eight of the measures introduced in 2019 were sponsored by individuals rather than legislative committees. Of those, Republican legislators authored 41 percent, Democrats authored 44 percent, and 15 percent were bipartisan.

Testing amid COVID

Given the months of time spent outside of school buildings during the 2019–20 school year, the most immediate concern this fall is how to capture where students are in their learning. One study estimated that the average student could lose roughly half of what they were expected to learn in math during the 2019–20 school year and close to a third of what they would have gained in reading.[5]

As schools reopen, educators should turn to diagnostic assessments and other measures of reading and math proficiency to help teachers identify individual student needs and know where to pick up instruction. These early-year assessments would be best kept closely connected to the school's curriculum and used solely to inform instruction and allocate resources.[6]

In the spring, states should restart their annual assessments mandated by ESSA. Just as widespread coronavirus testing will guide a return to normal life, state testing systems have a valuable role to play in helping leaders map education strategy, track progress, and back the nation's neediest students. But if federal waivers permit, states should not use these results for high-stakes accountability decisions. Attaching stakes to test results too quickly—as schools, teachers, and students are readjusting to the "new normal"—would play into the hands of accountability opponents at a time when we need smart testing more than ever.

Deploying high-quality assessments to support the nation's students and teachers and to inform policymaking in the wake of the COVID-19 crisis is one of the most significant challenges facing the education sector. I hope the articles in this issue of the *Standard* will help guide state policymakers as they face critical decisions about testing. In addition, FutureEd

cont'd on page 42

State testing systems have a valuable role to play in helping leaders map education strategy, track progress, and back the nation's neediest students.

Lynn Olson is a senior fellow at FutureEd.

With IADA approval in June 2019, North Carolina will expand this program from a first-year sample of participating students in through-grade assessments for mathematics and ELA to two districts participating in the second year, and a sample of at least 15 percent of students statewide in years three and four.

Early on, the task force engaged stakeholders as members and discussants, which provided vital input on the use, development, and technical requirements for a more balanced assessment system.[10]

Massachusetts

Massachusetts received its waiver from federal assessment requirements in April 2020, and state leaders plan to focus on increasing access to deeper learning, piloting an assessment system in science and technology/engineering for grades 5 and 8.[11] The new design will combine the current Massachusetts Comprehensive Assessment System (MCAS) with a new hands-on session, where students will be assessed in dynamic, interactive simulations, much as they might experience a task in a science class. "We want to make sure the assessment reflects the kinds of tasks that we're asking teachers to give their students," said Sam Ribnick, special advisor for innovative assessments and data at the Massachusetts Department of Elementary and Secondary Education.

The IADA pilot will expand another Massachusetts pilot, the Kaleidoscope Collective for Learning, in which select schools and districts organized classroom instruction around deeper learning. About half of the IADA pilot group are Kaleidoscope schools.

Its pilot is part of a broader state effort to reduce disparities and make deeper learning more ingrained into the instruction all students receive. "Any state that's getting into this process has to be really clear on what are the goals and the intended purpose of doing this and let that drive the initial conversation about what the new design will look like," Ribnick said. ■

[1]Office of Elementary and Secondary Education, "Innovative Assessment Demonstration Authority," webpage (Washington, DC: U.S. Department of Education, 2020), https://oese.ed.gov/offices/office-of-formula-grants/school-support-and-accountability/iada/.

[2]Louisiana Believes, "Louisiana Innovative Assessment Pilot," fact sheet (Baton Rouge, LA: Louisiana Department of Education, N.d.), https://www.louisianabelieves.com/docs/default-source/key-initiatives/louisianas-key-initiatives---innovative-assessment-pilot.pdf?sfvrsn=a6219f1f_18.

[3]Daniel T. Willingham, "Knowledge and Practice: The Real Keys to Critical Thinking," issue brief (Knowledge Matters, March 2016), http://knowledgematterscampaign.org/wp-content/uploads/2016/05/Willingham-brief.pdf.

[4]New Hampshire Department of Education, "New Hampshire: Application for the New Authorities under the Innovative Assessment Demonstration Authority" (Concord, NH: author, 2018), https://www.education.nh.gov/sites/g/files/ehbemt326/files/files/inline-documents/nhpaceapplication.pdf.

[5]New Hampshire Department of Education, "Performance Assessment of Competency Education," webpage (Concord, NH: author, N.d.), https://www.education.nh.gov/who-we-are/division-of-learner-support/bureau-of-instructional-support/performance-assessment-for-competency-education.

[6]Office of Elementary and Secondary Education, Georgia's IADA approval letter (Washington, DC: U.S. Department of Education, 2019), https://www2.ed.gov/admins/lead/account/iada/gaiadaapproval2019.pdf.

[7]Ibid.

[8]Office of Elementary and Secondary Education, "North Carolina: Application for New Authorities under the Innovative Assessment Demonstration Authority" (Washington, DC: U.S. Department of Education, 2018), https://www2.ed.gov/admins/lead/account/iada/nciadaappdec2018.pdf.

[9]North Carolina Department of Public Instruction, "Items that Require Additional Information or Revision in North Carolina's Innovative Assessment Demonstration Authority Plan" (Raleigh, NC: author, April 2019), https://www2.ed.gov/admins/lead/account/iada/ncapplicationaddendum.pdf.

[10]Office of Elementary and Secondary Education, "NC: Application for New Authorities."

[11]U.S. Department of Education, "Massachusetts Becomes First State to Qualify to Test New and Innovative Ways to Assess Student Achievement Next School Year," press release (Washington, DC: author, 2020), https://www.ed.gov/news/press-releases/massachusetts-becomes-first-state-qualify-test-new-and-innovative-ways-assess-student-achievement-next-school-year.

---

cont'd from page 11...A Shifting Landscape

recently published "Blueprint for Testing: How Schools Should Assess Students During the Covid Crisis," drawing on recent work, to provide a ready reference for busy education leaders and policymakers as they plan for standardized testing now and into the future.[7] ■

[1]Lynn Olson and Craig Jerald, "The Big Test: The Future of State Standardized Assessment" (Washington, DC: FutureEd, April 2020).

[2]Deepa Srikantaiah, "How State and Federal Accountability Policies Have Influenced Curriculum and Instruction in Three States: Common Findings from Rhode Island, Illinois, and Washington" (Washington, DC: Center on Education Policy, October 2009).

[3]Lynn Olson, "The New Testing Landscape: State Assessment under the Every Student Succeeds Act" (Washington, DC: FutureEd, September 2019).

[4]Olson and Jerald, "Big Test."

[5]Megan Kuhfeld and Beth Tarasaw, "The Covid Slide: What Summer Learning Loss Can Tell Us about the Potential Impact of School Closures on Academic Achievement" (Portland, OR: NWEA, April 2020).

[6]Thomas Toch, "Don't Abandon Standardized Testing in Schools Next Year: Rethink It," op-ed, *The Hill* (May 26, 2020).

[7]Lynn Olson, "Blueprint for Testing: How Schools Should Assess Students during the Covid Crisis," (Washington, DC: FutureEd, July 2020).

Joseph Hedger is NASBE's associate editor.