

THE BIG TEST

THE FUTURE OF STATEWIDE STANDARDIZED ASSESSMENTS

BY LYNN OLSON AND CRAIG JERALD

APRIL 2020

FutureEd
GEORGETOWN UNIVERSITY

About the Authors

Lynn Olson is a FutureEd senior fellow.
Craig Jerald is an education analyst.

About FutureEd

FutureEd is an independent, solution-oriented think tank at Georgetown University's McCourt School of Public Policy, committed to bringing fresh energy to the causes of excellence, equity, and efficiency in K-12 and higher education. Follow us on Twitter at @FutureEdGU.

Use

The non-commercial use, reproduction, and distribution of this report is permitted.

© 2020 FutureEd

THE BIG TEST

THE FUTURE OF STATEWIDE STANDARDIZED ASSESSMENTS

BY LYNN OLSON AND CRAIG JERALD

APRIL 2020

FutureEd
GEORGETOWN UNIVERSITY

FOREWORD

School reformers and state and federal policymakers turned to standardized testing over the years to get a clearer sense of the return on a national investment in public education that reached \$680 billion in 2018-19.

They embraced testing to spur school improvement and to ensure the educational needs of traditionally underserved students were being met. Testing was a way to highlight performance gaps among student groups, compare achievement objectively across states, districts, and schools, and identify needed adjustments to instructional programs.

But a widespread backlash against standardized testing has left the future of statewide assessments and the contributions they make in doubt. This report by Senior Fellow Lynn Olson and education analyst Craig Jerald examines the nature and scope of the anti-testing movement, its origins, and how state testing systems must change to survive. It is clear that if state testing systems do not evolve in significant ways Congress may abandon the statewide standardized testing requirements in the federal Every Student Succeeds Act when it next reauthorizes the law.

The report includes a comprehensive new analysis of state legislative action on testing from 2014 through 2019 that Jerald conducted with the help of Olson, Brooke LePage, Phyllis Jordan, and Rachel Grich. Molly Breen edited the report. Merry Alderman designed it. The report is the latest in a series of FutureEd initiatives on the future of standardized testing. We are grateful to the Bill & Melinda Gates Foundation for funding the project.

Thomas Toch

Director

In February, before the coronavirus pandemic upended the nation's education system, Georgia's Republican governor, Brian P. Kemp, announced plans to cut state-mandated high school end-of-course tests in half, confine state elementary and middle school testing to the last five weeks of the school year, shorten tests, and encourage school districts to work with the state's education agency to reduce local testing. "When you look at the big picture, it's clear," Kemp declared. "Georgia simply tests too much."

Governor Kemp's announcement is the latest expression of a backlash against standardized testing in public education that has left the future of state testing, a cornerstone of school reform for nearly three decades, in doubt.

Pressure to reduce testing has come from many, often confounding sources: teachers' unions and their progressive allies opposed to test-based consequences for schools and teachers; conservatives opposed to what they consider an inappropriate federal role in testing; suburban parents who have rallied against tests they believe overly stress their children and narrow instruction; and educators who support testing but don't believe current regimes are sufficiently helpful given how much teaching time they consume.

School reformers and state and federal policymakers turned to standardized testing over the years to get a clearer sense of the return on a national investment in public education that reached \$680 billion last year, to spur school improvement, and to ensure the educational needs of traditionally underserved students were being met. Testing was a way to highlight performance gaps among student groups, compare achievement objectively across states, districts, and schools, and identify needed adjustments to instructional programs.

But the pushback against testing in recent years—fueled by union communications and lobbying campaigns, right wing media personalities, and misconceptions about the extent of state testing—has led to a substantial retreat on testing among state policymakers. A new national analysis by FutureEd has found that between 2014 and 2019, lawmakers in 36 states passed legislation to respond to the testing backlash, including reducing testing in a variety of ways, a direction also taken by many state boards of education and state education agencies.

Likely Democratic presidential nominee Joe Biden expressed concerns about standardized testing during a televised forum in October. And in the wake of mass school closures due to the coronavirus pandemic, U.S. Secretary of Education Betsy DeVos announced on March 20 that her department would waive federal standardized testing requirements for the 2019-20 school year in states requesting testing relief. The move, and the consequent loss of a year's worth of longitudinal data, could further reduce the standing of state tests. The chances are strong that there won't be any end-of-year state testing anywhere in the nation for the first time in half a century. Already, teacher union leaders, like the president of the Massachusetts

Teachers Association, are signaling the suspension could provide an opening to cancel state tests permanently.¹

Given the testing climate in recent years, the federal Every Student Succeeds Act has become a bulwark against further reductions in the measurement of school performance, even as DeVos suspends the law's requirements for 2019-20. The law requires that every state test every student in seven grades annually and report the results in a way that bars school districts from masking the performance of disadvantaged students.

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 reconvince policymakers and the public that state testing is an important ingredient of school improvement and integral to advancing educational equity, annual state tests and the safeguards they provide are clearly at risk.

Toward Transparency

Before the 1960s, states had scant information about how well their students were performing. But in 1965, as part of the War on Poverty, the federal Elementary and Secondary Education Act poured millions of dollars into schools and required studies of the impact of those funds. In 1969, Congress mandated a federally funded snapshot of student performance, the National Assessment of Educational Progress.

State governments, meanwhile, began requiring tests to determine if students were making progress in core subjects. Michigan launched the first statewide testing program, also in 1969. But there were no state achievement standards for how well students should perform and no explicit consequences for schools if students performed poorly.

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 students graduated with the requisite basic skills. By the 1980s, there was growing alarm about student performance as the nation transitioned to a knowledge-based economy. Gaps in the

achievement of long-neglected student populations gained prominence in the wake of the civil rights movement. The concerns spurred the publication of *A Nation at Risk* and a host of other reform manifestos that urged a more demanding curriculum and more rigorous graduation requirements for every student.

In the ensuing years, policymakers became increasingly frustrated as local districts watered down those requirements with courses like "business math" and "the fundamentals of science." So, in the late 1980s and 1990s, Republican and Democratic presidents, as well as the nation's governors, began to push for higher educational standards and national goals, as well as more student testing and efforts to hold schools and districts accountable for results. The Clinton administration's reauthorization of the Elementary and Secondary Education Act 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.

But not all states responded to the requirements with equal rigor.² 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 federal 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 test them 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. The law effectively tripled the size of the state testing market over the six years after it was passed.³

The increased requirements reflected a belief that for every child in America to achieve high standards, schools needed to track the learning of every student every year against those standards and be held accountable for the results—a fight against what then-President Bush called "the soft bigotry of low expectations." National leaders simply didn't trust local educators to do the right thing for low-income

students and students of color, so they tried to force them to act, in part, by imposing far more transparency and accountability via testing.

“It has been clear to the civil rights community for a very long time that without comparable, annual statewide assessments, it is very difficult to know whether there is equal opportunity in education,” said Liz King, director of Education Policy for the Leadership Conference on Civil and Human Rights. “Even if you don’t care about equity or remedying discrimination, I think there are a lot of people who care about knowing whether the education system is working, whether expenditures are supporting the outcomes that families expect.”⁴

But if NCLB was designed to shed a bright light on educational inequities, states, districts, and schools frequently responded to that pressure in ways that jeopardized student learning and kindled anti-testing sentiment. Schools emphasized instruction in tested subjects at the expense of untested subjects and stressed test-taking skills. School districts piled on new benchmark tests to gauge how students would perform on end-of-year exams. States began to rely heavily on simplistic multiple-choice tests because they were

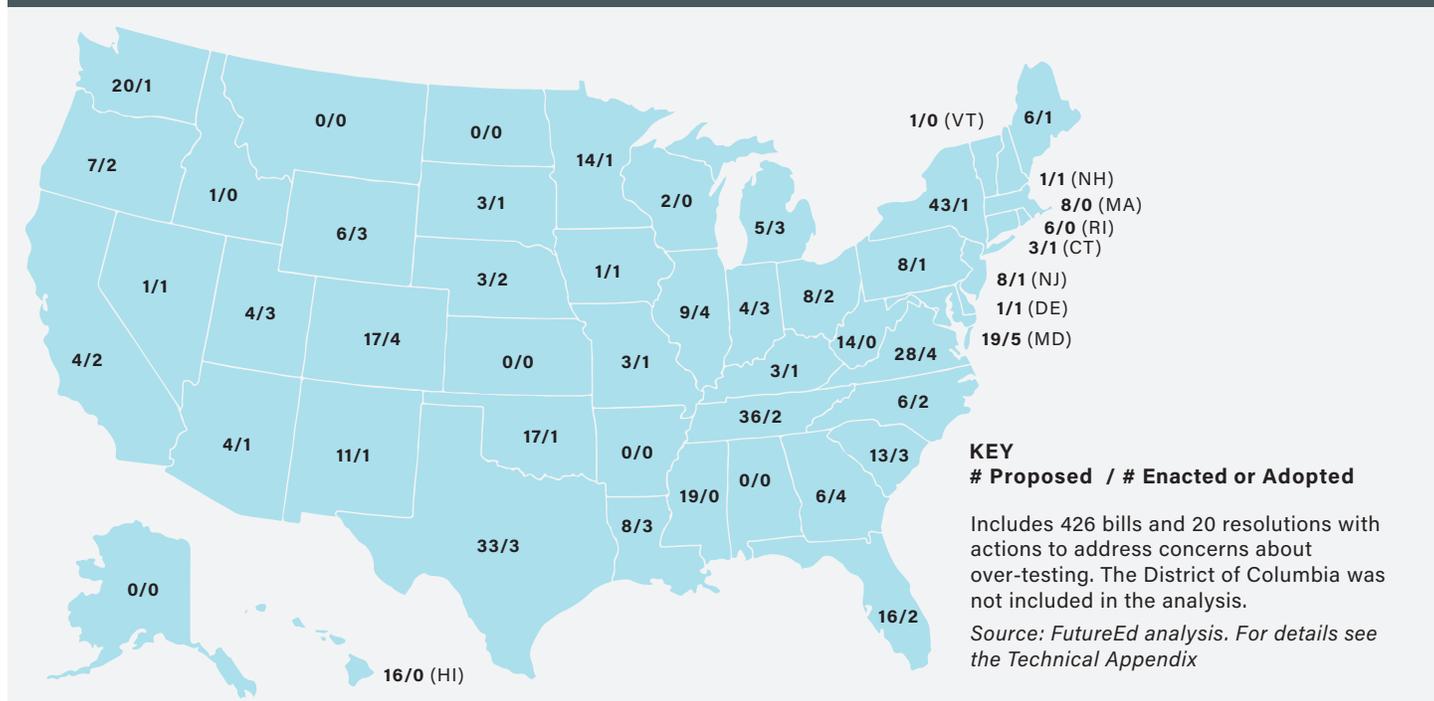
cheaper and easier to administer in the face of NCLB’s tight testing timelines. And many states lowered their testing standards to get more students over the proficiency bar.⁵

The Backlash Builds

Many states’ indifferent commitment to higher standards and counterproductive responses to NCLB led national leaders to push harder. In 1996, a bipartisan group of governors and business leaders had founded an organization called Achieve to help states ensure all students graduated high school ready for college and careers. In 2001, the same year that NCLB became law, a group of states began collaborating with Achieve to identify the “must have” knowledge and skills needed 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 to jointly develop demanding voluntary standards in English language arts and math—what became the Common Core State Standards.

As a result, the two organizations were finishing work on the Common Core standards at the same time the Obama

State Measures Introduced and Adopted to Address Over-Testing Concerns, 2014 through 2019



administration was drafting the Race to the Top initiative, which provided billions of dollars in education funding to states to help address the 2008-09 economic crisis.

Governors asked that states be allowed to use some of the federal funding to implement the Common Core and to develop aligned tests, an expensive undertaking. In the end, the \$4.35 billion Race to the Top competitive grant program incentivized states to adopt tougher academic standards and more rigorous tests aligned with those standards, by making the grants contingent on states adopting the reforms.

Spurred in part by the prospect of federal largesse, 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).

As the opt-out movement spread, generating a squall of sensational headlines, the U.S. Department of Education was compelled to warn a dozen states that they risked federal sanctions for not having enough students take their standardized tests.

The Obama administration also made competitive Race to the Top grants contingent on states creating new teacher evaluation systems that judged teachers “in significant part” on the basis of their students’ test scores, even though the tougher standards and tests wouldn’t be fully in place before the new evaluation systems were launched—a demand that Kate Walsh, the director of the National Council on Teacher Quality and a proponent of the move at the time, now calls “a strategic blunder.”

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 as state 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 outraged that their livelihoods suddenly depended on how well their students performed on brand new standards and tests.

In 2014, the 3-million-member National Education Association launched a national “Campaign to End ‘Toxic Testing’” that would “seek to end the abuse and overuse of high-stakes standardized tests and reduce the amount of student and instructional time consumed by them.”⁶

The NEA and the American Federation of Teachers pumped millions of dollars into state lobbying campaigns. They channeled money to outside organizations like FairTest to attack testing. And they launched grassroots efforts to encourage parents and their children to boycott state testing.

In 2015, the NEA’s Maryland affiliate launched a “Less Testing, More Learning” campaign to scale back testing and water down accountability—a lobbying blitz that included 50,000 e-mails, 4,000 phone calls, and 2,000 letters and postcards from MSEA members to their representatives, working alongside groups like the PTA, the ACLU, and the NAACP, organizations that have championed the reporting of test data by race and class to uncover educational inequities, but have opposed the use of state tests to make high-stakes decisions about schools and students, often out of concern about racial bias in standardized testing.⁷

“I would urge parents...to opt out of testing,” Karen Magee, the then-president of New York State United Teachers, told an Albany public affairs show in 2015, after thousands of vocal parents and students, many of them in more-affluent suburban school districts, refused to take New York’s standardized tests on the grounds that the tests were too long, too stressful, and sidetracked instruction.⁸

As the opt-out movement spread, generating a squall of sensational headlines, the U.S. Department of Education was compelled to warn a dozen states that they risked federal

sanctions for not having enough students take their standardized tests.

At the same time, several book-length critiques of testing added fuel to the anti-assessment fires, including *The Test: Why Our Schools Are Obsessed with Standardized Tests*, by Anya Kamenetz, an NPR reporter; *The Testing Charade*, by Daniel Koretz, a professor emeritus at the Harvard Graduate School of Education; and *Beyond Test Scores*, by Jack Schneider, an assistant professor in the college of education at University of Massachusetts-Lowell.

The Obama Administration Retreats

The avalanche of opposition forced the Obama administration to retreat on testing. In August 2015, U.S. Secretary of Education Arne Duncan announced in the blog post “Listening to Teachers on Testing” his department’s decision to grant states with NCLB waivers a one-year delay in incorporating scores into teacher evaluations. Duncan, who had pressed for test scores to be part of Race to the Top teacher evaluations, said he shared teachers’ concerns that “testing—and test preparation—takes up too much time.”⁹

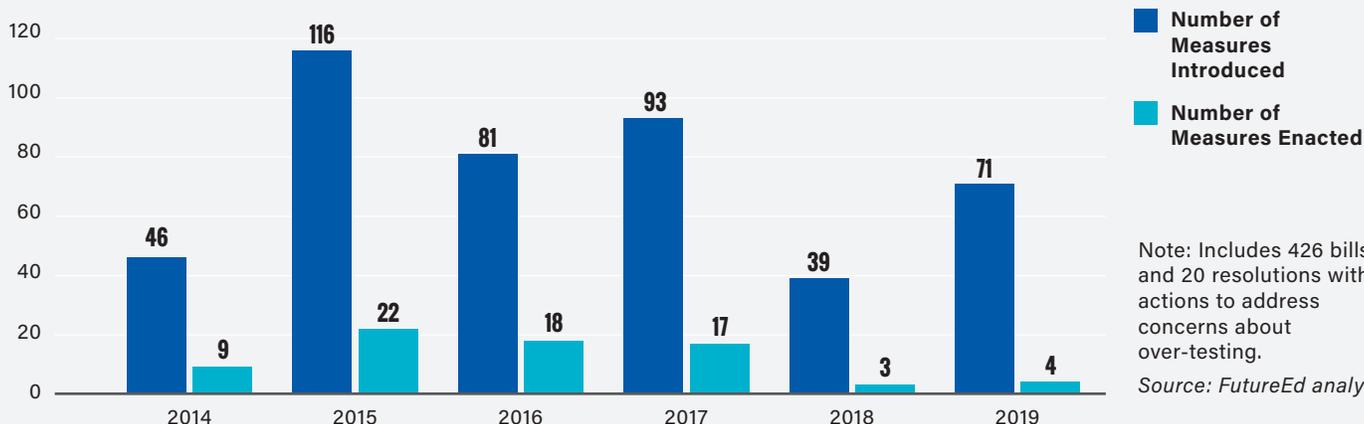
Even so, testing opposition remained strong, and the White House pressed for a more substantial response after the president told his advisors that testing was coming up in conversations as he traveled the country. In October, the administration announced a Testing Action Plan “to correct

the balance” between the “vital role that good assessment plays ... while providing help in unwinding practices that have burdened classroom time or not served students or educators well.”¹⁰ The administration announced grants for state and local testing audits, based on the recognition that state- and district-required tests, many of them not mandated by the federal government, had piled up over time and lost their strategic value. The administration also recommended that states cap the percentage of instructional time students spend taking required statewide standardized assessments, “to ensure that no child spends more than 2 percent of her classroom time taking these tests.”¹¹

That same month, the Council of Chief State School Officers 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.”¹²

In December, the president signed the federal Every Student Succeeds Act (ESSA), which replaced the No Child Left Behind Act. During congressional negotiations, Democratic leaders of the House and Senate education committees—Virginia Rep. Bobby Scott, the ranking Democrat on the House Education and Workforce Committee and Washington Sen. Patty Murray, the ranking Democrat on the Senate Health, Education, Labor and Pensions Committee (HELP)—demanded that the new law preserve NCLB’s annual state testing of every student in reading and math in seven grades

State Legislative Activity on Over-Testing Concerns, 2014 through 2019



and that states break down each school’s results by student race, income, English language-learner status, and disability status—a demand to which Tennessee Sen. Lamar Alexander, Republican chairman of the Senate HELP committee, eventually acceded. Their support proved crucial to a wing of the civil rights community that, since the drafting of NCLB, had advocated transparency as the strongest defense of neglected student populations.

The new federal law gave states and districts far more power to craft their own education solutions than they had wielded under NCLB. 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, even though the college admissions measures weren’t necessarily aligned to states’ high school standards. And it allowed states to explore new ways of testing students under an Innovative Assessment Demonstration Authority.

State Legislators React

State legislators responded to the testing backlash and the easing of federal mandates with a wave of legislation to restrict testing within their borders. FutureEd conducted a comprehensive analysis of state legislation and resolutions

from 2014 through 2019 and found a striking number of measures to shrink the standardized testing footprint in schools. (See Technical Appendix for details.)

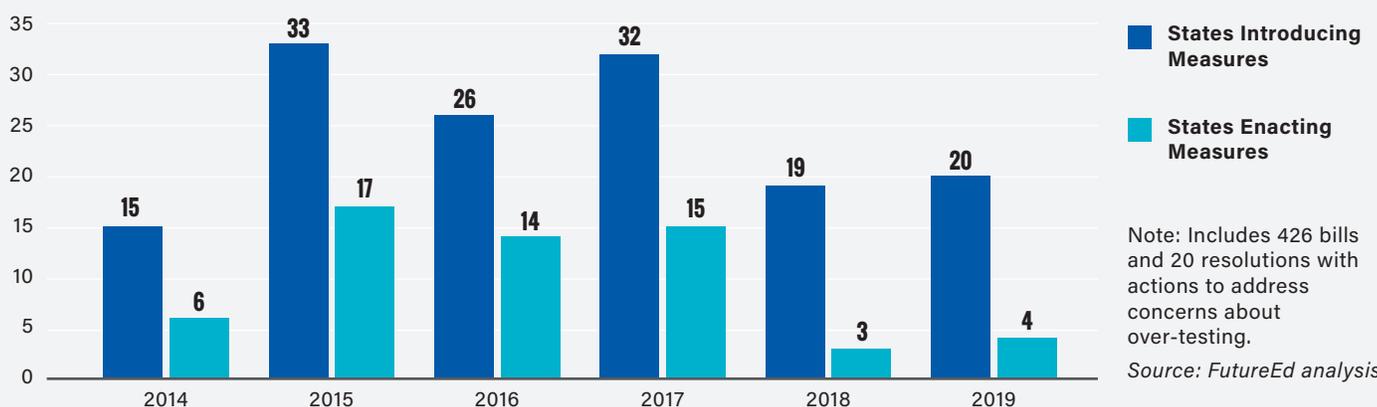
Lawmakers introduced no fewer than 426 bills and 20 resolutions in 44 states in response to critics’ claims of over-testing, and measures were adopted or enacted in 36 states. There were more bills in 2019 than in 2018, an indication that anti-testing sentiment remains strong in state capitals five years after the signing of the Every Student Succeeds Act. This is especially true in southern states such as Mississippi, South Carolina, Tennessee, Texas and Brian Kemp’s Georgia.

The FutureEd 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 departments of education.

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 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 the federal ESSA and the NCLB before it. Others targeted tests in grades and subjects not covered by

State Legislative Activity on Over-Testing Concerns, 2014 through 2019



ESSA—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.

North Carolina's Testing Reduction Act of 2019, signed into law by Democratic Governor Roy Cooper, captures today's state testing climate. The legislation eliminates nearly two dozen statewide exams used primarily for teacher evaluations. It directs the North Carolina department of education to report the numbers and types of state and local tests every year and publish state and local testing types/calendars on the agency's website. It compels local school boards to cut local standardized testing if the number of tests or hours of testing exceed the state's two-year average. And it establishes the intent of the North Carolina legislature to move away from a single, long testing event at the end of the year and toward multiple, short testing events throughout the year.

Policymakers' pushback against testing is reflected in recent declines in state testing sales after years of expansion. The market shrank by 3.5 percent from 2017-18 to 2018-19, according to a recent analysis by Simba Information, a marketing research firm.¹³ "The pendulum has swung from developing new tests to reducing the number of exams that are required and shortening the length of existing exams," according to Simba.¹⁴

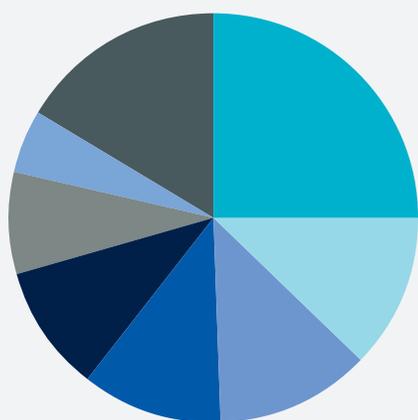
Bipartisan Opposition

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.

Despite its Republican-dominated legislature, Mississippi lawmakers from both parties in 2019 proposed a slew of strategies to reduce high school testing. Two bills introduced in the state senate, one by a Democrat and another by a Republican, offered nearly identical language to limit state and local testing to the minimum required under federal law, among other things.

In some states, like Georgia, pressure to reduce testing has come from Tea-Party conservatives. State School Superintendent Richard Woods ran on an anti-Common Core platform. In other states, legislative actions came in direct response to union pressure. Following union president Karen Magee's urging of New York parents to boycott the state's tests in 2015, some 200,000 students—roughly one in five—opted out, the largest proportion of students in the country. By the end of the year, a task force convened by Governor Andrew Cuomo recommended a delay in using the tests in teacher evaluations. New York State has had the most

Strategies to Address Over-Testing in State Laws, 2014 through 2019



25%
Reduce Number of Tests

11%
Multiple Actions

6%
Adopt Multi-Purpose Tests

12%
Shorten Tests

10%
Study or Publicly Report
on Testing Time

16%
Other

12%
Cap Testing Time

8%
Recommend Ways to
Reduce Testing

Note: Includes 67 bills enacted and 6 resolutions adopted with actions to address concerns about over-testing. "Other" includes limiting testing in grades K-2, giving districts flexibility to minimize testing, waiving tests for students who meet other criteria, capping state testing, requiring districts to take action to limit testing, and delaying or suspending testing due to concerns about over-testing.

Source: FutureEd analysis

legislation designed to address over-testing introduced in the nation in the past six years—43 bills.

The Maryland State Education Association’s lobbying also paid off, in 2017, when Republican Governor Larry Hogan signed the “More Learning, Less Testing Act” into law. It requires school boards and unions to set time limits for federal, state, and local standardized tests and caps testing at 2.2 percent of the school calendar for districts that fail to reach agreements. The union claimed the legislation would eliminate 730 hours of standardized testing across 17 school districts annually.¹⁵

In New Jersey, traditional civil rights groups, the New Jersey Education Association, and opt-out groups, like Save our Schools New Jersey, have joined forces to push the state to abandon PARCC testing, reduce the number of end-of-course tests in high school, and eliminate exit tests as a high school graduation requirement. Democratic Governor Phillip Dunton Murphy replaced Republican Chris Christie in January 2019, after pledging during his campaign to get rid of both PARCC and the high school graduation requirement. New Jersey is one of only a dozen states to still require a high school exit exam.

In addition, some states have backed away from using student test scores to rate teachers since the 2015 passage of ESSA. The National Council on Teacher Quality reports that 26 states now use results from state standardized tests as part of their teacher evaluation systems, down from 37 states in 2015.¹⁶

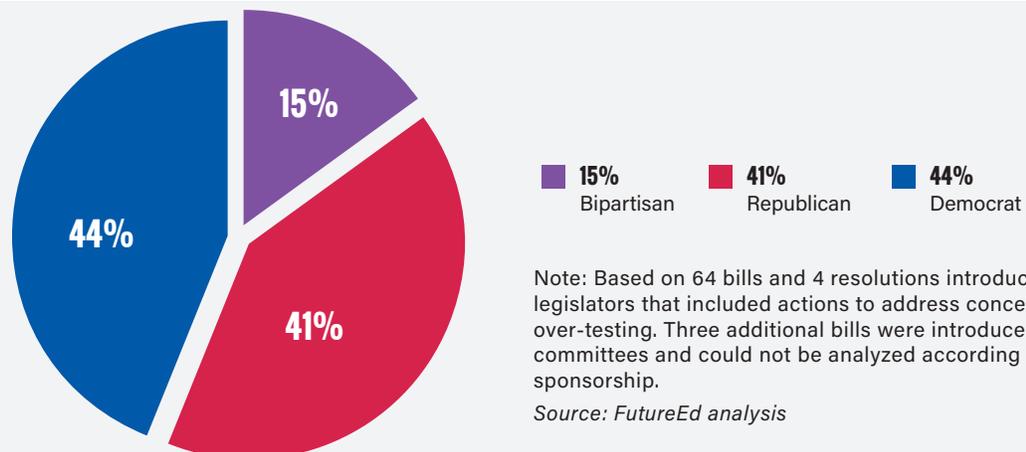
Teachers and Teacher Unions

Teachers’ take on testing is complex. In general, teachers want measures of student progress and support standardized testing, but many don’t think the existing tests are doing the job, particularly when it comes to informing instruction and reflecting their curriculum.

Teacher union leaders are forthright about not wanting their teachers held accountable for their students’ achievement on standardized tests, and about their opposition to high-stakes school accountability more generally. That’s why they often sided with the opt-out movement. “We have actually pushed for appropriate testing and we’ve pushed for reductions of extraneous tests, extraneous paperwork,” Randi Weingarten, the president of the American Federation of Teachers, says. “The reason we were part of the opt-out movement, in some places, was because of the high-stakes nature of standardized testing. It wasn’t a snapshot to see where kids were at any point in time ... Basically, there was a fixation on the teachers and the consequences for the teachers rather than a fixation on what children needed.”

Weingarten says she leans toward the type of sample testing in key grades that many high-performing countries use and away from testing every student every year, as required under ESSA—a move that would reduce transparency into school performance and eliminate the use of test results in teacher evaluations. The National Education Association advocated for such a shift during congressional drafting of ESSA.

Primary Sponsors of Measures Introduced in 2019 to Address Over-Testing Concerns



Surveyed directly, many rank-and-file teachers also express concerns about the usefulness of standardized tests in their classrooms.

Sixty-two percent of teachers responding to a 2015-16 survey by the Center on Education Policy said they spent too much time preparing students for state tests, and 81 percent said students spent too much time taking state and local tests.¹⁷ While 46 percent of teachers in a 2019 survey by the publication Education Next agreed that federal testing requirements should remain in place, 49 percent disagreed.¹⁸ In a national teacher poll released in January 2020 by Educators for Excellence, fewer than half of the respondents described tests administered by their state, district, or charter management organization as “very accurate” when asked if the tests aligned with their classroom curriculum.¹⁹

Still, teachers support standardized testing. “Many teachers would prefer to cut the frequency and length of state- and district-mandated tests rather than eliminate them altogether,” the Center on Education Policy survey found.²⁰ More than nine in 10 teachers in the 2020 Educators for Excellence survey responded that students should have a summative measure of their annual learning.

The public also supports standardized testing and the window into school and teacher performance it affords. Seventy-four percent of the respondents to Education Next’s 2019 survey approved of the federal government’s testing requirements. And a 2016 NWEA/Gallup poll found three-quarters of students themselves believed they spent either the right amount of time or too little time taking tests.²¹

But it’s clear in this climate that teacher unions are going to continue to wield their considerable influence against standardized tests unless the tests play a larger role in helping teachers be more effective—by improving their instruction and by helping them more fully understand students’ needs—and aren’t used merely for teacher and school accountability.

“This is the tension,” says Evan Stone, co-founder and co-chief executive of Educators for Excellence. “Teachers want to measure student learning. But they don’t have faith in our current state tests ... We need to fundamentally rethink how assessments can be integrated into the curriculum and into the flow of a school year.” As Adam Ezring, the deputy

director of the Collaborative for Student Success, puts it: “We’ve got to connect testing to improving instruction.”

A Tale of Two Tests

One reason for the perception that there’s too much state testing is that both parents and the public more generally often conflate state-mandated testing with the many interim, benchmark, and diagnostic tests that school districts deploy. Even as policymakers decrease time devoted to state testing, it’s difficult to rein in local assessments, which comprise an increasing share of the testing market, according to Simba.

“When we asked teachers, ‘What would you get rid of?’ they wanted to get rid of the state assessment because they didn’t find the state assessment immediately useful to their instruction,” says Rachel Cantor, executive director of Mississippi First, a state education advocacy group. “But the state assessments are not what actually takes all their time. It’s district assessments.” For their part, “Parents make no distinction whatsoever between state tests, district tests, teacher-made tests,” Cantor says. “It’s all just, ‘Why do you have so many tests?’”

Studies have found that the average classroom spends about 2 percent of instructional time taking mandated tests, a small fraction of the school year. But some schools and school districts spend much more, contributing to perceptions of “over-testing.” A 2014 study of a dozen urban districts by the nonprofit organization Teach Plus found that test-taking ranged from three days in one district to two weeks in another. Test preparation ranged from 16 to 30 school days. The report was especially critical of the time spent on low-quality local benchmark tests that were often misaligned with state standards.²²

An inventory of standardized testing in 66 urban districts the following year by the Council of Great City Schools found that testing was most prevalent in the eighth grade, where students spent just over 2 percent of school time taking standardized tests.²³ To the council’s executive director, Michael Casserly, the biggest problem wasn’t excessive testing time. It was results that weren’t coherent and useable: “I haven’t seen a lot of states, or anybody else, ask themselves

the question, “Does this portfolio of standardized tests we’re using really measure what we want to have measured?”

Yet the foundational argument for statewide testing—that the absence of objective insights into school performance leaves traditionally underserved students vulnerable—is no less valid today than at the outset of the school reform movement. And there’s a danger that shortening tests could reduce their quality. Ed Roeber, the executive director of the Michigan Assessment Consortium, says caps on testing time could cause policymakers to abandon constructed-response items that are truer measures of students’ abilities (for example by asking students to write), or to test only a narrow band of topics within a subject, as many states did under NCLB as they sought to save money and meet the law’s tight testing timelines. “If you cut a test down to two hours and you’ve got to cover the domain of math or English as defined by state standards,” Roeber says, “you’re going to have so few items on any standard that you’ll only get fairly global reports. Teachers gets frustrated” because such high-level results aren’t particularly helpful at the classroom level.

The Obama administration funded the development of the PARCC and Smart Balanced tests to address the mismatch

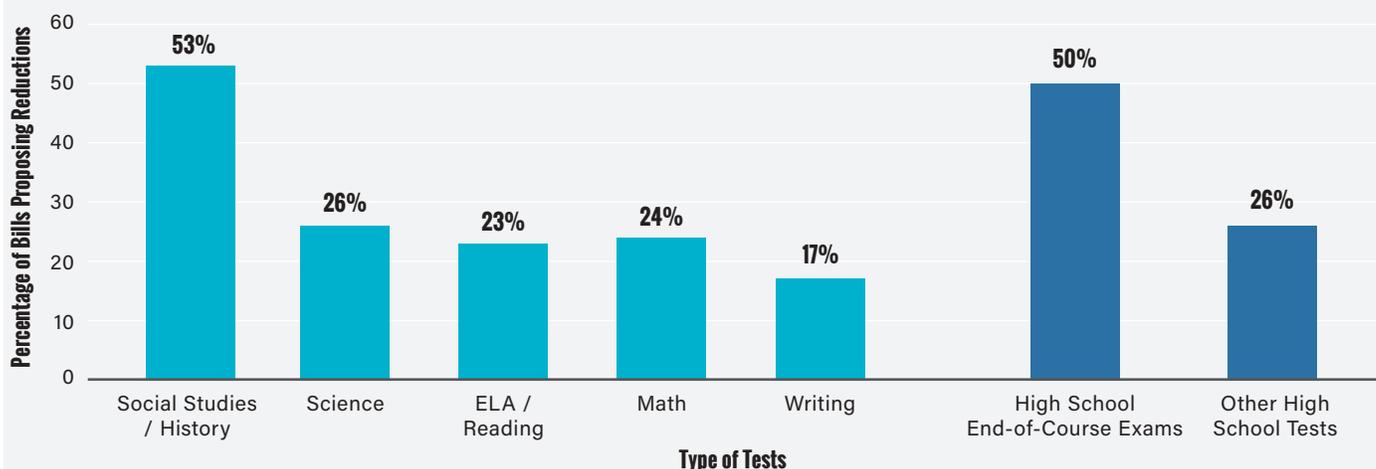
between higher standards embodied in the Common Core and the superficiality of the standardized tests that emerged in many states under NCLB, though the consortia tests were typically longer than the tests they replaced. So as states abandon the consortia assessments the risk of tests driving down classroom standards is reemerging. “We know what gets tested is likely what gets taught,” Roeber says. “The corollary is ‘how it gets tested is how it gets taught.’”

A New Testing Landscape

The contours of a new generation of statewide standardized testing that promotes both accountability and instructional improvement has emerged from the testing debates of the past several years and from teacher surveys.

Teachers identify “capturing student learning over time, rather than a single snapshot at the end of one year” as a key way to make standardized testing more useful.²⁴ Georgia, Louisiana, New Hampshire, and North Carolina are piloting moves in that direction under the federal Innovative Assessment Demonstration Authority, exploring ways to give more frequent, instructionally useful assessments that can

Targets of Measures to Reduce State Testing, 2014 through 2019



Note: Based on 167 bills proposing to eliminate one or more state-required tests. The subject areas include any testing that would be eliminated or reduced by the bill in question at any grade level. For example, “Math” includes any grade-level testing in math in K-12 as well as high school end-of-course exams in specific subjects such as Geometry or Algebra 2. “Writing” includes only stand-alone writing tests, not writing components of comprehensive ELA tests. For 22 of the 167 bills, it was not possible to determine precisely which state tests or types of testing would be eliminated based on available information.

Source: FutureEd analysis

roll up into a summative, year-end score. Similarly, Nebraska is working with the testing company NWEA to roll up the state's computer-adaptive tests, given periodically throughout the year, into a cumulative result. In January, the U.S. Department of Education proposed to expand experimentation by increasing federal funding for such work.

To Emma Vadehra, former chief of staff at the U.S. Department of Education during the Obama administration, that means developing more helpful tests, but also new measures to be used alongside tests in gauging school performance and supporting teachers' work. Teachers agree. More than seven in 10 teachers in the Educators for Excellence survey favored multiple measures of school performance that included district or state standardized tests.²⁵

The Every Student Succeeds Act incentivizes states to develop new measures by requiring them to include non-academic factors in judging schools' performance. So far, 36 states have added student absenteeism to their accountability systems, making it by far the most popular measure. A dozen states have moved toward measuring school climate and student engagement through annual surveys of students, teachers and parents, in some cases holding schools accountable for the results. Though researchers warn that surveys shouldn't yet be deployed in that way, school climate and student engagement are emerging as important aspects of student success.

Meanwhile, the layering of many local tests on top of state testing regimes, and the conflating of state and local testing in the minds of many parents and policymakers, points to the importance of more effective coordination on standardized testing between state and local education leaders. This could help reduce the level of standardized testing in the nation's classrooms. In Roeber's words: "States, with the cooperation and collaboration of local districts, need to develop systems of assessment that balance the state [accountability] program with assessments that actually help kids learn."

Ary Amerikaner, a former Obama administration education official and now a vice president at the Education Trust, also points out that school improvement continues to be a component of federal standardized testing mandates under ESSA, and "that requires actually getting

the school-improvement side of state accountability systems right." Emerging tests and testing systems also need to be developed with much more teacher involvement than has been the case so far. Testing systems must also stay attuned to the needs of parents, who tend to value their children's progress more highly than the performance of the school systems they live in or even the schools their children attend.

A Race Against Time

There is not a tremendous amount of time to address these challenges.

The pincer movement from teacher unions on the left and the Tea Party on the right put federally mandated annual testing at risk during the 2015 drafting of the Every Student Succeeds Act. The president's personal commitment to the measure, the insistence of democratic congressional leaders, and the support of education organizations like the Education Trust and the Council of Great City Schools won the day on Capitol Hill. And that was after the Obama administration abandoned the mandatory use of test scores in teacher evaluation and agreed to let states decide what actions to take against schools that tested badly under ESSA.

Amerikaner of the Education Trust says those compromises, combined with calls for state and local testing audits, and other "release valves" written into ESSA may reduce demands to abandon the law's annual testing requirement in the future.

Maybe. But the Georgia governor's recent move to slash the state's testing system suggests the issue is very much alive among conservative policymakers. Testing critics on the left have kept up a relentless drumbeat against high-stakes testing, supported in recent months by progressive democratic presidential candidates Bernie Sanders and Elizabeth Warren, as well as by Biden. The polemicist Diane Ravitch recently published *Slaying Goliath*, the latest in a series of book-length attacks on school reform, in which she claims that the federal testing requirements are part of a "decades-long campaign by right-wingers who hate public schools."²⁶

Given the political gridlock in Washington and the current pandemic, it's likely to be several years before ESSA is

reauthorized. But the depth and breadth of the backlash against statewide standardized testing in recent years and the continued attacks on testing from both the left and the right suggest there's a very real chance that Congress could abandon the ESSA testing mandate when it next rewrites the law. That, in turn, could mean the end of annual statewide testing in many parts of the country and an end to the transparency and civil rights protections it was designed to produce.

That leaves school reformers in a race against the clock to create testing systems that are more valuable to educators and parents and that offer meaningful windows into school and student performance without overwhelming teachers and principals. That is, they're in a race to change the national narrative on standardized testing.

ENDNOTES

- 1 Merrie Najimi, Facebook Post, (April 10, 2010) <https://www.facebook.com/profile.php?id=100005070708557>. Najimi, president of the Massachusetts Teachers Association announced the suspension of spring 2020 testing and wrote: "Onward to cancelling MCAS PERMANENTLY and reclaiming our schools as joyful places, where our curriculum returns to being developmentally appropriate as it was before we were shackled by MCAS, and where ALL kids can have their recess back since they won't have to spend their days preparing for MCAS."
- 2 Robert Rothman, "Something in Common: The Common Core Standards and the Next Chapter in American Education", 2011, Cambridge, MA: Harvard Education Press.
- 3 Thomas Toch, "Margins of Error: The Education Testing Industry in the No Child Left Behind Era," 2006, Washington, DC: Education Sector.
- 4 The quotes in this report, except when otherwise noted, are based on interviews by the authors.
- 5 See, for example, Jane Hannaway and Laura Hamilton, "Performance-Based Accountability Policies: Implications for School and Classroom Practices," 2008, Washington DC: Urban Institute and RAND Corporation; Joseph J. Pedulla et al., "Perceived Effects of State-Mandated Testing Programs on Teaching and Learning: Findings from a National Survey of Teachers," 2003, Chestnut Hill, MA: National Board on Educational Testing and Public Policy; Richard Rothstein et al., "Grading Education: Getting Accountability Right," 2008, New York: Teachers College Press; Thomas Toch, "Margins of Error"; and Victor Bandeira de Mello et al., "Mapping State Proficiency Standards on NAEP Scales: 2005-2007," October, 2009, Washington DC: Institute for Education Sciences, National Center for Education Statistics.
- 6 Cindy Long and Sara Robertson, "NEA Launches Campaign to End 'Toxic Testing,'" <http://www.nea.org/home/59747.htm>.
- 7 Monty Neill and Lisa Guisbond, "Test Reform Victories Surge in 2017: What's Behind the Winning Strategies?" November 2017, Boston, MA: FairTest, <http://www.fairtest.org/sites/default/files/FairTest-TestReformVictoriesReport2017.pdf>.
- 8 Patrick Wall, "As NYSUT Endorses Testing Opt-Outs, City Union Holds Back," April 1, 2015, Chalkbeat, <https://chalkbeat.org/posts/ny/2015/04/01/as-nysut-leader-endorses-testing-opt-outs-city-teachers-union-walks-a-fine-line/>.
- 9 Arne Duncan, "Listening to Teachers on Testing," August 2014, Washington DC: U.S. Department of Education.
- 10 U.S. Department of Education, "Fact Sheet: Testing Action Plan," October 24, 2015, Washington DC: U.S. Department of Education, <https://www.ed.gov/news/press-releases/fact-sheet-testing-action-plan>.
- 11 Ibid.
- 12 Liana Heitin Loewus, "State and District Leaders Vow to Reduce Testing, Stick with Yearly Assessments," October 15, 2014, Curriculum Matters, http://blogs.edweek.org/edweek/curriculum/2014/10/state_and_district_leaders_vow.html.
- 13 Simba Information, "PreK-12 Testing Market Forecast 2019-2020," October 2019, Rockville, MD: Simba Information.
- 14 Ibid.
- 15 Maryland Education Association, "Maryland Legislation to Limit Testing in Schools Signed into Law," May 25, 2018, <https://www.marylandeducators.org/press/maryland-legislation-limit-testing-schools-signed-law>.
- 16 National Center for Teacher Quality, "State of the States 2019: Teacher & Principal Evaluation Policy, 2019, Washington D.C.: NCTQ, https://www.nctq.org/dmsView/NCTQ_State_of_the_States_2019_print_ready.
- 17 Diane Stark Rentner et al., "Listen to Us: Teacher Views and Voices," May 2016, Washington DC: Center on Education Policy, <https://www.cep-dc.org/displayDocument.cfm?DocumentID=1456>.
- 18 Michael Henderson et al., "School Choice in the Trump Era: Results from the 2019 EdNext Poll," 2019, Cambridge, MA: Program on Education, Policy and Governance, Harvard Kennedy School, <https://www.educationnext.org/files/2019ednextpoll.pdf>
- 19 Ibid.
- 20 Rentner et al., 2016.
- 21 NWEA, "Make Assessment Work for All Students: Multiple Measures Matter," May 2016, Portland, OR: NWEA, https://www.nwea.org/content/uploads/2016/05/Make_Assessment_Work_for_All_Students_2016.pdf.
- 22 Mark Teoh et al., "The Student and the Stopwatch: How Much Time Do American Students Spend on Testing?" Winter 2014, Boston: Teach Plus, http://www.teachplus.org/sites/default/files/publication/pdf/the_student_and_the_stopwatch.pdf.
- 23 Ray Hart et al., "Student Testing in America's Great City Schools: An Inventory and Preliminary Analysis," October 2015, Washington DC: Council of the Great City Schools, <https://www.cgcs.org/cms/lib/DC00001581/Centricity/Domain/87/Testing%20Report.pdf>.
- 24 Ibid.
- 25 Educators for Excellence, "Voices from the Classroom 2020: A Survey of America's Educators," 2020, New York: Educators for Excellence, https://e4e.org/sites/default/files/voices_from_the_classroom_2020.pdf.
- 26 Diane Ravitch, "Slaying Goliath: The Passionate Resistance to Privatization and the Fight to Save America's Public Schools," 2020, New York: Knopf.

APPENDIX

Legislative Measures to Address Concerns about Over-Testing, 2014 through 2019

	Total Introduced	Introduced						Total Enacted
	2014-2019	2014	2015	2016	2017	2018	2019	2014-2019
ALABAMA	-	-	-	-	-	-	-	-
ALASKA	-	-	-	-	-	-	-	-
ARIZONA	4	-	-	2	1	-	1	1
ARKANSAS	-	-	-	-	-	-	-	-
CALIFORNIA	4	-	1	-	1	1	1	2
COLORADO	17	3	8	1	4	1	-	4
CONNECTICUT	3	-	3	-	-	-	-	1
DELAWARE	1	-	1	-	-	-	-	1
FLORIDA	16	2	8	-	6	-	-	2
GEORGIA	6	-	2	3	1	-	-	4
HAWAII	16	-	8	4	-	2	2	-
IDAHO	1	-	1	-	-	-	-	-
ILLINOIS	9	3	2	2	-	2	-	4
INDIANA	4	-	-	2	1	1	-	3
IOWA	1	-	-	-	1	-	-	1
KANSAS	-	-	-	-	-	-	-	-
KENTUCKY	3	-	-	2	1	-	-	1
LOUISIANA	8	-	2	3	3	-	-	3
MAINE	6	-	2	-	4	-	-	1
MARYLAND	19	-	4	10	2	2	1	5
MASSACHUSETTS	8	-	2	-	3	-	3	-
MICHIGAN	5	1	-	-	3	-	1	3
MINNESOTA	14	2	6	2	2	2	-	1
MISSISSIPPI	19	1	4	-	3	3	8	-
MISSOURI	3	1	2	-	-	-	-	1
MONTANA	-	-	-	-	-	-	-	-
NEBRASKA	3	2	-	1	-	-	-	2
NEVADA	1	-	-	-	1	-	-	1
NEW HAMPSHIRE	1	-	-	-	1	-	-	1
NEW JERSEY	8	3	3	1	-	1	-	1
NEW MEXICO	11	1	6	1	1	1	1	1
NEW YORK	43	11	6	7	8	6	5	1
NORTH CAROLINA	6	-	2	-	1	1	2	2
NORTH DAKOTA	-	-	-	-	-	-	-	-
OHIO	8	-	3	1	2	-	2	2
OKLAHOMA	17	4	3	5	2	1	2	1
OREGON	7	-	2	-	1	-	4	2
PENNSYLVANIA	8	1	2	1	3	1	-	1
RHODE ISLAND	6	1	2	3	-	-	-	-

Legislative Measures to Address Concerns about Over-Testing, 2014 through 2019

	Total Introduced	Introduced						Total Enacted
	2014-2019	2014	2015	2016	2017	2018	2019	2014-2019
SOUTH CAROLINA	13	-	-	3	-	1	9	3
SOUTH DAKOTA	3	-	-	1	-	-	2	1
TENNESSEE	36	-	4	8	9	6	9	2
TEXAS	33	-	14	-	5	-	14	3
UTAH	4	-	1	1	2	-	-	3
VERMONT	1	-	-	-	1	-	-	-
VIRGINIA	28	10	3	5	4	5	1	4
WASHINGTON	20	-	4	7	6	1	2	1
WEST VIRGINIA	14	-	1	4	9	-	-	-
WISCONSIN	2	-	1	-	1	-	-	-
WYOMING	6	-	3	1	-	1	1	3

Note: First column includes 426 bills and 20 resolutions. The District of Columbia is not included in this analysis.

Source: FutureEd analysis. For details see *Technical Appendix*.

STATE LEGISLATIVE TECHNICAL APPENDIX

FutureEd conducted a scan to identify state bills and resolutions introduced from 2014 through 2019 to address concerns about “over-testing,” excluding parental “opt out” bills.

First, we conducted an initial search for examples using Google, state legislative websites, media websites, and other sources. That search yielded 100 bills, which we used to construct a preliminary understanding of the range of actions state legislators have proposed to address concerns about over-testing.

Second, we conducted a more systematic 50-state scan using Quorum, a proprietary public affairs website with a state bill database that includes records for all bills introduced beginning in 2014. We used a broad search term that yielded 54,215 bills with words like “assessment,” “test,” or “exam” in the bill text or bill summary.

Third, we analyzed each of those bills to identify ones that qualified as addressing concerns about over-testing, employing a preliminary operational definition based on the initial scan and analysis of 100 bills. In addition to the bill text, we used other sources as necessary to determine the context and intent of the sponsors in order to confirm that the bill was intended to address concerns about over-testing.

Our final results yielded 446 bills based on the following decisions and criteria:

Resolutions: We included resolutions as well as bills because they indicate legislative concern about an issue. Moreover, some types of resolutions, if adopted, enact requirements similar to a bill. The final results include 20 resolutions, or about 4.5 percent of the total.

Carryover Bills: Just over half of state legislatures allow bills to carry over from the first year of a biennium to the second. We counted the carry-over version of such bills as an additional bill only if the legislature amended the bill and gave it a new identifier or the legislature took some

kind of action to move the bill forward. The total includes fewer than 20 such bills.

Enacted: For bills, we included all bills passed by the full legislature and signed by the governor. We did not include bills vetoed by the governor. For resolutions of one body, we included all resolutions formally adopted by that body. For concurrent resolutions, we included all concurrent resolutions adopted by both bodies. For joint resolutions, we included all joint resolutions adopted by the full legislature and signed by the governor.

Bill Types Included in Final Results:

Reduce Number of State Tests

Bills that proposed to eliminate one or more state tests or tests required by the state, along with any bills that would have reduced the number of grade levels in which tests are administered or reduce the frequency of state testing (for example, bills to test students biennially rather than annually). We did not include bills that proposed to eliminate a test as a graduation requirement unless the bill would have ended the administration of the test altogether.

Study or Publicly Report on Testing Time

Bills to require a state official or entity to conduct a study of the amount of testing and testing time in public schools or require public reporting of the amount of testing and testing time in schools. To be sure that the bill was meant to address concerns about over-testing, specifically, the bill had to include study or reporting of time devoted to test administration and/or preparation. We did not include bills that would only require study or reporting of other aspects of testing, such as technical quality or alignment with state standards.

Cap Testing Time

Bills to establish a limit on the amount of time that could be devoted to test administration or test preparation, typically defined as a percentage of instructional time but sometimes in units of time such as hours or days.

Recommend Ways to Reduce Testings

Bills directing a state official or entity to take action to reduce testing generally, along with bills to require a state official or entity to make recommendations about strategies to reduce testing.

Provide Local Flexibility on Testing

Bills permitting districts to choose not to administer certain state tests as well as bills allowing districts to substitute or make other choices about tests as a strategy to reduce testing.

Adopt Multi-Purpose Tests

Bills to replace a current state test with one that a significant number of students already take for another purpose, which would reduce the number of tests students have to take. We also included bills to pilot such strategies. Most often, these bills proposed to replace a current high school test with the ACT or SAT exam. Because reduction in testing for high school students was a primary selling point of such strategies, we included all bills that proposed to replace a current state high school test with the ACT or the SAT.

Shorten State Test

Bills directing the state to shorten an existing exam as well as bills that proposed to replace a current test with a shorter test if we found evidence that the intent was to reduce testing.

Waive Test for Students Who Meet Other Criteria

Bills to waive test-taking or test administration requirements for students who achieve a designated score on an alternative exam or fulfil some other academic requirement. Most of these bills focused on high school students.

Restrict Testing in Grades K-2

Bills to prohibit or restrict standardized testing, in any of grades K-2, as well as bills to make required state tests in such grades sample tests instead of tests to be administered to all students in the grade in question.

Restrict Amount of Local Testing

Bills to limit local testing to the minimum required to comply with the letter or the intent of state laws and/or federal and state laws. It also includes several bills to restrict the amount of local testing during periods when state tests are scheduled to be administered.

Cap Number of State Tests

Bills to establish limits on state testing based on current or former levels of state testing, either indefinitely or for some period of time.

Delay or Suspend State Testing

Bills to delay, suspend, or place a moratorium on state testing explicitly due to concerns about over-testing based on evidence in the bill text, supplemental legislative documents, or media sources.

Recommend Best Practices to Minimize Testing

Bills to require a state official or entity to issue guidance or to recommend best practices to help local districts or schools minimize standardized testing.

Require Districts to Take Action to Address Over-Testing

Bills to require local districts to take some action to reduce or to minimize the amount of standardized testing, for example by agreeing to local caps on time for testing, to establish policies regarding testing, or to submit plans to reduce testing to the state.



**THE BIG TEST:
THE FUTURE OF STATEWIDE
STANDARDIZED ASSESSMENTS**