



# Reshaping Assessment and Accountability in 2021 and Beyond

COVID-19 gave the country a crash course in public health data. It underscored the value of data that can support well-informed, timely action in a crisis. The experience with virus testing, contact tracing, and vaccine distribution illustrated the dangers of decision making with insufficient information. Even when the news is bad, it is always better for those in charge—and the public—to know the truth. Good data are just as important for education systems in crisis.

And they are in crisis. Schools and districts have been struggling to reach and educate students safely. Students who were disproportionately underserved before the pandemic—including low-income students, students of color, English learners, and students with disabilities—are experiencing more severe educational setbacks now. Although it may change as the public health situation changes, the education crisis is unlikely to end anytime soon. Its effects will almost

*Data about system performance will continue to lie at the heart of school improvement.*

---

**Bonnie O’Keefe, Andrew Rotherham, and Jennifer O’Neal Schiess**

---

## State leaders must beware the dangers of acting without the right information.

certainly outlive the medical part of this crisis in the United States.

As state and local education leaders, including state boards of education, contemplate the future, they should identify the data that parents, educators, and leaders need in order to act. Systems of assessment and accountability must evolve to meet that need, and state leaders must beware the dangers of acting without the right information.

In February, the U.S. Department of Education issued guidance to states that they could not wholly waive assessments this spring, as some states had hoped to do.<sup>1</sup> This guidance balanced the critical importance of transparent information on student learning amid this crisis and offered states substantial flexibility to adapt testing to the current moment:

- waivers from the accountability requirements to identify lowest-performing schools and implement status quo school accountability consequences and 95 percent test participation requirements; and
- test design and administration flexibility, including options to shorten tests, push testing to the fall, administer tests remotely, or broaden the test window.

Assuming the Biden administration holds states to a meaningful standard in the waiver process, this approach creates the conditions for states to successfully administer low-stakes, modified assessments that give educators, parents, and leaders useful data on student performance with which to make decisions. With a pause on accountability, no student or school should be penalized for the circumstances of the pandemic. These waivers could start an important conversation about how to modify and refocus state assessments and accountability for long-term success.

As always, there is a risk that districts, advocates, or states will misuse or misinterpret 2021 test data, creating unwanted educational consequences in an already stressful time. Although 2021 tests could be a valuable source of information on student learning, among many other sources, state boards and state superintendents should set cautious expectations for educators and the public about test results. Substantially modified or shortened state assessments, or those given remotely, may not be comparable

to prior or subsequent years.<sup>2</sup> Consequently, any analysis of trends would need to come with caveats. Detailed performance information or growth data may also be limited.<sup>3</sup> For these reasons, states should consider how to present results in new ways and combine test results with a variety of other metrics that reflect the unique learning circumstances of 2021.<sup>4</sup> As in any other year, state test results are just one set of data points that parents, educators, and leaders should consider about student learning and school conditions in order to reflect on the past and make informed decisions for the future.

What might come next, beyond 2021? Longtime opponents of standards-based accountability and assessments are pushing to permanently remove the best sources of data about student performance and progress. Robust state systems of assessment and data-driven school improvement efforts are still vitally important, and state and federal leaders should not give in to pressure to dismantle them. But this multiyear disruption may nonetheless create an opportunity to reshape assessment and accountability systems in a productive way.

State boards should lead the way toward new systems of assessment and accountability. In the past, accountability and assessment systems attempted to do too much for too many. The most valuable focus for these systems moving forward will be high-quality instruction, and resources and support for the schools and students that need them most.

## A Short History of Accountability

Early iterations of nationwide accountability policy under the Improving America's Schools Act and the subsequent No Child Left Behind (NCLB) in 2001 brought unprecedented transparency to reading and math performance and graduation rates for individual students and subgroups, shining a spotlight on persistent equity and achievement gaps for students of color, low-income students, English learners, and students with disabilities. NCLB focused school intervention efforts on the lowest performing schools and largest achievement gaps, as primarily defined by math and reading proficiency rates. It also prompted states to overhaul their standards to align with college and career readiness.

This approach produced some meaningful gains in student outcomes on measures such as math proficiency and high school graduation rates.<sup>5</sup> While the focus on standards and data moved results in the right direction, the implementation and application of that focus created new problems. NCLB's narrow set of success measures led some schools to implement poor pedagogy or narrow what students were taught in favor of tested subjects.<sup>6</sup> Its system of school-level interventions was at once flexible enough for some states to do little to improve schools but also restrictive enough to get in the way of efforts by ambitious states.<sup>7</sup> All of this, plus general reform fatigue, prompted Congress to increase flexibility for states when the Every Student Succeeds Act (ESSA) replaced NCLB in late 2015.

ESSA retained most of NCLB's annual testing and reporting requirements but gave states much more flexibility in how they designed and implemented accountability for schools. It thus created a much more varied accountability landscape. A few states created their own rigorous accountability systems, but many others took the opportunity ESSA offered to create opaque quality measurement systems and

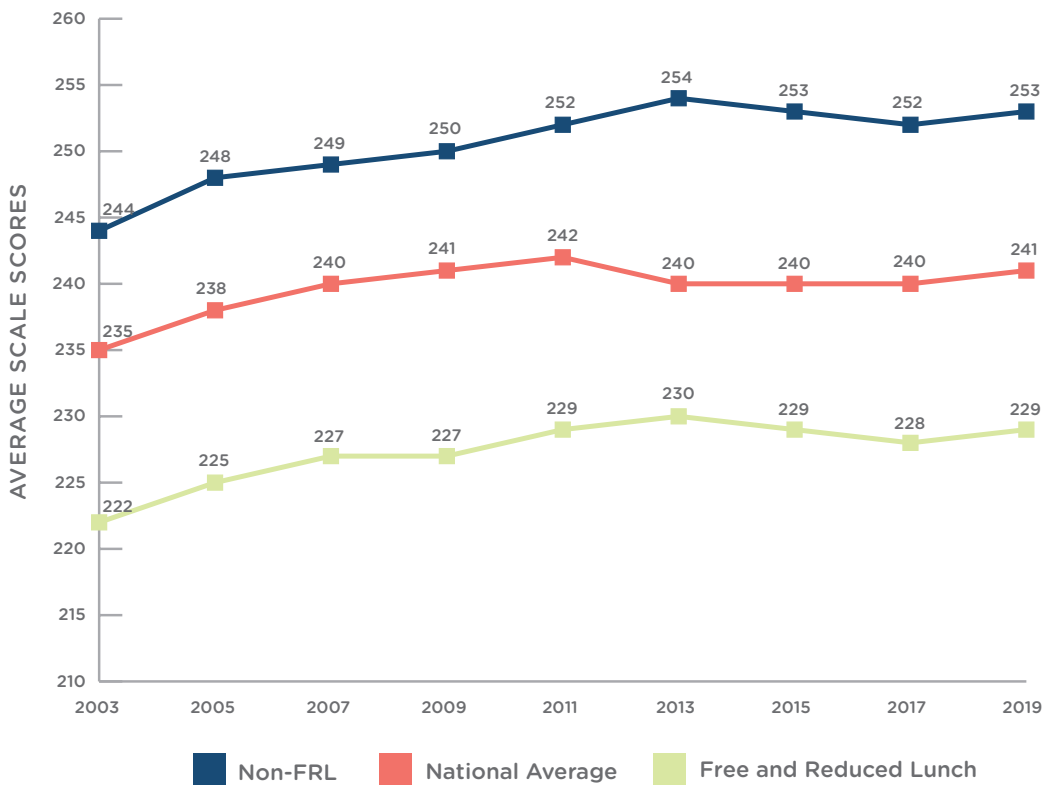
minimize interventions.<sup>8</sup> For example, in some states a school could earn the lowest possible rating for multiple years before getting any additional support or being required to do anything differently.<sup>9</sup>

Modest but measurable progress on many student achievement metrics stalled during NCLB's later years, and these concerning student performance trends have continued under ESSA. Average math and reading scores among middle and elementary school students have been essentially flat, while gaps between low-income and higher income students have grown.<sup>10</sup> For example, in the early years of NCLB (2003–11) average fourth grade math scores rose,<sup>11</sup> but they plateaued around 2011 (figure 1), with slightly widened differences between higher income and lower income students (as measured by free- and reduced-price lunch status). Under ESSA, this plateau has continued, and the hoped-for benefits of state-level flexibility have not yet materialized.

Flat achievement and persistent or widening gaps also appear in other grades and subjects.<sup>12</sup> As a result, only one in three American high school seniors in 2019 was proficient in reading, and fewer than one in four was proficient in

**Only one in three American high school seniors in 2019 was proficient in reading, and fewer than one in four was proficient in math.**

**Figure 1. Fourth-Grade NAEP Math Scores, 2003–19 (average scale scores)**



Source: National Assessment of Educational Progress, Data Explorer.

---

**State boards should focus first on directing accountability resources and energy toward supporting effective instruction in all schools.**

math, a decline from the preceding four years.<sup>13</sup> These trends cannot establish causal links with federal policy changes, but these metrics suggest that current policies and systems meant to close achievement gaps and improve student performance are not translating into better student outcomes.

Both NCLB and ESSA tried to serve multiple stakeholder groups, whose needs are often in tension: State and federal policymakers want a high-level understanding to guide resource and policy decisions and interventions. Schools and districts want faster, more detailed information to guide action in schools and to direct resources to respond to problems when they arise. Parents want to know how their children are performing and want to understand what is happening inside their schools and districts. When states try to do all these things within one framework, often no one has their needs fully met.

Politicized debates have mostly focused on removing the very data sources that educators, leaders, and the public use to understand student performance and growth. Yet we already know what education looked like before the introduction of rigorous state assessments and accountability systems: It was on the whole less equitable, less transparent, and lower performing. A revised framework for accountability should keep the best components of prior assessment and accountability systems, learn from experience, and respond thoughtfully to students' needs in the foreseeable future.

## **A New Vision of Accountability and Assessment**

As schools attempt to move forward from COVID-19, it is already abundantly clear that millions of students lost significant opportunities for learning and endured multiple destabilizing events in their time outside school walls.<sup>14</sup> These opportunity gaps are very likely to affect student learning but not in evenly distributed ways. And we know that curriculum and classroom instructional practices are among the biggest in-school factors that determine student learning.

For these reasons, state boards should focus first on directing accountability resources and energy toward supporting effective instruction

in all schools and second on intervening to change outcomes for the lowest performing schools and most underserved students. This kind of system would set out a clear, succinct set of high-level performance, progress, and equity indicators to guide state actions. Beyond those indicators, assessments and accountability would focus on what schools need to support effective, responsive instruction. Consequently, assessment systems must be fast, flexible, and embedded within high-quality curricula, and accountability systems must value continuous improvement and growth.

Such accountability systems could take the form of more streamlined end-of-year assessments, with other progress-monitoring assessments embedded into the school year. Parents and state leaders would still get the high-level information they need, but these systems would prioritize getting information to educators in a timely manner so they can act on it in close to real time.

Reliable, comparative information on student learning remains essential for state leaders to monitor educational equity and quality. The biggest potential trade-offs to faster, more flexible assessments are possible threats to precision in measurement or comparability. These are risks that assessment designers and researchers can anticipate and investigate. And states could offset these concerns with additional sources of nonacademic school performance data, sampled end-of-year assessments, or both, as well as new tools or frameworks to present data appropriately for different audiences.

Several states have taken steps toward this kind of approach by piloting more flexible assessments that can be administered in pieces throughout the school year. This way, teachers and school leaders get results faster, and tests are not such a disruptive end-of-year event. One example is New Hampshire's performance-based assessment and accountability pilot. The state streamlined standardized tests and supplemented them with assessments designed and administered by local educators, integrated into teaching and learning. All this work is aligned with the states' curricular and standards-based emphasis on deeper learning.<sup>15</sup> However, this work is complex and can take time, especially under current policy structures: The New Hampshire pilot has grown and evolved over

more than five years but has not yet rolled out to the whole state.

A stronger emphasis on instructional improvement would require most states to amp up their typically hands-off role in curriculum and teaching practices. For example, Louisiana's humanities exam pilot, which combines social studies and English language arts content and is administered in pieces throughout the school year, would not be possible if the state had not already invested substantially to create its own ELA curriculum, now in use by 75 percent of the state's school districts.<sup>16</sup>

The biggest risk of this kind of approach is that, even with high-quality curriculum and faster, more flexible assessments, students in some schools or subgroups might not be making sufficient progress to achieve college and career readiness by the time they graduate. ESSA creates a floor for states, but state boards, governors, state education chiefs, and state legislators should exceed that floor, intervening aggressively with an eye toward equity for underserved students.

One way in which state intervention could improve on current practice is by incorporating a stronger focus on district-level policies and practices. One big weakness of school-level improvement plans is that most school leaders do not control staffing, budget, enrollment policies, and curriculum—huge factors that constrain or enable improvement. States can correct for this by monitoring performance and equity indicators at a district level and requiring that district-level policy changes be part of improvement actions.

Finally, another important form of accountability is where families opt to send their children to school. In the past year, family choices have come to the forefront as COVID presented wholly new options, risks, and benefits for families to weigh. Families have made choices to send their children back into classrooms, continue virtual learning, homeschool, or switch school systems. As in the past, families with more economic resources had more choices. State and local policy should ensure that all families have an accurate picture of the educational options available to them and an equitable chance to make decisions for their children. Litigating the merits of various school choice structures is beyond the scope of this article, but

it is worth noting that education is an outlier. In most areas of American life, we balance choice and public regulation to protect the public interest while also harnessing the benefits and signaling of a choice-driven system. At a minimum, state boards focused on equity should examine who has the ability to vote with their feet on educational options and who does not.

## Role of State Boards

State boards can advance a new vision of assessment and accountability in several ways. Perhaps most important, board leaders can clarify the role and mission of boards as it relates to improving outcomes for students. State boards serve different functions in different states and have very different contexts. Depending on the state, they also have varying powers and responsibilities relating to assessment, accountability, curriculum, and standards. Some are highly politicized, while others see themselves primarily as advocates for schools rather than regulators who base their decisions on evidence and equity. This role confusion can cloud board decision making. A good regulator seeks to protect the regulated from counterproductive or ill-considered regulations but never loses sight of its fundamental role to protect the public interest.

Boards can also take specific steps. The first is advocacy and communication with other education leaders at the state and federal level. Past federal administrations' interpretations of ESSA would not allow for full implementation of an assessment and accountability framework like the one described above. It seems likely, however, that the Biden administration's approach will differ from that of its predecessors, which might bode well for more state flexibility, especially around assessments, in the coming years.

State boards could have a powerful voice in this process, especially if they add theirs to those of state and district superintendents, educators, legislative leaders, civil rights advocates, parent advocates, and other critical stakeholders. Together, these groups can set a vision that will make assessment and accountability more effective and durable, without backing away from underlying principles of equity and transparency. They can also jointly communicate to the

---

**State intervention could improve on current practice by incorporating a stronger focus on district-level policies and practices.**

**Very different approaches are possible, even under current federal law, but they require time, resources, and careful implementation.**

public that assessment results are a valuable source of information but not the sole determinant of school quality and student success. This nuance, on which the vast majority of education stakeholders agree, has often been lost in polarized arguments to abolish or uphold standardize testing.

State boards must also maintain or adopt rigorous college- and career-ready standards and approve, incentivize, or adopt high-quality curricular materials and resources, if that is part of their purview. The multiyear process to reshape curriculum and instruction from the state level in places such as Louisiana and New Hampshire show that very different approaches are possible, even under current federal law, but they require time, resources, and careful implementation.

Assessment and accountability alone will not deliver the school improvements students need, and significant changes in accountability and assessment will intersect with many other factors. For example, how should school choice and enrollment systems use information from assessments to communicate to parents effectively and equitably? And how can parents and students play a key role in defining school quality based on a variety of academic and nonacademic factors, and partner on plans for school and systemic improvement? Other critical factors—equitable school finance, teacher and leader preparation, and professional development, for example—also shape student outcomes and school quality. However, well-designed assessment and accountability systems can help states, educators, parents, and others better understand how students are progressing and make decisions with the best information available.

State boards can lay the groundwork now for reexamining, revising, and rolling out changes to current accountability and assessment plans that will bring states closer to a stronger, revised vision for accountability and assessment in the future. ■

<sup>1</sup>U.S. Department of Education, “U.S. Department of Education Releases Guidance to States on Assessing Student Learning During the Pandemic,” press release, February 22, 2021, <https://www.ed.gov/news/press-releases/us-department-education-releases-guidance-states-assessing-student-learning-during-pandemic>.

<sup>2</sup>“New Data from Curriculum Associates Quantifies Impact of COVID Learning Loss; Raises Questions about At-Home

Testing,” Curriculum Associates, press release, October 5, 2020.

<sup>3</sup>Ishtiaque Fazlul et al., “Bridging the Covid Divide: How States Can Measure Student Achievement Growth in the Absence of 2020 Test Scores” (Washington, DC: Thomas B. Fordham Institute, January 13, 2021).

<sup>4</sup>See, for example, Andrew Ho, “A Smart Role for State Standardized Testing in 2021,” *Sound and Fury* blog (Washington, DC: FutureEd, March 23, 2021), <https://www.future-ed.org/a-smart-role-for-state-standardized-testing-in-2021/>.

<sup>5</sup>Thomas Dee and Brian Jacob, “The Impact of No Child Left Behind on Student Achievement,” *National Bureau of Economic Research Working Paper #15531* (November 2009); Douglas Harris, “Are America’s Rising High School Graduation Rates Real—or Just an Accountability-Fueled Mirage?” Brown Center Chalkboard blog (Washington, DC: Brookings Institution, March 2, 2020).

<sup>6</sup>Bellwether Education Partners, “Non-Tested Subjects: Evidence of Curriculum Narrowing,” in chapter 2 of *The Learning Landscape* (Washington, DC, author, 2017); Thomas S. Dee, Brian Jacob, and Nathaniel Schwarz, “The Effects of NCLB on School Resources and Practices,” *Educational Evaluation and Policy Analysis* 35, no. 2 (June 2013): 252–79.

<sup>7</sup>E.g., Brian M. Stecher, Georges Vernex, and Paul S. Steinberg, “What Can We Learn from the Implementation of No Child Left Behind?” research brief (Santa Monica, CA: RAND, 2010); Kerstin Carlson Le Floch, “Did School Improvement Grants Work Anywhere?” blog post (Washington, DC: AIR, May 11, 2017), <https://www.air.org/resource/did-school-improvement-grants-work-anywhere>; “In Need of Improvement: NCLB and High Schools,” policy brief (Washington, DC: Alliance for Excellent Education, November 2007).

<sup>8</sup>Chad Aldeman et al., “An Independent Review of ESSA State Plans” (Washington, DC: Bellwether Education Partners and the Collaborative for Student Success, December 2017), <https://bellwethereducation.org/publication/independent-review-essa-state-plans>.

<sup>9</sup>“When Equity Is Optional: How State Choices Affect School Ratings and Identifications for Support Under ESSA” (Washington, DC: Alliance for Excellent Education, 2020).

<sup>10</sup>Kalyn Belsha, “Reading Scores Fall on ‘Nation’s Report Card’ While Disparities Grow between High and Low Performers,” *Chalkbeat*, October 30, 2019.

<sup>11</sup>Dee and Jacob, “Impact of No Child Left Behind.”

<sup>12</sup>Michael Hansen et al., “2018 Brown Center Report on American Education: Trends in NAEP Math, Reading, and Civics Scores” (Washington, DC: Brookings Institution, June 27, 2018).

<sup>13</sup>Sarah D. Sparks, “Even before Pandemic, National Test Finds Most Seniors Unready for College Reading, Math,” *Education Week*, October 28, 2020.

<sup>14</sup>Hailly T. N. Korman, Bonnie O’Keefe, Matt Repka, “Missing in the Margins: Estimating the Scale of the COVID-19 Attendance Crisis” (Washington, DC: Bellwether Education Partners, October 21, 2020).

<sup>15</sup>Scott Marion and Paul Leather, “Assessment and Accountability to Support Meaningful Learning,” *Education Policy Analysis Archives* 29, no. 9 (February, 2015), <https://doi.org/10.14507/epaa.v23.1984>.

<sup>16</sup>Louisiana Department of Education, “Louisiana’s Innovative Assessment Pilot National Update,” presentation, August 15, 2019, [https://www.louisianabelieves.com/docs/default-source/assessment/national-update-on-la’s-innovative-assessment-pilot.pdf?sfvrsn=9e849d1f\\_4](https://www.louisianabelieves.com/docs/default-source/assessment/national-update-on-la’s-innovative-assessment-pilot.pdf?sfvrsn=9e849d1f_4).