BTS
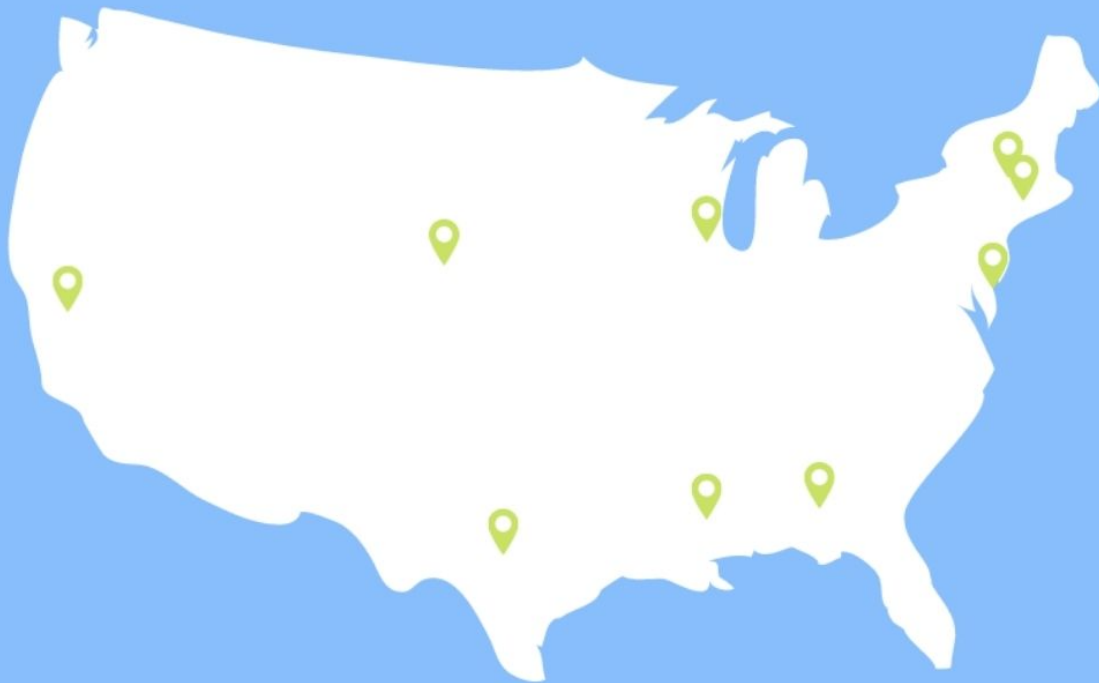
Alternative Educational Accountability *Right Now*: Examples from the Field

Ashley Carey & Jack Schneider
University of Massachusetts Lowell

Spring 2022

TABLE OF CONTENTS

Introduction

In cities and states across the country, there is a burgeoning movement to expand how we measure student learning and school quality. Over the past two decades, a body of evidence has emerged illustrating the shortcomings of the dominant approach to measurement and accountability, which remains driven primarily by student standardized test scores.[1]

In response, coalitions of researchers, practitioners, and activists have begun to develop more holistic and democratic alternatives. Such efforts have, to date, been largely regionalized, and generally operate in isolation from one another. This report draws those separate efforts together, framing them as components of a larger movement. Drawing on interviews with the leadership of these efforts, this report offers a general overview of each, including background information about their structures, successes, and challenges. This report also looks *across* these efforts to identify responses that might collectively address shortcomings of existing measurement and accountability systems in the United States.

By elevating these alternatives, we hope not only to advance similar endeavors in new places, but also to sketch out a comprehensive set of reforms that leverage the flexibility afforded by the Every Student Succeeds Act. While some fundamental changes in state measurement and accountability systems will require a reauthorization of the Elementary and Secondary Education Act, there is much that we can do *right now*—to engage communities, improve outcomes for young people, and strengthen our schools.

---

[1] Jennifer Booher-Jennings, "Below the Bubble: 'Educational Triage' and the Texas Accountability System," *American Educational Research Journal* 42, no. 2, (2005): 231-268; Thomas Dee, Brian Jacob, and Nathaniel Schwartz, "The Effects of NCLB on School Resources and Practices," *Educational Evaluation and Policy Analysis* 35, no. 2, (2013): 252; Jennifer Jennings and Jonathan Bearak, "'Teaching to the Test' in NCLB Era: How Test Predictability Affects our Understanding of Student Performance," *Educational Researcher* 43, no. 8, (2014): 381; Daniel Koretz, *Measuring Up: What Educational Testing Tells Us* (Cambridge, MA: Harvard University Press, 2008); Jack Schneider, *Beyond Test Scores: A Better Way to Measure School Quality* (Cambridge, MA: Harvard University Press, 2017).

What is ESSA?

I n 2015, Congress reauthorized the Elementary and Secondary Education Act (ESEA) as the Every Student Succeeds Act (ESSA). Initially signed into law by President Lyndon Johnson in 1965, ESEA provided an assortment of federal grants to support educational programs. Grants were distributed chiefly through Title I funds for schools and districts serving high proportions of low-income students. Although assessment was not initially a focus of ESEA, policy leaders began to push for the inclusion of standardized testing as a mechanism for assessing the extent to which federal funds were helping students as intended.[2] This shift toward standardized testing as an accountability tool became even more pronounced over the next several decades, culminating in the 2002 reauthorization of ESEA as the No Child Left Behind Act (NCLB).[3]

NCLB Changes the Game

As was the case with ESEA, the intention of NCLB was to close gaps in performance between student subgroups, and particularly between white students and students of color. NCLB pushed the nation's obsession with the "achievement gap" to new heights, increasing the emphasis on standardized testing through mandated progress benchmarks.

NCLB's embedded theory of change suggested that transparency and the threat of formal and informal sanctions would prod schools and districts to take action and improve outcomes. The law required all states to annually test students in grades 3-8, as well as in one year of high school. Schools repeatedly failing to meet adequate yearly progress (AYP) benchmarks, which rose each year toward the goal of universal proficiency, would be subject to closure or reconstitution.[4]

Because state accountability systems mandated by NCLB relied chiefly on student standardized test scores, a number of unintended consequences ensued. Many schools, for example, narrowed the curriculum in order to focus on tested subjects like math and

---

[2] Milbrey McLaughlin, "Evaluation and Reform: The Elementary and Secondary Education Act of 1965, Title I," RAND Corporation, (1975).

[3] Maris Vinovskis, "History of Testing in the United States: PK-12 Education," *The ANNALS of the American Academy of Political and Social Science* 683, (2019): 22; Koretz (2008).

[4] Frederick Hess and Michael Petrilli. *No Child Left Behind Primer* (Peter Lang, 2006).

English.[5] Other forms of gaming included teaching to the test, "educational triage"—where students closest to proficiency are targeted for interventions—and even outright cheating.[6] Although such consequences were unintended, they were not unforeseeable. In the late 1970s, for instance, social scientist Donald Campbell cautioned that a narrow set of high-stakes measures would not only "lose their value as indicators of educational status," but also "distort the educational process in undesirable ways."[7] Several decades later, millions of educators are familiar with Campbell's Law.

Fixing NCLB's Flaws

After a long delay in reauthorizing ESEA, Congress replaced NCLB in late 2015 with the Every Student Succeeds Act (ESSA). In response to critiques of the prior law, ESSA eliminated AYP requirements, pared back formal sanctions, and offered additional flexibility with regard to measurement. Under the new law, each state would choose at least five indicators to measure school performance.

<div style="border: 2px solid #8dc63f; padding: 10px;">

Overview of ESSA Indicator Requirements

1. Academic achievement in reading & Math

2. A second academic indicator, such as student growth in reading & math

3. Four-year high school graduation rates with the option to included extended year rates

4. Progress towards English-language proficiency for English learners

5. At least one measure of school quality or student success

</div>

Source: Center for American Progress, "Measuring Success: An Overview of New School Classification Indicators Under ESSA," (2008).

For the fifth measure, states were directed to choose at least one measure of school quality or student success, which might include areas like kindergarten readiness, access to and

---

[5] Dee, Jacob, and Schwartz (2013); Heinrich Mintrop and Gail Sunderman, "Predictable Failure of Federal Sanctions-Driven Accountability for School Improvement—And Why We May Retain It Anyway," *Educational Researcher* 38, no. 5, (2009); Monika Shealey, "The Promises and Perils of 'Scientifically Based' Research for Urban Schools," *Urban Education* 41, no. 1, (2006); Douglas Downey, Paul von Hippel, and Melanie Hughes, "Are 'Failing' Schools Really Failing? Using Seasonal Comparison to Evaluate School Effectiveness," *Sociology of Education* 81, (2008).

[6] Brian Jacob, "Accountability, Incentives and Behavior: The Impact of High-Stakes Testing in the Chicago Public Schools," *Journal of Public Economics* 89, no. 5-6, (2005):761–796; Jennifer Jennings and Heeju Sohn, "Measure for Measure: How Proficiency-Based Accountability Systems Affect Inequality in Academic Achievement," *Sociology of Education* 87, no. 2, (2014).

[7] Donald Campbell, "Assessing the Impact of Planned Social Change," *Evaluation and Program Planning* 2, no. 1, (1979): 67.

completion of advanced coursework, college readiness, discipline rates, or chronic absenteeism rates. This fifth measure breaks rank with the first four in that it can include one or more non-academic measures. Although ESSA introduced new flexibility with the fifth measure, the first four academic measures are weighted most heavily.[8]

ESSA Falls Short

Despite improving upon NCLB, ESSA has fallen short in a few key ways. First, like its predecessor, the law relies heavily on a narrow set of academic indicators and maintains the use of standardized tests as the primary measure of academic achievement. And while incentives to game the system have been reduced, they still exist.

The heavy reliance on standardized tests also raises questions about the extent to which present measurement and accountability systems are truly measuring school quality. As a number of scholars have pointed out, proficiency scores tend to correlate strongly with student demographic variables like socioeconomic status and race.[9] That being the case, schools serving racially minoritized and low-income populations will almost invariably produce lower aggregate scores than their whiter and more affluent counterparts. How validly, then, do existing systems measure what they purport to measure?

An additional concern is that the narrow approach to measurement outlined by ESSA fails to account for aspects of school quality that members of the public view as important. Americans have long viewed school quality in a broad manner that goes well beyond academic learning and standardized test scores.[10] Richard Rothstein and Rebecca Jacobsen, for example, surveyed a nationally representative sample of adults, asking them to rank a range of goals that schools can pursue—academic skills, critical thinking, social skills, citizenship, physical health, and more. As they concluded, an accountability system relying exclusively on standardized tests "is a travesty and a betrayal of our historic commitments."[11]

Finally, the theory of change embedded in NCLB—that stigmas and sanctions will force schools to improve—remains in place. This is perhaps best illustrated in the A-F rating systems that many states have adopted to "grade" school performance. As scholars suggest,

---

[8] For an extended overview of accountability requirements under ESSA, visit: https://www.americanprogress.org/article/measuring-success-overview-new-school-classification-indicators-essa/

[9] Koretz (2008); Pamela Davis-Kean, "The Influence of Parent Education and Family Income Child Achievement: The Indirect Role of Parental Expectations and the Home Environment," *Journal of Family Psychology* 19, no. 2, (2005); Andrew Hegedus, "Evaluating the Relationships Between Poverty and School Performance," Center for School and Student Progress, (2018); Schneider (2017).

[10] Schneider (2017).

[11] Richard Rothstein and Rebecca Jacobsen, "The Goals of Education," *Phi Delta Kappan*, (2006): 271.

these systems have not actually improved schools.[12] Moreover, labeling schools with designations like "chronically underperforming" can actually undermine their performance by driving away particular groups of families and educators. Ultimately, such practices may disproportionately harm the very communities that ESEA was intended to support.[13]

Emergence of Alternative Accountability Efforts

Alternative accountability efforts emerged as a direct rejoinder to NCLB and ESSA, and many of these efforts frame their work as a response to the failures of the dominant approach to measurement and accountability. Leaders of these efforts have created systems that seek to more accurately and comprehensively assess school quality, as well as to shift away from punitive forms of accountability.

Broadly speaking, these efforts fall into two main categories:

1) Several alternative measurement and accountability projects seek to *add more to the mix*. They offer comprehensive school quality frameworks, which include a broader range of constructs. And they utilize new forms of measurement, including student and teacher perception surveys.

2) A number of other projects seek specifically to *change the way that student learning is measured*. They chiefly rely on performance-based assessment (PBA), seeking to measure what a student knows and can do via teacher-created, curriculum-embedded "tasks."

Viewing these efforts together as a movement allows for a more complete understanding of what an alternative measurement and accountability system might look like. Together, they represent a fairly comprehensive approach to addressing the shortcomings of present systems. As such, they may offer a roadmap for improvement that does not require the reauthorization of ESEA.

---

[12] Curt Adams et al., Next Generation School Accountability: A Report Commissioned by the Oklahoma State Department of Education, Oklahoma Center for Education Policy & The Center for Educational Research and Evaluation, (2015).

[13] Eve Ewing, *Ghosts in the Schoolyard: Racism and School Closing on Chicago's South Side* (Chicago: The University of Chicago Press, 2018).

Existing Efforts: An Overview

Alternative accountability efforts are primarily collaborations between researchers, non-profit organizations, and districts that wish to engage in alternative ways of assessing student learning and school quality. Drawing on interviews with organization leadership, this section provides overviews of seven efforts currently piloting some form of alternative measurement and accountability effort in the U.S.

bravEd's True Accountability Consortia

bravEd has existing consortia in Texas, Georgia, and Mississippi, and has recently convened consortia in Pennsylvania, Ohio, and Kentucky. As an organization, bravEd helps school leaders achieve school improvement goals by measuring and tracking progress on key areas or "benefits." The bravEd system, True Accountability (TA), operates on the idea that there are two forms of accountability in organizations: (1) an institutional accountability that organizations use to shape themselves for the future and (2) a protectionist accountability that helps ensure the organizations act in a way that does not harm society

at large. Institutional accountability speaks directly to the goals and values of the organization and its stakeholders, while protectionist accountability speaks to compliance with an oversight agency. bravEd labels the policy environment stemming from NCLB and ESSA as purely protectionist and, as a result, not in alignment with how change and improvement actually occur within organizations.

The TA system focuses on the key components of the work schools do, which bravEd leadership refers to as "benefits." TA's benefits are chosen using stakeholder groups who identify what is most important to them in regard to their local school.

After school stakeholders select the benefits that represent the areas of change or improvement they desire, they determine how to measure those benefits and develop a plan to track progress on their benefits. bravEd schools draw on a variety of measures that go beyond standardized testing to measure school performance on their benefits. bravEd staff support schools in building capacity for measuring and tracking key areas that a school is focused on improving.

Relationship with the State

bravEd views their work as independent from the work of state education agencies (SEAs). While bravEd intends to initiate policy work in the future, they also worry that their model could be co-opted into the current model if policy work is not approached carefully.

Notably, the Kentucky State Board of Education, Kentucky's SEA, has recently begun its own journey into the world of alternative accountability. bravEd finds that their work in Kentucky dovetails nicely with the work of the Kentucky State Board of Education because they view their role as focused more on the process of school improvement. The Kentucky State Board of Education, by contrast, is focused on determining what areas are important for the state to track in a more protectionist fashion. bravEd leadership views their work as less concerned with the areas that states, districts, or schools deem important and more concerned with how the data are collected, analyzed, and ultimately harnessed to inform improvement in those areas.

Challenges

bravEd leadership has found shifting educators' mindsets to be one of the greatest hurdles in implementing an alternative accountability system. Under the existing accountability system, bravEd leaders find that educators working in schools considered compliant are encouraged to maintain the status quo. This dynamic, while sufficient under existing systems, stifles innovation and prevents schools from changing in meaningful ways.

Colorado Education Initiative (CEI)

The Colorado Education Initiative (CEI), since its 2007 start, has more than a decade of experience convening educators, community advocates, and policymakers to develop local-level solutions to educational challenges in their state. CEI works to convene, support, and facilitate learning among schools and districts that are piloting alternative assessment practices and accountability systems.

CEI's alternative accountability work is founded on the idea that an accountability system communicates public priorities and that public priorities are not only absent from traditional systems, but also in conflict with them.[14] Additionally, CEI recognizes multiple shortcomings of existing systems, particularly their failure to ensure equitable and high-quality learning environments for Colorado's students.

CEI has supported districts in piloting local accountability efforts and has also worked to convene stakeholder groups that often include parents, school board representatives, students, administrators, higher education institutions, and teachers. Drawing on research and institutional knowledge, districts have worked to launch metrics for use in school improvement and reporting to stakeholders about progress. Districts that have participated in a Colorado legislative pilot have chosen to include a range of metrics in their accountability work, such as access to learning opportunities, discipline disparities, data on student social and emotional learning, and more.

Relationship with the State

In 2019, the Colorado State Legislature created the Local Accountability System Grant[15], which provides funds to local education agencies that adopt local accountability systems as a way to supplement the state accountability system. CEI provided technical assistance for grantees during the first year of operation and further supported grantees by holding convenings to facilitate networking and technical assistance.

Challenges

While CEI has found success in implementing local accountability systems, they have struggled to translate this success into calls for change at a larger level. CEI leadership believes that amplifying successful local-level work can create a thirst for change on a

---

[14] Rebecca Holmes and Elliott Asp, "The Nature of Accountability is Ready for Change," Colorado Education Initiative, (2018).

[15] Colorado Department of Education, *Local Accountability System Grant,* http://www.cde.state.co.us/localaccountabilitysystemgrant, (2019).

larger scale, in addition to creating a supply of solutions from which state or federal systems might draw from.

CEI leadership aspires to support widening both local and state leaders' understanding of what is possible in the current system, while managing the political realities and concerns over the current system. Further, CEI leadership is concerned about the lack of federal clarity regarding a path forward on accountability and assessment, believing this creates confusion around how much latitude states have to innovate and explore within the current context.

California Office to Reform Education (CORE)

In 2013, California's eight founding CORE districts received a federal waiver from NCLB "to use more than just test scores to measure strengths and weaknesses in schools and to identify those in need of improvement."[16] Since then, leaders across member districts have developed a nationally-recognized school quality framework that reaches more than one million students. While the CORE framework began as an accountability system, CORE leadership has begun to transition to a system focused on school improvement.

CORE's school quality framework includes 11 areas co-constructed by educators and experts: (1) student academic growth, (2) student social and emotional learning, (3) high school readiness, (4) school culture/ climate, (5) college and career readiness, (6) graduation rates, (7) suspension rates, (8) English learner progress, (9) chronic absenteeism, (10) academic performance, and (11) disproportionality in Special Education. To measure these areas, CORE relies on administrative data and perception surveys that are administered to students, educators, and caregivers.[17] Some administrative data items are drawn from the California Longitudinal Data System, while others are unique to CORE.

Relationship with the State

CORE was launched as a response to what districts viewed as ineffective support from the California Department of Education for school improvement. Despite this beginning, CORE and the California Department of Education often work collaboratively. For example, CORE has supported the development of multiple projects initiated by the California Department of Education due to CORE leadership's strong working relationships with member districts. CORE and the California Department of Education also share findings with one another and strategize about important educational issues.

---

[16] California Office to Reform Education (CORE), "Building Continuous Improvement in the CORE Districts," https://coredistricts.org/wpcontent/uploads/2020/03/CORETheoryofActionandFocusAreas2020.pdf, (2020).

[17] For descriptions and more details on CORE's measures, visit their improvement measures webpage.

Challenges

As an early leader in alternative school quality frameworks, the CORE organization has experienced its share of challenges. CORE leadership cites managing the tension between attending to factors in schools that can create meaningful change while attending to the state and federal accountability systems as a key challenge. Additionally, while CORE leadership believes school improvement is a critical process, they recognize that the day-to-day logistics of running schools and districts take precedence. To put it simply, working with schools and districts is complex work that requires time, resources, and patience.

Massachusetts Consortium for Innovative Education Assessment (MCIEA)

Launched in 2016, the Massachusetts Consortium for Innovative Education Assessment (MCIEA) aims to expand the conception of school quality in a way that is more useful to school leaders and educators and more in tune with what the public believes schools should do. The consortium was conceived as an opposing force to the existing measurement and accountability system, which persistently labels schools serving higher proportions of low-income students and students of color as "underperforming."

Consisting of eight districts—Attleboro, Boston, Lowell, Milford, Revere, Somerville, Wareham, and Winchester—the consortium reaches one of every 10 students in Massachusetts. MCIEA employs a collaborative leadership model that brings together superintendents and teacher union presidents from each of the districts. A Boston-based nonprofit, the Center for Collaborative Education, is the fiscal agent for MCIEA, which is also supported by the University of Massachusetts Lowell. MCIEA is unique in that it draws on both an alternative school quality framework (School Quality Measures), and a performance-based assessment system (Quality Performance Assessments).

To develop the School Quality Measures (SQM) framework, SQM researchers turned to reviews of scholarly research, national polling, and local focus groups—conducted with educators, administrators, parents, and other community members—to expand the concept of school quality. The development of the framework was a democratic process designed to reflect what stakeholders believe about the purpose and function of school.

The final framework includes five main categories that are separated into inputs: (1) teachers and leadership, (2) school culture, and (3) resources, and outputs: (4) academic learning and (5) community and well-being. For a complete overview of the SQM framework, see Appendix A. To assess these categories, SQM draws from surveys of students and teachers, as well as from school-level administrative data, like the ratio of guidance counselors to students or the percentage of students enrolled in advanced courses.

SQM's sister project, Quality Performance Assessment (QPA), uses teacher-generated "tasks" to allow for a richer means of assessing student achievement. Classroom-embedded assessments allow students to apply and transfer learning by creating original products or solutions. For instance, rather than taking a subject matter test, students might demonstrate their learning by creating an original podcast that showcases learning standards or skills they are working on in class.

Collaboration remains a theme in both strands of MCIEA's work. SQM's project director works closely with administrators and educators interested in harnessing data to inform meaningful changes in their schools. One tool MCIEA staff use to help facilitate conversations about school change is their data dashboard, which displays each school's yearly progress on the areas included in the SQM framework. Similarly, QPA coaches offer extensive support for consortium educators who are using performance-based assessment tasks in their classrooms. The QPA team also maintains a task bank where educators can upload, review, and share performance tasks as a way to increase collaboration and facilitate reviews of student work.

Relationship with the State

MCIEA districts, like all Massachusetts districts, still operate under the state's existing accountability system, which is heavily reliant on student standardized test scores. While MCIEA initially had its sights set on a federal waiver from ESSA, the consortium's governing board recently decided that the pursuit of a waiver might hinder MCIEA's ability to measure school quality and student learning in alignment with consortium values. The board believed that using performance assessments to measure proficiency towards state standards might limit the ability of teachers and schools to design assessments that are tailored to their local context or learning needs. At the same time, the decision not to pursue a waiver means schools and districts are accountable to the existing state system. For educators and administrators in the consortium, this poses a dilemma: should educators focus on improving student performance on standardized tests, given the stakes associated with them, or should they focus on consortium work that more holistically addresses school improvement and student learning?

Challenges

One challenge for MCIEA members is sustaining a cohesive vision while maintaining collaborative leadership. This tension is one the consortium navigates regularly. For instance, MCIEA leadership cites challenges in approaching their work from a racial equity lens, given the complexity of the topic and their desire to maintain collaborative leadership. It is not uncommon for district partners or consortium leadership to have slightly different visions of what constitutes social and racial equity and how consortium practices can best advance that aim.

A second challenge for MCIEA members is deepening school-level engagement with SQM and QPA. While the consortium has had success engaging with district-level leaders, maintaining regular contact with school-based leaders and educators across the consortium remains a challenge. Recently, the consortium has begun offering small grants to schools as a means to spark engagement with SQM and QPA projects.

NY Performance Standards Consortium

The New York Performance Standards Consortium (NYPSC) was founded two decades ago as an effort to reflect a fuller picture of what students know and can do through performance-based assessments (PBA). Currently, NYPSC serves approximately 30,000 students across New York's five boroughs, with additional partner schools located in the Ithaca and Rochester school districts and some pilot districts outside of the state.

Members of the NYPSC believe that PBA produces better information about academic progress because it allows students to demonstrate what they have learned in a more authentic and open way. Standardized tests, by contrast, limit the capacity of students to demonstrate what they have learned. A test asks, "Did you learn X, Y, and Z?" A PBA task asks, "Can you show me what you learned?" Such openness allows for student voice, pride, and ownership of work.

NYPSC leadership regards PBA as more than just an assessment tool, viewing the approach as a system that supports and enhances curriculum, instruction, and learning through resources, external support, teacher collaboration, and policy initiatives. For a complete overview of their system, see Appendix B.

Consortium leaders believe that the success of PBA in schools is due, in no small part, to the fact that performance-based assessments are not tacked onto existing accountability systems. Instead, consortium schools are free from the external pressures of existing accountability systems that are disconnected from curriculum, instruction, and learning.

Relationship with the State

The Consortium operates under a waiver granted by the New York State Board of Regents and the New York State Education Department (NYSED), which allows member schools to graduate students using performance-based assessments in lieu of four of the five required state exams. That waiver has been extended by NYSED and the Board of Regents for the past quarter-century. Currently, the Regents are reviewing test-based graduation requirements for the entire state with a strong interest in performance-based assessment.

Challenges

For those schools contemplating a transition from a high-stakes testing approach to a PBA approach, NYPSC leaders caution that one of the greatest challenges is the change in culture that is required. For example, even with a state waiver, many educators do not have experience working without the pressures of high-stakes testing. Helping them to maintain the focus on PBA as they move from test-driven structures to inquiry-based pedagogy, via PBA, is a substantial challenge that requires full faculty support and a sense of ownership over the process.

Performance Assessment of Competency Education (PACE)

Initiated in 2011, Performance Assessment of Competency Education (PACE) is an accountability and assessment system created in collaboration with the New Hampshire Department of Education. PACE currently serves 180,000 students across 10 school districts in the state, where the organization works to increase educator, school, and district capacity for using performance-based assessments (PBA).[18] Rooted in research on how students learn, how to assess what students know, and how to foster organizational learning and change,[19] PACE's aims to achieve positive changes to the instructional core of classroom practices through collaboration, support, policy, and through the use of high-quality PBA (see PACE's full theory of action in Appendix C). PACE leadership believes building capacity for PBA increases educator efficacy and leads to greater collaborative practices among educators and administrators.

As an alternative accountability system, PACE reduces the need for frequent standardized testing by replacing some mandated tests with PBA. Altogether, their system uses a combination of locally developed PBA tasks, PACE common tasks, and state standardized tests (NH SAS). PACE common tasks are collaboratively developed and administered by all participating schools and districts, whereas local PBA tasks are developed within the locale in which they are administered. By adding local and PACE common performance tasks strategically throughout grade levels, testing occurs less frequently.[20] Research suggests that PACE meets standards for technical quality for accountability purposes when compared to traditional test-based systems.[21]

---

[18] PACE districts include Sanborn Regional, Rochester, Epping, Souhegan, SAU 39, Concord, Laconia, Newport, North Conway, and Hinsdale.

[19] Carla Evans, "Effects of New Hampshire's Innovative Assessment and Accountability System on Student Achievement Outcomes After Three Years," *Education Policy Analysis Archives* 27, no. 10, (2019).

[20] For an overview of PACE Innovative Assessment and Accountability System by grade and subject, visit: https://www.education.nh.gov/sites/g/files/ehbemt326/files/files/inline-documents/pacegradessubjects052720_4.pdf

[21] Carla Evans and Susan Lyons, "Comparability in Balanced Assessment Systems for State Accountability," *Educational Measurement: Issues and Practice* 36, no. 3, (2017): 24-34.

Relationship with the State

PACE was developed as a collaborative project with the New Hampshire Department of Education (NHDOE) to meet the basic requirements of a federal innovation waiver under NCLB. In 2018, the U.S. Department of Education approved NHDOE participation in the Innovative Assessment Demonstration Authority (IADA) program under ESSA.[22] This waiver allowed NHDOE and PACE to pilot an assessment and accountability system, where standardized testing frequency was decreased with the addition of performance-based assessments to measure student academic success.

While the program experienced early success as a collaborative effort with the state, it has struggled to navigate the transition to a state administration that favors privatization of education and de-funding of public-school programs. The NHDOE has since granted funds to New Hampshire Learning Initiative to support NH "PLACE"—a curriculum and performance assessment model that is classroom based, but not used for accountability purposes.

Challenges

One of PACE's greatest hurdles is the political landscape. PACE's existence as a state-supported program in a state with strong support for local control makes the program susceptible to changes in administration. In 2020, the Governor of New Hampshire, Chris Sununu (R), suspended assessments citing the need for increased flexibility during the pandemic. This suspension also included the assessment work of PACE. As a result, PACE's focus shifted away from assessment and towards support and coaching work with district partners. Given that the governor and education commissioner both currently support policies like school vouchers, PACE's future as an accountability system remains at risk. Understanding this political risk, PACE maintains relationships with organizations like the New Hampshire Learning Initiative (NHLI) to ensure the movement for alternative accountability continues.

A second challenge experienced by PACE is that the use of PBAs and standardized testing concurrently has created two conflicting goals. On one side, schools and educators were tasked with achieving satisfactory test scores, which are susceptible to practices like teaching to the test and educational triage, while on the other they were attempting to implement a holistic system that prioritized authentic teaching and learning. This contradiction is especially problematic for educators who want to engage meaningfully with PBAs, but who work with students in a tested grade level. In short, educators have to split their efforts between teaching to the test and engaging meaningfully with PBAs.

---

[22] Charles Barone, "Lessons from New Hampshire's Innovative Approach to Assessment and Accountability," Fordham Institute. https://fordhaminstitute.org/national/commentary/lessons-new-hampshires-innovative-approach-assessment-and-accountability

5Essentials Framework & Survey

The 5Essentials is an improvement framework and diagnostic survey that provides insights into schools' strengths across five areas: (1) effective leadership, (2) collaborative teachers, (3) involved families, (4) supportive environment, and (5) ambitious instruction. The 5Essentials survey, which is administered to students, teachers, and parents, has been used in more than 6,000 schools across 22 states.

The 5Essentials framework was created by University of Chicago Consortium on School Research in the 1990s and consortium leaders believe that their tools continue to serve as important indicators of school improvement. They find that effectiveness in these five areas promotes positive student outcomes like attendance, standardized test scores, GPA, and college enrollment. Consortium leaders view their framework and diagnostic survey as a school improvement tool rather than an accountability structure.

Relationship with the State

While the consortium does not maintain a formal relationship with the Illinois State Board of Education (ISBE), the 5Essentials survey is the preferred tool of the ISBE for measuring instructional environments within the state's schools. In 2014, the Illinois state legislature passed an act mandating that districts annually administer surveys to students and teachers in order to assess the instructional environments of the state's schools. While districts may choose an alternate survey, the state recommends the 5Essentials survey.

Challenges

According to UCCSR leadership, one major challenge has been navigating the usage of the 5Essentials framework for both school improvement and accountability. While consortium leaders find that the 5Essentials framework maintains value even when it is used for accountability purposes, they also believe there is much to be learned from the 5Essentials implementation in Chicago Public Schools: "Our findings show that despite the district's goals for implementation and use of the 5Essentials as both an accountability metric and school improvement tool, constraints in schools' capacity and tensions stemming from the survey's role in SQRP detracted from the data's use in practice."[23]

The Chicago Public Schools (CPS) include 5Essentials data in the district's high-stakes accountability system. Since CPS operates on a school choice model, where families can opt

---

[23] Laura Davis, Andria Shyjka, Holly Hart, Vanessa Gutierrez, and Naureen Kheraj, "5Essentials Survey in CPS Using School Climate Survey Results to Guide Practice," UChicago Consortium on School Research Report, (2021), https://consortium.uchicago.edu/sites/default/files/2021-12/5Essentials%20Survey%20in%20CPS_Practice-Dec%202021-Consortium.pdf

to send their children to schools other than their neighborhood school, families often use these ratings to "shop" for schools. In one instance, a school used their 5Essentials score in a billboard advertisement to attract families. Other unintended consequences include an attempt to game the survey in at least one school.[24] This serves as a reminder that alternative accountability is not immune to the unintended consequences observed in present measurement and accountability systems, underscoring the need for support and resources in order for systems like 5Essentials to succeed.

Alternative Accountability Inside State Offices of Education

In addition to consortia-based organizations, a new wave of alternative accountability has emerged within state offices of education in places like Colorado, New Mexico, California, Virginia, Kentucky, and Illinois.

Some efforts appear to be drawing on new flexibility under ESSA by including school climate in their accountability formulas and measuring climate through student and teacher perception surveys. This appears to be the case in New Mexico, which eliminated the existing A-F school grading system in 2019 and replaced it with a new accountability system based on multiple measures: (1) academic achievement, (2) academic progress, (3) English language proficiency, (4) school quality/student success, and (5) graduation rate. Included under school quality/student success are five indicators: (1) science proficiency, (2) chronic absenteeism, (3) college and career readiness, (4) educational climate, and (5) growth in 4-year graduation rate. These indicators account for 20% of the accountability score at the elementary and middle school level and 30% of the score at the high school level. The state also launched an online dashboard in the Fall of 2019 that shows the performance of the state and each school district in each of the five accountability measures.

Illinois became the first state in the nation to include arts as a weighted indicator in its school accountability metric under ESSA.[25] The arts indicator, which considers participation in arts coursework, quality of instruction, and student voice, will account for 5% of each school's overall accountability score. Connecticut also includes an arts indicator in their accountability formula, but not for all grade levels.[26]

---

[24] Matt Masterson, "South Loop Principal Told Teachers How to Answer School Survey: Email," WTTW, (2017), https://news.wttw.com/2017/11/01/south-loop-principal-told-teachers-how-answer-school-survey-email

[25] Ingenuity, "Illinois Makes History, Advances Arts Education for All Students Through ESSA, Press Release, (2020), https://www.ingenuity-inc.org/about/news-update/essa-arts-indicator/

[26] Connecticut State Department of Education (CSDE), "Using Accountability Results to Guide Improvement," Guide, (2016), https://portal.ct.gov/-/media/SDE/Evaluation-and-Support/EvalResearch/using_accountability_results_to_guide_improvement_20160228.pdf

Other efforts seem to be incorporating new forms of data in the wake of ESSA, but it is not entirely clear how these new data are presently being used. For instance, in 2018, the state of Illinois passed legislation mandating the Illinois State Board of Education (ISBE) "administer a climate survey to provide feedback from, at minimum, students in grades 4 through 12 and teachers on the instructional environment within a school."[27] As a result, the ISBE recommends that districts use the 5Essentials survey for this purpose and that the state cover costs for districts using the tool. At this point, however, it is unclear how the survey data are used at the state or district level. At least in these cases, there is a demonstrated desire among states to approach educational measurement and accountability in a new way. At the same time, however, states seem hesitant to fully embrace the opportunity. Fortunately, states are in a position to learn from the consortia and organizations who have already embarked on this work.

What Can We Learn from Existing Efforts?

States can more effectively measure student learning and school quality, moving beyond current systems and their shortcomings. Drawing on the knowledge and expertise of groups featured in this report, states wishing to engage in similar efforts can take clear steps to design systems that are more valid, more democratic, and less plagued by unintended consequences. In this section, we organize our recommendations into three areas of focus: (1) aligning measurement and accountability systems with public values, (2) introducing multiple measures of student learning and school quality, and (3) embracing human decision making.

---

[27] Illinois General Assembly, Public Act 100-1046, https://www.ilga.gov/legislation/ilcs/documents/010500050K2-3.153.htm#:~:text=2%2D3.153.,instructional%20environment%20within%20a%20school

Recommendation 1: Align Measurement and Accountability Systems with Public Values

Currently, the dominant approach to measurement and accountability fails to speak to the broad values Americans have for public education.[28] In order to make stronger claims on legitimacy and validity, present systems must expand to more fully capture the domains of student learning and school quality. Despite the flexibility afforded by ESSA, most educational measurement and accountability systems fail to answer basic questions stakeholders have about schools, and often present a misleading picture of school quality and student learning.

Recommendation 2: Introduce Multiple Measures of Student Learning and School Quality

Improving educational measurement and accountability systems demands not just alignment with public values, but also a more comprehensive approach to assessing student learning and school quality. A broad set of aims requires a varied set of measures, and the best systems will use multiple forms of measurement for each construct of interest. Such an approach promises to be more accurate, while also minimizing the consequences that so often result from the overreliance on a small number of metrics.

Care must also be taken to develop and use measures that do not simply repackage student demography in a different form. The use of multiple measures can help identify the degree to which particular instruments are assessing in-school vs. out-of-school variables. Thus, it might be possible to gauge present levels of student and school performance without mistakenly attributing the influence of environmental contexts to the educational context.

Recommendation 3: Embrace Human Decision Making

All of the groups featured in this report emphasize the need for care in how educational data are used for decision-making, communication, and accountability. While many existing accountability systems rely on algorithmic thinking and formulas to rate and rank schools, the organizations we spoke with suggest moving in the direction of dialogue, conversation, and human judgment. Rather than designing a formula to calculate how well a school is doing, states can use data to structure inquiry, support, and action.

---

[28] Schneider (2017).

Questions & Hurdles

I n addition to the insights identified in the previous section, our conversations with the groups leading alternative accountability efforts raised important questions that remain unanswered. They also identified significant hurdles in the path to improved educational accountability.

What about Standardized Tests and "Objectivity"?

Measurement and accountability have become synonymous with standardized testing over the past two decades. As a result, convincing stakeholders that test scores are not objective measures of student learning and school quality will be a significant challenge.

Some of the groups featured in this report have moved to drop the use of standardized tests entirely. Most of them draw on constructs like "social and emotional well-being," which have not typically been included in measurement and accountability systems. Most also use new tools, like perception surveys, which are often perceived as being "soft" or subjective in nature. Whatever the value of these approaches, they nevertheless represent a major cultural shift that most Americans will need time and support to understand and embrace.

Tearing Down the "Data Wall"

Schools, too, will need support in changing their longstanding relationships with measurement and accountability systems. Standardized testing has become woven into the structure and culture of schools, shaping the way educators and administrators approach their work. Perhaps nothing exemplifies this so clearly as the "data walls" found in so many of the nation's schools, which focus almost exclusively on student standardized test scores in math and English.[29]

---

[29] "Data walls" are visual representations of student performance on standardized tests and other metrics displayed in the classroom or hallway. The commonness of practices like "data walls" or test pep-rallies exhibits how standardized testing has become part of the cultural fabric of schools.

Shifting to a more holistic and robust system of measurement and accountability, and adjusting to a system not dominated by high-stakes testing, will take time and will likely face considerable pushback in some quarters. Any new approach must be accompanied by significant investments in professional development, support for new approaches to school improvement, and attention to the culture of data use.

How Will We Know When to Intervene?

While the efforts featured in this report offer fairer and more comprehensive measures of student learning and school quality, they do not always offer the same level of comparability as systems built narrowly around standardized tests. No Child Left Behind was designed to advance the interests of historically marginalized student groups by facilitating comparisons across schools and districts, as well as by focusing attention on a single source of data. For this reason, some civil rights groups have expressed concerns about a shift away from the existing approach to measurement and accountability, even despite its demonstrable flaws.

It is undoubtedly possible to design measurement and accountability systems that offer clearer and more complete information about students and schools. It remains essential, however, to consider how states will know when to intervene. Which measures will trigger action? How will strengths be balanced against weaknesses? How will standards for school performance be determined? All of these are important questions.

Are We Throwing out the Baby with the Bathwater?

Understanding that federal and state accountability systems emerged as a response to abuses against historically marginalized students is central to conversations regarding the role of educational accountability. While a system based almost entirely around high-stakes standardized testing has exhibited myriad flaws, the introduction of educational accountability and measurement has also led to some positive outcomes. Under NCLB, test scores increased modestly for some student groups and educational policy has shifted in other arguably positive ways, like increased per pupil spending and increased teacher compensation.[30] These outcomes would seem to suggest that existing accountability systems are doing *something* right.

While some might be eager to overturn measurement and accountability, it is important to remember that such systems exist for a reason. Rather than dismissing them, we should consider who is being held accountable, what they are being held accountable for, how they are being held accountable, and to whom they are being held accountable.

---

[30] Dale Ballou and Matthew Springer, "Achievement Trade-Offs and No Child Left Behind," working paper, (2008); Dee, Jacob, and Schwartz (2013).

Who's in Charge Here?

Over the past two decades, states have assumed unprecedented control over the nation's schools. Meanwhile, alternative accountability systems are emerging at different levels, often among consortia of districts. In some cases, multiple measurement and accountability systems are even emerging at different levels in the same places. This creates confusion over who, exactly, is in charge.

As states begin to engage more fully with new approaches to measurement and accountability, will they be willing to cede some of the control they currently wield? If so, will new governance structures emerge at different levels? And what will the role of the public be?

Conclusion

Alternative accountability emerged in response to demonstrated shortcomings. State-level measurement and accountability systems typically rely on a narrow set of measures and elevate a limited set of values. They operate algorithmically, excluding human judgment. And their underlying theory of change appears to assume that school improvement is driven by pressure and threats. Not surprisingly, these systems have produced a slew of unintended consequences.

Alternative approaches to measurement and accountability are not without their own potential shortcomings. Yet they also have a great deal to teach us. In profiling the models featured in this report, our intention is not to suggest that any one of them is ideal. Rather, we believe that they offer a range of tools and practices that might be adopted by states seeking to improve present systems.

A number of significant hurdles remain. Standardized testing will, at least for the time being, remain at the core of all state measurement and accountability systems—maintained by federal law, as well as by a range of other structural and cultural supports. Rating and ranking systems, run by the state as well as by third parties like GreatSchools.org, will continue to shape and constrain our collective thinking about how we measure school quality. And the use of sanctions will continue to incentivize various forms of gaming.

Despite these barriers, however, it is clear that immediate improvement is possible. Moreover, such steps may pave the way for broader reforms to educational measurement and accountability—reforms that may be hard to imagine today, but which may be easier to imagine tomorrow. Using existing efforts as a roadmap, policymakers and state officials can pave the way forward towards a future that is more democratic, more humane, and more likely to support the aim of school improvement.

Appendix A

Massachusetts Consortium for Innovative Education Assessment's
School Quality Measures (SQM) Framework

SCHOOL QUALITY MEASURES FRAMEWORK

1 Teachers and Leadership

1A Teachers and the Teaching Environment
1A-i Professional qualifications
1A-ii Effective practices
1A-iii Professional community

1B Leadership
1B-i Effective leadership
1B-ii Support for teaching development & growth

2 School Culture

2A Safety
2A-i Student physical safety
2A-ii Student emotional safety

2B Relationships
2B-i Student sense of belonging
2B-ii Student-teacher relationships

2C Academic Orientation
2C-i Valuing of learning
2C-ii Academic challenge

3 Resources

3A Facilities and Personnel
3A-i Physical space and materials
3A-ii Content specialists and support staff

3B Learning Resources
3B-i Curricular strength and variety
3B-ii Cultural responsiveness
3B-iii Co-curricular activities

3C Community Support
3C-i Family-school relationships
3C-ii Community involvement, external partners

4 Academic Learning

4A Performance
4A-i Performance growth
4A-ii Performance assessment proficiency rates

4B Student Commitment to Learning
4B-i Engagement in school
4B-ii Degree completion

4C Critical Thinking
4C-i Problem solving emphasis
4C-ii Problem solving skills

4D College and Career Readiness
4D-i College-going and persistence
4D-ii Career preparation and placement

5 Community and Wellbeing

5A Civic Engagement
5A-i Appreciation for diversity
5A-ii Civic participation

5B Work Ethic
5B-i Perseverance and determination
5B-ii Growth mindset

5C Creative and Performing Arts
5C-i Participation in creative and performing arts
5C-ii Valuing creative and performing arts

5D Health
5D-i Social and emotional health
5D-ii Physical health

www.mciea.org

Appendix B
New York Performance Standards Consortium's
Performance-Based Assessment System

New York Performance Standards Consortium

Performance-Based Assessment System

ADDITIONAL STRATEGIES
- Advisory
- HW labs
- Tutoring
- Conferences
- Narratives

RUBRICS
- Consortium-based
- Staff-developed & revised
- Mediated
- Across & within schools

POLICY
- Performance assessment
- Interim & summative assessment
- College prep

GRADUATION REQUIREMENTS
- Consortium & school PBATs
- Oral presentations
- ELA Regents

STUDENT WORK
- Analysis
- Extensive reading & writing
- Discussion
- Research

CURRICULUM
- Inquiry-based
- Teacher-developed
- Hands-on
- Active learning

PROFESSIONAL DEVELOPMENT
- School and Center-based
- Collorataive
- Moderation study

EXTERNAL
- External evaluators
- PAR Board
- College courses

COLLABORATION
- Team teaching
- Class visits
- School visits
- Projects

Center for Inquiry

Appendix C
PACE Theory of Action

Explicit involvement of local educational leaders in designing and implementing the accountability system	Fosters positive organizational learning and change by supporting internally-driven motivation	
Reciprocal support for local districts including technical, policy, and practical guidance	Builds local capacity of teachers and administrators	Changes to the instructional core of classroom practices
Use of competency-based approaches to instruction, learning and assessment.	Restructures the rigor and content representation of curriculum, instruction, and assessment	
Use of curriculum-embedded, high-quality performance-based assessments	Provides specific feedback to teachers, students, and parents on student progress towards proficiency	

Students are college and career ready

BEYOND TEST SCORES PROJECT

University of Massachusetts Lowell
Coburn Hall
850 Broadway St
Lowell, MA 01854

@Beyond_Scores | BeyondTestScores.org

The Beyond Test Scores Project works to strengthen public education and advance equity through more humane assessment. Standardized test scores measure only a fraction of what we all agree schools should do. Moreover, they often tell us more about student background variables than they do about school quality. If we're serious about collecting data about student learning and school performance, it's time to look beyond test scores. The Beyond Test Scores Project works with educators and communities across the United States to build fairer, more accurate, and more comprehensive measures of school quality. BTS also works with policy leaders and educational researchers to use data more responsibly and more democratically.