BTS
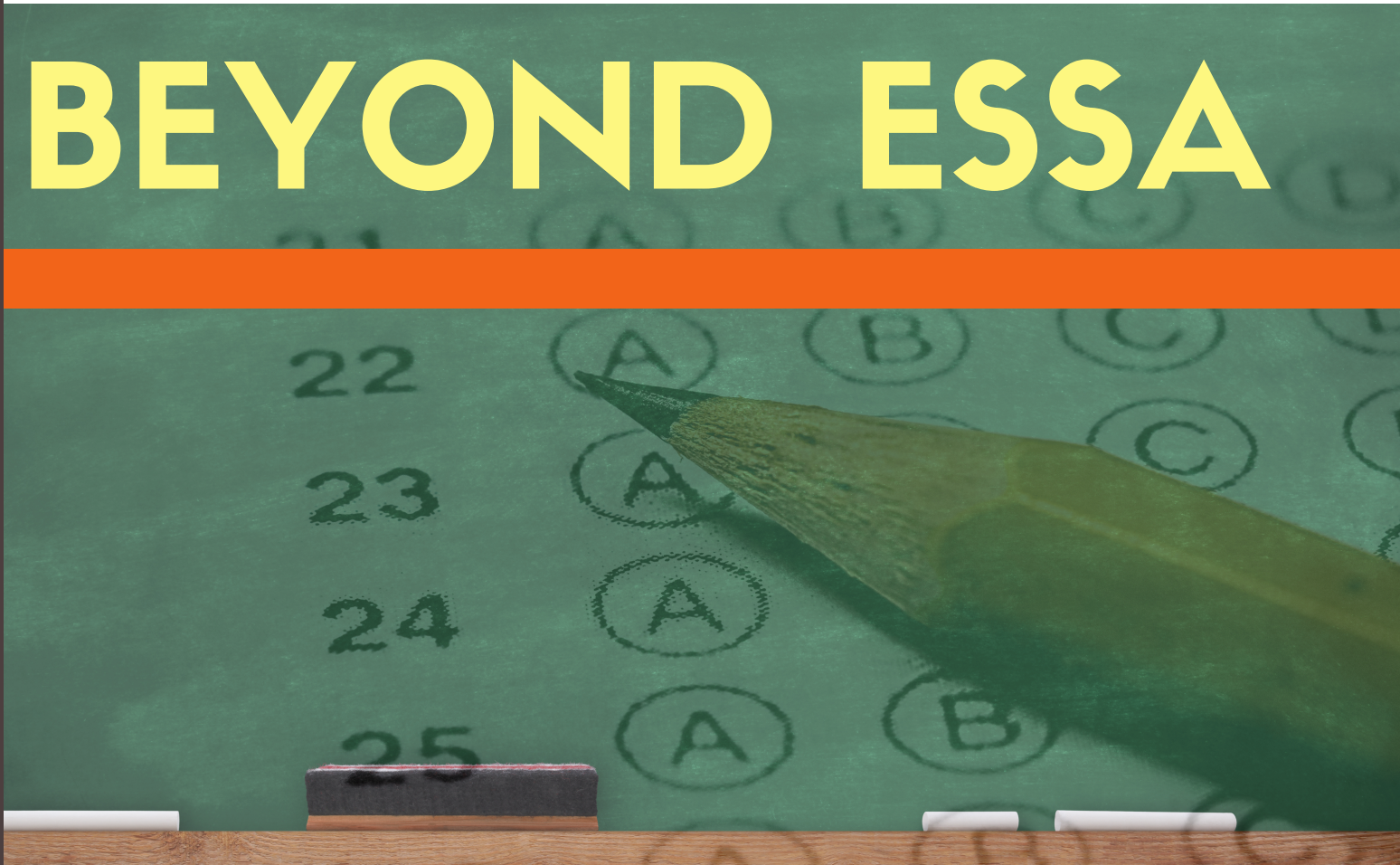BEYOND TEST SCORES PROJECT

NEPC
NATIONAL EDUCATION POLICY CENTER

EDUCATIONAL ACCOUNTABILITY 3.0

BEYOND ESSA

Working Group

Alfredo J. Artiles, Derek Briggs, Elena Diaz-Bilello, Edward García Fierros, David R. Garcia, Derek Gottlieb, Julian Vasquez Heilig, Ethan L. Hutt, Nathan Jones, Jaekyung Lee, Scott Marion, Jeff MacSwan, Katie McDermott, Bill Penuel, Jennifer Randall, Ricardo D. Rosa, Christopher Saldaña, Andrew Saultz, Jack Schneider, Benjamin Shear, Lorrie A. Shepard, David C. Webb, Kevin G. Welner, and Rachel S. White

Research Assistants

Ashley J. Carey and Kaitlin R. Mork

Suggested Citation

This report was primarily authored by four people: Ashley Carey, Kaitlin Mork, Jack Schneider, and Kevin Welner. Schneider and Welner are the Principal Investigators of the project. Carey and Mork are the Research Assistants who worked on the project. However, the ideas in the report—as well as some of the report's language —were generated by the full working group, which included the authors as well as Alfredo J. Artiles, Derek Briggs, Elena Diaz-Bilello, Edward García Fierros, David R. Garcia, Derek Gottlieb, Julian Vasquez Heilig, Ethan L. Hutt, Nathan Jones, Jaekyung Lee, Scott Marion, Jeff MacSwan, Katie McDermott, Bill Penuel, Jennifer Randall, Ricardo D. Rosa, Christopher Saldaña, Andrew Saultz, Benjamin Shear, Lorrie A. Shepard, David C. Webb, and Rachel S. White.

Contact

Jack Schneider
University of Massachusetts Amherst
Schneider@educ.umass.edu

Kevin Welner
University of Colorado Boulder
Kevin.Welner@colorado.edu

Table of Contents

Executive Summary

Several of the most problematic elements of No Child Left Behind (NCLB) were repealed when the Elementary and Secondary Education Act (ESEA) was reauthorized in 2015 as the Every Student Succeeds Act (ESSA). Nevertheless, the nation's most important education legislation remains deeply committed to test-focused approaches to accountability. In both incarnations—as NCLB and ESSA—this key federal law has been framed as an effort to strengthen schools and close opportunity gaps, yet it has constrained state and local officials, produced a host of unintended consequences, and largely failed to realize its aims. Moreover, whatever their benefits in data transparency and attention to student subgroups, these ESEA iterations have failed to meaningfully address systematic inequalities affecting racially minoritized students, low-income students, students with disabilities, emerging bilingual students, and others.

This report seeks to outline what a more effective and equitable approach to assessment of student learning and accountability for schools and districts might look like. Drawing together roughly two dozen leading scholars, it sets forth a policy agenda for the next reauthorization of ESEA. At the same time—in light of the lengthy delay in reauthorization that is likely to occur—it considers how local and state leaders might leverage some of the underutilized flexibility available under ESSA.

The approach of this report is not to tinker with the details of current federal law; instead, it outlines six principles that should be prioritized in pursuit of equity and effectiveness. Our recommendations simply seek to enact those principles through present assessment and accountability systems.

Recommendations for Effective and Equitable Assessment and Accountability:

1. Align assessment policy with goals for high-quality curricula and instruction.
2. Develop a system with reciprocal accountability.
3. Ensure that representative community members play a meaningful role in the system.
4. Move toward a broader array of school quality indicators.
5. Ensure interpretable and actionable results.
6. Design a system that will evolve and improve.

As the report explains, policy leaders should strive to advance these principles, whatever the particular nature of their work around assessment and accountability. These principles are congruent with what we have learned from educational research. And they are aligned with the fundamental aim of equity.

As we move toward what might be termed Accountability 3.0, it is important to reflect on three primary shortcomings in the nation's present approach to assessment and accountability. First among those is a flawed theory of change. The existing theory of change, exemplified by NCLB but evident in many other state and federal policies, suggests that schools will improve if sanctioned for poor student performance on standardized tests. Two decades on, it is clear that this approach fails to strengthen schools or close opportunity gaps. In envisioning a new accountability system, this report recommends revisiting this theory of change with an eye towards improving opportunities to learn, as well as toward fostering the kinds of activities and outcomes that communities want from schools.

The second shortcoming is that stakeholders have been inadequately included and involved. The current approach largely strips educators of professional judgment and generally fails to empower families. Additionally, communities have had no say in what gets assessed, how accountability is determined, or what the consequences of accountability are. This report recommends an approach to assessment and accountability that empowers stakeholders by ensuring that community members play a meaningful role in the system and by implementing a system of reciprocal accountability.

Third, the rigid, top-down approach to accountability of the past two decades has promoted distrust among educators, stifled creativity, and limited the degree to which the system can evolve. This report envisions a system that balances federal, state, and local authority—creating space for experimentation, adaptation to local context, and the evolution of systems. An effective and equitable assessment and accountability system will be locally-relevant and tied to high-quality curricula. It will also foster not just the improvement of schools, but the improvement of the assessment and accountability system itself.

We hope that members of Congress will take up a now-overdue reauthorization of ESEA. When they do, we hope they will look to research for guidance and will elevate educational equity as their primary objective. In the meantime, school districts and states need not wait for a reauthorization of ESEA. And the U.S. Department of Education, for its part, can support such efforts by working with states to leverage ESSA in ways that advance the six principles outlined in this report. Ultimately, all of us are accountable for moving forward. Inaction is not preordained; it is a choice.

Definitions of Key Terms

Accountability System: As used in recent decades, the term has been defined as, "A system that imposes student performance-based rewards or sanctions on institutions such as schools or school systems or on individuals such as teachers or mental health care providers" (AERA, APA, & NCME, 2014, p. 215). However, we use the term more broadly to refer to a system that obliges all institutions and individuals to meet their responsibilities to students.

Assessment: "Any systematic method of obtaining information, used to draw inferences about characteristics of people, objects, or programs; a systematic process to measure or evaluate the characteristics or performance of individuals, programs, or other entities, for purposes of drawing inferences; sometimes used synonymously with test" (AERA, APA, & NCME, 2014, p. 216).

Evaluation: "The process of determining to what extent the educational objectives are actually being realized" (Tyler, 1950, p. 69).

Equity: "Educational equity requires that educational opportunity be calibrated to need, which may include additional and tailored resources and supports to create conditions of true educational opportunity" (NASEM, 2019).

Indicator: An indicator is a single or composite measure that relates to a particular area of interest (e.g., postsecondary readiness).

Test: "An evaluative device or procedure in which a systematic sample of a test taker's behavior in a specified domain is obtained and scored using a standardized process" (AERA, APA, & NCME, 2014, p. 224).

Introduction

The federal Elementary and Secondary Education Act (ESEA) has been reauthorized eight times since 1965. Each reauthorization has attempted to balance the tension between a federally directed system and a system that promotes local autonomy (Gamson et al., 2015; Pinder, 2010). Since 2002, when the Elementary and Secondary Education Act (ESEA) was reauthorized as No Child Left Behind (NCLB), all states have been required to use student standardized test scores as a means of holding schools accountable. Although the most recent reauthorization, the Every Student Succeeds Act (ESSA), thoughtfully addressed some glaring problems that had arisen under NCLB, its reauthorization was greatly overdue and it ultimately preserved most of NCLB's test-focused approach (Schneider & Saultz, 2020).

The top-down nature of the last two decades of education reform has prompted educators, students, parents, and citizens alike to question the ways in which federal, state, and local authorities hold districts and schools accountable for student learning and performance (Heilig, Khalifa & Tillman, 2013; Horsford & Heilig, 2014; Schneider, 2017). Eight years after the passage of ESSA, it seems unlikely that ESEA reauthorization, which is presently overdue, will be taken up quickly by Congress. Whatever the delay this time around, however, we see value in developing a coherent vision for what reauthorization might achieve and what states and school districts can do currently—prior to any reauthorization—to address existing challenges and limitations.

Given the mounting dissatisfaction with test-based accountability and a greater readiness (at least in some quarters) to confront systemic, racialized inequality, this report begins with the premise that policy leaders must seek new approaches for school improvement. ESEA was, from the outset, intended to address large and troubling inequities in the educational opportunities provided to children throughout the nation (Thomas & Brady, 2005). In key ways, the shift toward standards and accountability—a shift that started with the 1989 Charlottesville Summit, continued with the 1994 Improving America's Schools Act, and then ramped up with NCLB—was in line with that original intent. This is particularly true with regard to the requirement that test results be disaggregated to expose the gaps affecting marginalized subgroups. But two decades of experience have also revealed what many researchers predicted: opportunity gaps and the resulting achievement disparities were not closed merely by emphasizing higher standards or by holding teachers, schools, and districts accountable for raising students' test scores.

The fundamental concern guiding this report is access to educational opportunity. We therefore seek to consider how assessment and accountability systems might be augmented to better address systemic inequalities affecting racially minoritized students, low-income students, students with disabilities,

emerging bilingual students, and others. Such inequalities, often intersecting in nature, have not been addressed in an effective way by the approach that has dominated for the past two decades.

The dominant form of educational accountability, presently enacted in all 50 states, falls short of the useful and equitable system we believe is possible. Despite some modifications made in the transition from NCLB to ESSA, federal law requires states to take a high-stakes and often punitive approach. Moreover, state systems largely fail to address the issue of school and district capacity. Additionally, despite modest changes to the accountability indicators included in the current federal system, it is still built chiefly around standardized tests. As was repeatedly documented during the NCLB years, an overreliance on tests for accountability purposes has produced a range of negative unintended consequences (Koretz, 2017; Loeb & Byun, 2019; Dee et al., 2013; Jennings & Bearak, 2014; Jennings & Sohn, 2014). Finally, present accountability systems measure only a fraction of what schools seek to do for young people. By taking such a narrow approach, current accountability systems both fail to capture the full work of schools and encourage leaders to reduce the mission of public education.

Starting with the assessment and accountability requirements of ESSA and then considering associated state- and district-level policies, we seek to outline what we see as a better approach to assessment and accountability, rooted in research evidence. What elements of current federal assessment and accountability requirements should be preserved? What elements should be peeled away? What new elements should be introduced? How can assessment and accountability best support socially just school improvement efforts? We answer these questions by thinking specifically about what states are currently required to do by federal law, what they are currently prevented from doing in light of ESSA's mandates, and how they might use flexibility within the ESSA legislative regime. What should be measured at the state level? How should those measures be used in service of accountability? What kinds of flexibility should be built into federal law? We ask a similar set of questions concerning district-level requirements and flexibility.

To answer such questions, we convened a working group of two dozen leading scholars in the field of educational assessment and accountability. Over the course of nearly a year, scholars engaged in conversation guided in part by a desire to reimagine accountability in service of an ESEA reauthorization. The working group was also committed to exploring state and local policies that could

advance a more useful and equitable system within the current federal framework. Conversations highlighted the importance of a coherent, unified policy—changes across domains, which work together in a mutually reinforcing manner.

This report brings together scholars' recommendations for federal, state, and local action. These recommendations flow from six principles for effective and equitable assessment and accountability. The body of the report is organized around six recommendations, each of which seeks to advance these principles.

Principles for Effective and Equitable Assessment and Accountability:

1. Assessment and accountability should support student learning by beginning with engaging, challenging, and relevant curriculum and instruction in each classroom.
2. While the system holds schools accountable for the education of students, it should correspondingly hold elected officials and other leaders accountable for providing schools with what they need to succeed.
3. Community members should play a meaningful role in both the crafting and enactment of educational accountability.
4. Assessment and accountability should reflect a broad array of input and output indicators that measure what we care about in schools, and that decrease the current overreliance on high stakes standardized tests.
5. Assessment and accountability systems should provide interpretable and actionable results.
6. Whatever the particular nature of our next educational assessment and accountability system, it should be designed to evolve and improve over time.

The report begins with a recognition of similar initiatives focused on ESSA reauthorization that have been undertaken over the past few years. These multiple efforts suggest a broad desire for a new approach to the nation's most important education legislation. The overview of similar initiatives also highlights the ways in which our work makes a uniquely contribution. In the main body of this report, we present our recommendations, explaining how they fit together to form a school improvement system that includes all of the components necessary for student success.

Review of Initiatives to Improve Assessment and Accountability

This report is not the result of recent revelations. Since the dawn of the NCLB era two decades ago, scholars, activists, and policy leaders have been identifying shortcomings in federal accountability legislation and the systems enacted by all 50 states. In fact, many of the scholars who contributed to this report have spent much of the past 20 years assembling powerful evidence from research, identifying the ways in which the dominant approach to assessment and accountability might be strengthened to advance the aims of educational equity and educational excellence.

This report is also not a lone effort. In light of the fact that we now have two decades of evidence, it should not be surprising that a number of efforts have emerged seeking to outline a more useful, valid, and equitable approach to assessment and accountability. The emergence of these multiple efforts is also a product of the fact that ESEA is now overdue for reauthorization. Congress may not act for several more years, as was the case in the transition from NCLB to ESSA; yet reauthorization is now on the horizon.

Organizations across the country have explored key considerations for a reimagined federal accountability system. Among them are: the National Urban League & UnidosUS (the Future of Assessment and Accountability Project, 2022), the National Coalition for School Diversity (PRRAC, 2011), the U.S. Chamber of Commerce Foundation, and the Racially Just Assessment and Accountability Initiative (Partnership for Future of Learning, 2022).

We do not here attempt to review these or other earlier reports. In general, however, these organizations emphasize the importance of centering racial and economic equity at the heart of modifications to federal and state accountability systems. As part of this, these organizations have, to varying extents, sought out the voices and experiences of students, communities, educators, and parents in proposing reimagined assessment and accountability systems. Proposed ideas put forth by these organizations include promoting school racial integration and other forms of diversity, decreasing the concentration of poverty in schools, imposing new requirements on the distribution of school resources, broadening the definitions of student and school success, engaging in continuous improvement and accountability, and shifting from a top-down approach to accountability to a reciprocal system. While our work centers many of these same priorities, as well as the priorities of equitably serving students with disabilities and emerging bilingual students, we intend for our proposed system to be holistic and coherent such that changes across domains mutually reinforce one another.[1]

---

[1] For instance, beginning with NCLB and continuing with ESSA, decisions about reclassification to English proficiency have been standards-based and tied to assessment instruments that, as a practical matter, measure language and academic achievement concurrently by focusing on what is generally referred to as "academic language" (Rolstad & MacSwan, 2023). As a result, students who know English well but do not yet have the level of academic achievement needed to

Policy Recommendations

In the pages that follow, we describe a system in which assessments and accountability serve to advance students' opportunities to learn, with clear links between research-based policies and their beneficial outcomes. Toward this end, we make six recommendations that should be prioritized in working towards a more effective and more just approach to assessment and accountability.

First, assessment and accountability should center student learning by beginning with an engaging, challenging, and relevant curriculum in each classroom. Student demonstrations of learning should further student learning and avoid the trap of assessments that distort or distract from the learning process. Accountability systems must not undermine this first imperative.

Second, we embrace the idea of reciprocal accountability, meaning that while the system holds schools accountable for the education of students, it must correspondingly hold elected officials and other leaders accountable for providing schools with what they need to succeed. Leaders, and those who put leaders in power, cannot reasonably demand unidirectional accountability; a system's demands on schools must be linked to the provision of capacity, support, and resources.

Third, we stress the importance of a meaningful role for community members. States and districts should actively involve community members—particularly those historically marginalized—in the design and implementation of assessment and accountability. When accountability systems call for the sanctioning of schools, the involvement of the community is again crucial. School improvement cannot be delivered to or imposed on communities; it must genuinely involve them.

Fourth, we point to the need for a broad array of input and output indicators that measure what we care about in schools, and that decrease the current overreliance on high stakes standardized tests. We stress the importance of creating assessment and accountability systems that are valid, in that they measure what they purport to be measuring, and that better characterize school quality overall. And we stress the importance of systems that are equitable, in that they build capacity and strengthen performance, rather than relying on a market logic of competition and the threat of sanction.

Fifth, we advocate for assessment and accountability systems that provide interpretable and actionable results. This includes assessments that are more tightly linked to the actual work of teachers in classrooms, rather than relying on externally-produced and managed standardized tests. It also entails providing results that are interpretable for the broader community, while avoiding the risk of distortion that is inherent to reductive approaches like A-F grades.

Sixth, we make the case for a system that can evolve and improve. This requires ensuring that a number of key processes are in place. Is there authentic, meaningful, and representative community engagement? Are leaders doing their part to provide resources that meet the needs of students, educators, and schools? Are schools implementing authentic and relevant demonstrations of learning? These and other elements of the system will require active monitoring and improvement, which requires ongoing evaluation.

---

score well on these tests end up as long-term English learners (LTELs), a group that has been growing dramatically. These LTELs generally have reduced access to challenging academic courses in high school.

Recommendation #1:
Align Assessment Policy with Goals for High-Quality Curricula and Instruction

Accountability systems serve both short-term and longer-range goals. As outlined in state and federal laws, they aim for the immediate goal of making public entities and employees answerable to the public and of making results transparent. NCLB, in demanding that schools attend more carefully to the needs of marginalized and vulnerable student subgroups, did send a powerful signal about the importance of equitable outcomes as measured by test scores. But beneficial accountability systems must also be designed to serve the larger purpose of helping to create schools that provide students with rich opportunities to learn. Little is gained and much is lost if an accountability system accomplishes the first goal in ways that undermine the second. And, as we have seen in study after study, this can happen when high-stakes assessments are developed separately from high-quality curricula and instruction.

Lesson Learned from NCLB/ESSA:

Don't let the accountability system drive poor practices in classrooms.
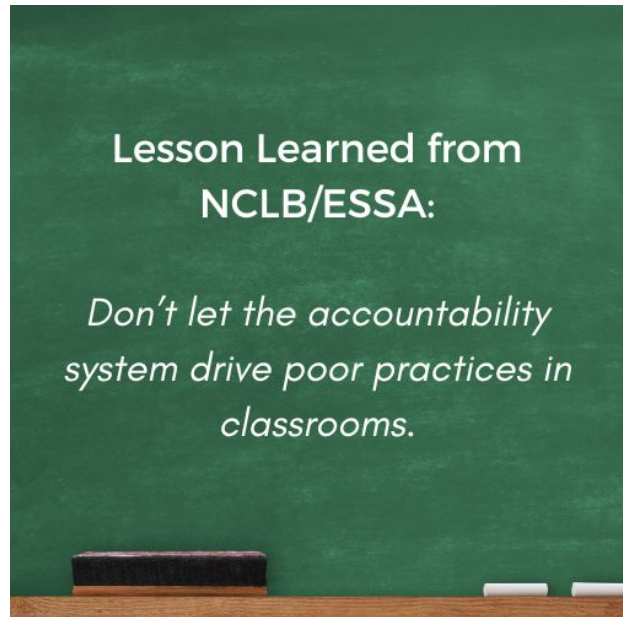
Districts Should Implement High-Quality Curricula and Instruction Guided by a Vision of Learning

Improving education requires strengthening opportunities to learn and offering supports for students owed an "education debt" (Ladson-Billings, 2006). We begin, then, with the central importance of engaging, challenging, and relevant curricula and instruction. High-quality curricula are essential for promoting equity and affirming students' rights to a meaningful education, allowing them to cultivate their intellectual well-being, humanity, and potential (Espinoza et al., 2020). Further, these curricula must be tied to instructional quality, as the Council of Chief State School Officers recognizes with its High-Quality Instructional Materials and Professional Development Network (CCSSO, 2022).

Ultimately, it must be school districts that cultivate and ensure such curricula and instruction. States can and should continue to provide academic standards and serve as clearinghouses for quality curricula. But district leaders, in collaboration with classroom educators, should develop a place-based vision of learning that connects with the broader purposes of education, goes beyond those academic standards, and guides all assessment. That vision should focus on equity in the quality of students' experiences, ensuring that all students maintain a deep sense of belonging in and beyond the classroom. High-quality curricula and instruction should also align with what we know about how people learn, highlighting key learning goals, as well as what students and educators should do in classrooms to meet those goals (NASEM, 2018).

Districts Should Develop Curriculum-Embedded Assessments

Districts should develop or adopt curriculum-embedded measures of student learning that are authentic and relevant to students' lives. This could include portfolios, performance-based assessments, and projects with public demonstrations—authentic demonstrations that enrich learning opportunities, as opposed to simply generating scores for high-stakes accountability. Assessments should be anchored in challenging and relevant curricula and should be designed to enhance student buy-in through engaging tasks and opportunities for students to document their learning (Penuel, 2021). For example, the project-based learning units at New Tech Network schools are collaboratively designed by teachers and seek to ensure that each project is meaningful for the school's students, so that a student never is left wondering, "Why are we learning this?" (Adams & Grand, 2019).

High-quality assessment materials are also a resource for educator learning that can help shift classroom instruction (Davis et al., 2017). Such shifts can support coherence and equity at the system level when high-quality materials and associated professional learning are widely available (Kirp, 2013). In addition, assessments anchored in high-quality instructional materials are more likely to be fairer to students, since students will have had a meaningful opportunity to learn the content and skills being assessed.

State education agencies can support the development and adoption of curriculum-embedded assessments. For example, SEAs could develop model units with embedded assessments, which districts could then adopt or use as an exemplar to develop their own assessments. Another way SEAs could support this work is by identifying high-quality curricula that include embedded assessments that districts could then adopt.

Districts Should Align Grading Practices with High-Quality Curricula and Instruction

In line with the implementation of high-quality curricula and instruction and aligned assessment practices, districts should promote the development of grading practices that support equitable learning. Grading criteria should encourage risk-taking and be grounded in the assumption that all students can be successful. Grading schemes, for example, might give students multiple chances to show what they know and can do (Aguilar et al., 2015; Feldman, 2018; Schneider & Hutt, 2023). This could include providing students opportunities to correct assignments and options for earning credit for meeting learning goals. The act of grading should not detract from high-quality teaching or from student learning.

Recommendation #2:
Develop a System with Reciprocal Accountability

The dominant approach to accountability, framed by NCLB and ESSA, has been unidirectional, focusing on making demands of teachers, principals, and district leaders. A reimagined accountability system must instead be grounded in the idea of reciprocal accountability (Elmore, 2002; Oakes & Welner, 2023). School and district leaders should be directly responsible for the education of students, while leaders in federal and state government have an obligation to provide schools with the capacity and resources needed to serve students. Leaders, and those who put leaders in power, cannot reasonably demand or accept unidirectional accountability.

Lesson Learned from NCLB/ESSA:

Resources and capacity building are at least as important as high standards and demands.

State and the Federal Government Should Be Held Accountable for Investing in Schools

Existing accountability systems suggest that educators and school leaders need to make use of the discretion they have to run classrooms and schools in a manner that advances student learning. We agree. Disparities in resources cannot be overcome by Title I funding alone. Thus, while teachers should be held accountable for providing a high-quality education for their students, state and federal offices should be held accountable for investing in basic necessities like educator development, school safety, and reasonable class sizes. States also have a role to play in serving as clearinghouses for high-quality curricula, for offering professional learning opportunities to leaders, and for building school district capacity.

In short, each set of actors within the system should be accountable not only to meet the obligations identified in state and federal policy, but also to meet obligations to schools, teachers, and students. Such reciprocal accountability is an important component in a system that ensures that schools and districts are adequately supported to carry out their key functions. Capacity-building, which often depends on a commitment from districts and states, is a core piece of standards-based reform that was lost in the shift towards high-stakes, test-based accountability (Shepard, Marion, & Saldaña, in press).

State and the Federal Government Should Expand Accountability Indicators to Include Inputs and Processes, Not Just Outcomes

Measures of school quality under NCLB and ESSA have placed an outsized emphasis on test scores, at the expense of other important measures (Lee & Lee, 2020; National Research Council, 2011; Schneider, 2017). This test-driven accountability policy has failed to encompass a whole-child development approach, including non-cognitive measures of achievement such as socioemotional skills, mental and physical well-being. This narrow set of measures also has failed to highlight the lack
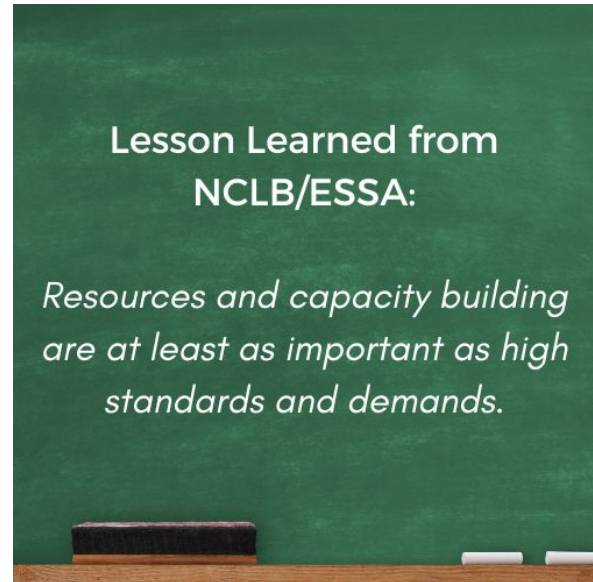
of resources and supports needed to improve student outcomes and close racial and socioeconomic achievement gaps (Kirsh & Braun, 2016; Lee & Wong, 2004). An effective and equity-focused accountability system must reach beyond achievement outcomes, to encompass the inputs and processes required to achieve desired results (Elmore, 2005; Jackson, 2020; Koretz, 2008; Lee & Reeves, 2012; Marion et al., 2017; Sebring et al., 2006). For example, legal protection of students' right to due process is an important type of accountability for processes. A recent National Education Policy Center brief on the future of the Elementary and Secondary Education Act recommended that the federal government require states to use the federal Civil Rights Data Collection to identify schools with racial disproportionalities in discipline or access to advanced courses and provide funding to help eliminate those disproportionalities (DeBray et al., 2022). In cases where inputs and processes are falling short of expectations, state and federal offices, not schools and districts, should be held accountable and required to act.

By not including indicators for inputs and processes, federal leaders are missing an opportunity to help states and localities identify inputs that are associated with valued outcomes. Expanding accountability indicators to include inputs and processes, will support school and district learning about the types of inputs that others have adopted, which may have led to improvements in particular outcome measures. In short, rather than telling educators the destination while removing all road signs and withholding directions that will get them there, input-focused accountability indicators can provide signals pointing educators to the desired destination.

Indicators of the quality of schools and the broader school system should provide stakeholders and the public with information about the extent to which schools have what they need to succeed. These indicators should highlight the extent to which states and the federal offices of education are fulfilling their responsibilities. The accountability system should, in this way, combine school-level accountability with attention to the overall functioning of the educational system and community needs. For example, in California, all districts and schools receiving funding under the Local Control Funding Formula (LCFF) are required to develop a local accountability plan (LCAP) that includes community-determined goals and priorities for districts to address community-based priorities developed at the local level (Heilig et al., 2014).

Further, if federal and state law impose content and performance standards for all students and particular student subgroups on school districts, stakeholders should be made aware of the gap between current spending and adequate levels of spending, with adequacy reflecting the resources needed to reach state and federal expectations. Illinois's Evidence-Based Funding Formula, for example, provides district-level estimates of adequate funding, providing stakeholders with information about how far state efforts are from fully funding public schools. Indicators like these could highlight the extent to which states and the federal government, are fulfilling their responsibilities to local stakeholders. The accountability system should, in this way, combine school-level accountability with attention to the overall functioning of the educational system.

Recommendation #3:
## Ensure that Representative Community Members Play a Meaningful Role in the System

For accountability systems to be fair and effective, communities should be the primary drivers in designing those systems. This is particularly important in marginalized and minoritized communities that have long been disenfranchised and that disproportionately have endured the brunt of accountability consequences. While ESSA requires states to include stakeholder groups, the next revision of the nation's accountability approach should fully embrace the idea of devolving important decision-making authority to representative community members.

States Should Incentivize Partnerships with Community

With this goal in mind, states should incentivize and facilitate school and district partnerships with parent and community organizations to ensure that members of underserved, marginalized, and disenfranchised communities are listened to and learned from. Examples of such partnerships can be found in North Carolina's Village of Wisdom, which prepares families to be advocates in their communities; similarly, the Center for Racial Justice and Youth Engaged Research at UMass Amherst prepares youth, educators, faculty, and community members to engage in research and social action. To create these partnerships, school districts—assisted by state offices—should take an active role. For instance, each district might have a community stakeholder engagement liaison who is integrated and invested in the district, but who is not managed by the district and not on the district payroll. To be clear, these liaisons would not be called upon to represent the voices of the communities they serve. Rather, their services would be contracted or made available to any community member seeking to acquire tools for effective advocacy in educational spaces. These efforts will take fiscal and human resources in which the state and federal governments must be willing to invest (e.g., compensation for community members' time and energy, childcare for meeting times, etc.).

Districts Should Elevate Marginalized Voices

Leaders must also adopt processes and structures to elevate marginalized voices and balance unequal distributions of power. Experienced professionals sometimes exercise their expertise with pre-prepared documents and policies, leaving little room for broader participation. More elite participants are also more likely to have familiarity and confidence in making arguments, presenting data and other evidence, and otherwise shaping discussions. Leaders can address potential imbalances by, for example, adopting processes for inclusive participation, or by training participants and stakeholders in Intergroup Dialogue processes (Zúñiga, Naagda, & Sevig, 2002). States or state-contracted providers might be available, at district request, to lead these efforts. State-level policies can also assist local
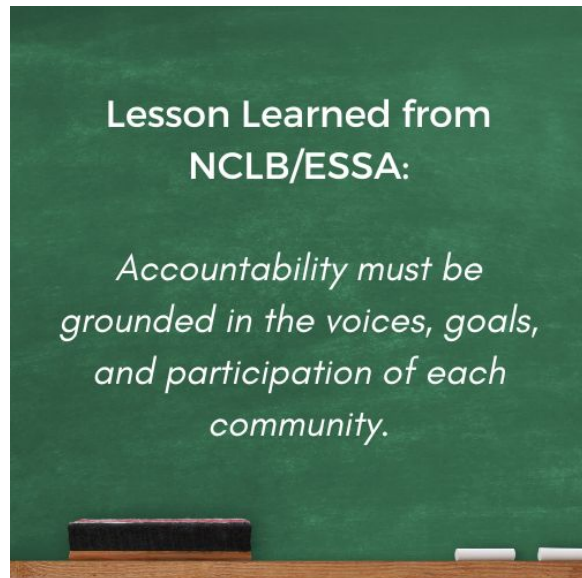
policymakers by imposing consequences for districts that fail to authentically include marginalized communities.

States and Districts Should Provide Supports to Facilitate Representative Community Participation and Responsiveness to the Community

In order for their participation to be meaningful, community members must be provided with the information, training, and opportunities to fully and successfully engage. Representative accountability bodies should have a deliberative democratic process for collecting and responding to data, including gathering and considering community feedback. In determining what measurements and data are important to collect, this feedback and deliberation should consider the communities' values, within societal constraints about what must be (or cannot be) included. And policymakers should put in place structures for community sense-making around data, including qualitative and quantitative measures.

States and the federal offices of education can also provide capacity, incentives, and requirements designed to ensure that local accountability processes are grounded in data and other evidence that are responsive to community feedback and values. Some types of data collection (e.g., regarding civil rights and services for students with disabilities) must be universal and mandatory; the legal accountability schools face for providing free and appropriate education for students with disabilities, for instance, plays an essential protective role for these students. To that end, and with pandemic-era failings in mind, it is important to expand data collection efforts in order to establish the full nature and extent of inequities in both opportunities and outcomes.[2] Other data collection decisions, however, should arise out of deliberative democracy rooted in community voices and values (Gottlieb & Schneider, 2018). Accountability and school improvement depend on data and evaluation—on a sound understanding of current inputs, processes, and outcomes. This is true whether decision-making is taking place at the state or federal level, or if it is devolved to local levels. Moreover, the process of engaging with data and evidence presents powerful learning opportunities that are particularly important for educators and reform buy-in (Neri, Lozano, & Gomez, 2019).

The resulting accountability system must create structures for participation that are inclusive for all families and that support and incentivize the lowering of barriers for meaningful engagement from community members. Many families from marginalized communities, for instance, feel alienated from the special education decision-making process (Ong-Dean et al., 2011; Waitoller & Super, 2017; Golson et al., 2022). At the most basic level, school districts should eliminate obvious obstacles to broad participation. These steps include translation services, childcare during meetings, and remote meeting options. In short, enabling people who have been systematically excluded from decision-making processes to advocate for their children is only half of the solution; the other half is building a system that works even for children whose parents have not had the opportunity to have their voices heard in a meaningful way through advocacy.
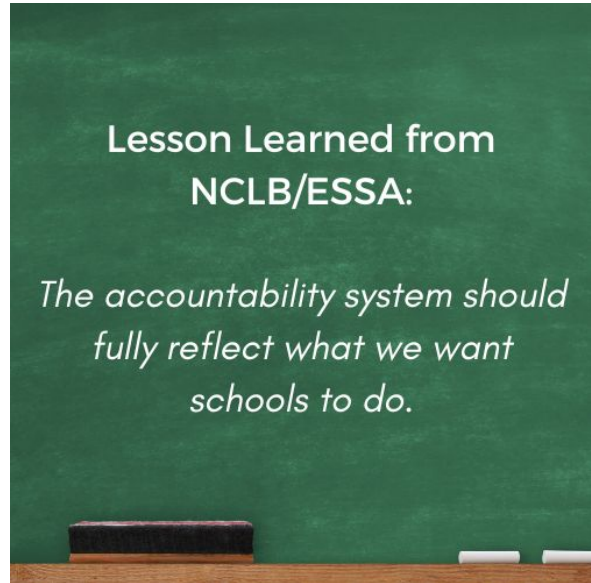
But beyond the obvious are other steps that leaders should take to address the actual dynamics of participation, protecting against domination and silencing tied to racism as well as unequal power. While current standards and assessment regulations require attention to diversity in establishing

---

[2] There is good reason to believe that inequity is rampant but currently unobservable. Beyond whether schools are fulfilling the formal letter of the IDEA law, there's the question of how services get written into students' IEPs. Students in higher income schools should not be systematically accessing more and better services, and the nation needs data systems to understand the extent to which this is happening (Ong-Dean et al. 2011).

content and performance standards (U.S. Department of Education Regulations), there have not been strong enough requirements or expectations around the inclusion of groups that have been historically marginalized from educational opportunities. States should actively involve such groups in system design, development, and evaluation. This could take the form, for example, of expanding the role of parents and community members that review state and local performance assessments in the development process and flag area of potential bias. We recognize the need for content-area expertise, and the logistical challenge of seeking to meaningfully engage members of the public. At the same time, we are heartened by the research on democratic deliberation, which, as James Fishkin writes, "might serve an advisory function for public policy" (Fishkin, 2002, p. 231). Although the particular details of such efforts will be the domain of future policymakers, scholars have outlined possible models for educational accountability (Gottlieb & Schneider, 2018).

Recommendation #4:
Move Toward a Broader Array of School Quality Indicators

The dominant approach to educational accountability for the past two decades has centered on standardized, quantitative measures of student learning and has weighted these measures disproportionately relative to other indicators. Centering these systems on test score results correspondingly excludes more holistic considerations of student learning and school performance (Lee & Lee, 2020; Gagnon & Schneider, 2019; Schneider, 2017). Multiple negative consequences emerge when focusing overwhelmingly on standardized tests: narrowing the curriculum to devote more time to tested subjects, teaching to the test, and taking valuable lesson time out of the school year (Booher-Jennings, 2005; Heilig & Darling-Hammond, 2008; Heilig, Young & Williams, 2012; Koretz, 2017).

Lesson Learned from NCLB/ESSA:
The accountability system should fully reflect what we want schools to do.

An accountability system that holds schools accountable using a comprehensive array of indicators requires four primary modifications to the current system, each of which is described below: lowering the stakes associated with standardized testing; decreasing the amount of testing that takes place; recognizing the many roles of our public schools; and maintaining comparability across schools, districts, and states.[3]

States and the Federal Government Should Lower the Stakes of Testing and Accountability

The dominant approach to assessment and accountability has been problematic in a number of different ways, as discussed throughout this report. Many of those problems, however, can be remedied or reduced by lowering the stakes that have been central to NCLB-style accountability. The theory of change underlying the past two decades of accountability has presumed that more pressure on educators and school leaders will lead inexorably to better results. This has not come to pass. Instead, and as predicted (e.g., Campbell, 1979), the high-stakes of modern accountability have led to a range of gaming behaviors. And, when coupled with the narrow design of present assessment and accountability systems—focusing chiefly on standardized test scores, which often indicate more about out-of-school factors than they do about in-school factors—these stakes can lead to actions that are directly opposed by the communities ostensibly being served. School closures, for instance, have disproportionately affected Black and brown families, as well as low-income families, who have often fought to keep their schools open.

---

[3] This recommendation focuses on the elements of the accountability system designed to hold schools accountable. Under a system of reciprocal accountability, other elements of the system address accountability to schools, their teachers, and their students.

Lowering the stakes of accountability is a logical response to the fact that theorized improvement in response to the threat of sanction has not materialized. Schools, it seems, do not improve in response to pressure; instead, and as discussed elsewhere in this report, they improve when their capacity is strengthened. Lowering stakes would also increase the accuracy and integrity of the data produced through assessment and accountability efforts. Reducing the incentive to game accountability indicators would foster a clearer and more precise picture of how schools are actually doing.

States and the Federal Government Should Lighten the Footprint of Summative State Testing

Along with lowering the stakes of standardized testing, a revised approach to assessment and accountability should allow states to lighten the footprint of summative state testing. This could be done by using matrix sampling, as exemplified by the approach used for the National Assessment for Educational Progress (NAEP). In a system with matrix sampling, a stratified sample of students can each answer a small subset of the entire assessment, minimizing the total number of hours dedicated to testing while preserving the ability to report on the performance of all students.

Administering assessments to only a sample of schools and students each year would free up school and district administrators from time dedicated to assessment preparation and administration. Recent legislation proposed in Congress would require states to use matrix or representative sampling, along with grade-span testing (More Teaching Less Testing Act, 2023). If the goal of state standardized testing is to hold schools, districts, and states accountable for closing achievement gaps and monitoring progress toward meeting this goal—rather than to continually assess every student's learning—then not every student needs to take the entire state battery of assessments. It is important to note, however, that transitioning from current models of census testing to matrix sampling-based approaches will limit many current uses of standardized tests. States and policymakers should be prepared to solve technical challenges this poses as well as revise the intended uses of standardized tests in an accountability system.

It is also important to note that lightening the footprint of annual testing would reduce the cost associated with standards-based accountability, allowing for reallocation of resources. The costs associated with alternatives to standardized tests—alternatives like performance assessments—would be reduced dramatically through such reallocation.

States and the Federal Government Should Include Indicators That Reflect the Many Roles of Public Schools

The nation's system of test-focused accountability has captured only narrow aspects of the nation's schooling goals (Rothstein & Jacobsen, 2006; Schneider et al., 2017, 2018, 2021, 2022). By shifting from a system that relies primarily on standardized test scores to one that is inclusive of all goals and aims of K-12 public schooling, accountability systems can offer a more holistic picture of student learning and school quality. This is important for a variety of reasons—it honors the full mission of public education, it stands to offer educators and the public better information for action, it is less likely to lead to corruption and gaming, and it will more accurately characterize school quality for the purpose of accountability.

A transformational approach to accountability indicators would require more than just new measures. As exemplified by California Local Control Accountability Plans (LCAPs), communities should be

central in the development of accountability approaches. With state procedural and financial support, districts (or consortia of districts, facilitated by the state) should be able to modify assessment and accountability frameworks in response to local priorities. School districts should be given the power to choose which indicators to hold schools accountable for, drawing from a set of options approved by the state. Just as importantly, the state must ensure that all stakeholders—particularly those historically and presently marginalized—are able to participate in shaping these frameworks (Anderson, 1998; Asen, 2021; Hargreaves & Braun, 2013; Marsh & Hall, 2018; Rothstein & Jacobsen, 2006; Schneider, 2017).

As discussed in Recommendation #2, this process also should include indicators of inputs and processes, at least at the district level (see, e.g., Carter & Welner, 2013; Elmore & Fuhrman, 1995; Jacob & Levitt, 2003; Koretz, 2017; McDonnell, 1995; Schneider, 2017). In particular, it should measure racial and socioeconomic disparities in the input and process factors that determine educational outcomes, arrived at through inclusive deliberation at the local level (Heilig et al., 2014; Lee & Wong, 2004). These factors might include, for example, equity in student disciplinary outcomes, access to high-quality teachers, access to enriching and culturally inclusive learning opportunities, and access to student support personnel.[4] Additionally, including opportunity to learn standards can highlight inputs and processes that schools, districts, and states should invest in and implement to provide students with equitable educational opportunities.

---

[4] States should also consider administering teacher perception surveys. Teacher retention and performance improves when they are afforded opportunities to make meaningful contributions to policy and practices (Bryk, 2010; Ingersoll, 2007, 2009). As such, states should be incentivized to collect statewide data on teachers' perceptions of their working conditions, to analyze those data, and to identify and rectify inequities across districts within the state and schools within districts.

Recommendation #5:
Ensure Interpretable and Actionable Results

The usefulness of assessment and accountability systems depends on the timeliness and trustworthiness of the information they provide, as well as on their helpfulness in informing classroom instruction. State testing systems have failed in this regard and should not be pursued for this purpose. A new vision of school quality measurement should prioritize providing useful information to the public and educators.

States Should Ensure that Assessments Exemplify Learning Goals and Support Program Improvement

States should design their accountability systems to help educators understand the content goals

and reasoning targets of the assessments, along with how students demonstrate those goals in various tasks and activities (Andrade, 2010; Black & Wiliam, 1998; Hattie & Timperley, 2007; William, 2010). To the maximum extent possible, educators should be involved in assessment scoring and moderation for the purpose of more deeply understanding student learning and, in turn, being able to use assessment data to revise local curricula and develop their own performance tasks and formative probes. Analyzing student work and applying criteria together with colleagues provides a means for professional development and more actionable insights for next steps in student learning (Gambell & Hunter, 2004; Gearhart & Osmundson, 2009; Palermo & Thomson, 2019).

As noted earlier in this report, states should support districts in developing or adopting high-quality curricula with embedded performance-based assessments. They should ensure that districts are able to provide professional learning opportunities that enable the integration of equitable classroom assessment practices and ambitious teaching (Shepard, Marion, & Saldaña, in press).

States Should Ensure that Results Are Interpretable for the Broader Public

Accountability data should be accessible to the broader public while also being useful for school leaders and other policymakers. Past attempts to use accountability systems to inform the public and stakeholders have been problematic, given the difficulty of balancing the sometimes-competing aims of simplicity and validity. For example, reducing school performance to an A-F letter grade fails to inform people because it fails to offer a valid and holistic picture of education or school performance (Schneider, 2017; Adams et al., 2015). This is particularly true in light of the fact that the primary components of most letter grade systems are student standardized test scores, which often indicate more about out-of-school factors than about in-school factors. While informing the public is a
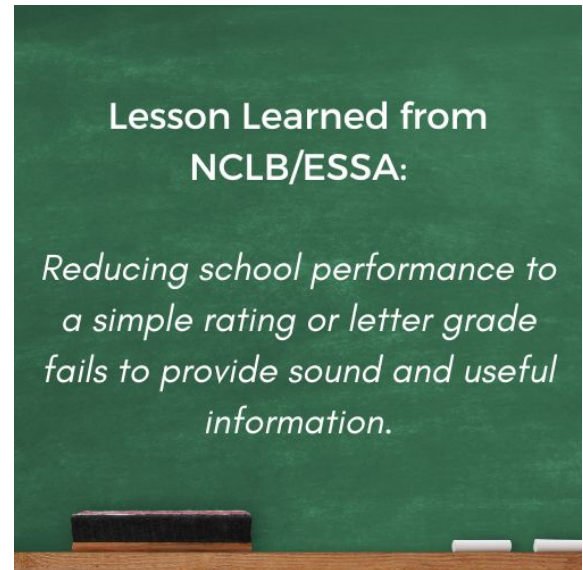
fundamental aim, states must recognize that when faced with the choice of offering more complex information or information that is misleading, they should opt for the former. Simplicity, in other words, is not informative if it offers a distorted picture of reality.

Rather than relying on obscure and summative evaluations of school quality, states can take steps to present data in a way that fosters understanding. One consideration in presenting data is the design of webpages or dashboards where data will be displayed. Thoughtfully considered visualizations can be informative, engaging, and accessible. A data dashboard, alone, though, is not enough. Supporting public understanding and deliberation with data is an ongoing process and states should create opportunities for the public to learn about and discuss data. Rather than conceptualizing accountability as a single metric that communicates to the public how a school is doing, we view accountability as a structure for inquiry, support, and dialogue.[5]

---

[5] Explaining the system and the data in languages in addition to English will often be necessary to accomplish this goal.

Recommendation #6:
Design a System That Will Evolve and Improve

A final important element of a revised accountability system involves the ongoing improvement of the system itself. Evaluative feedback loops should be built into every significant element of the accountability system in order to ensure that it is working as intended. For instance, feedback loops should help ensure that participation is authentic and representative, as set forth in Recommendation #3 (see, e.g., Abowitz, 2015; Abowitz & Mamlok, 2020), and to avoid capture of the process by special interest groups (Hartney, 2022; Murray et al., 2019; Polikoff, 2021; Sabin, 2021). The system should be designed to generate feedback, data, and evaluation-based responses as a regular part of each key process.

Lesson Learned from NCLB/ESSA:

A vitally important and complex accountability system requires continuous evaluation and improvment.

The K-12 public education system is extraordinarily dynamic, as are assessment and accountability systems. This complexity and dynamism counsels in favor of developing adaptive learning systems (see, e.g., Collinson & Cook, 2006; Fullan, 1994, 1999). Accordingly, there is no one single mechanism recommended by our participants for iteratively improving the system. Instead, we offer here one example, to concretize the sort of approaches that will be necessary.

We encourage strengthening federal- and state-level agencies that allow parents and other community members to report concerns of inequality and discrimination in schools (i.e., unmet needs or harmful practices that schools should be held accountable for). Such agencies must then work with local accountability bodies and community members to design and implement approaches to address problems. Conducting annual evaluations of this system is important, including a process evaluation to ensure that the system is being appropriately implemented, as well as an impact evaluation to ensure that results from the system are relevant, interpretable, and representative of the goals of public education. Results from the evaluation would provide insight into areas of success and areas that require changes. The federal and state agencies would then be responsible for attending to areas of the system requiring changes and continuing to monitor the system throughout the year.
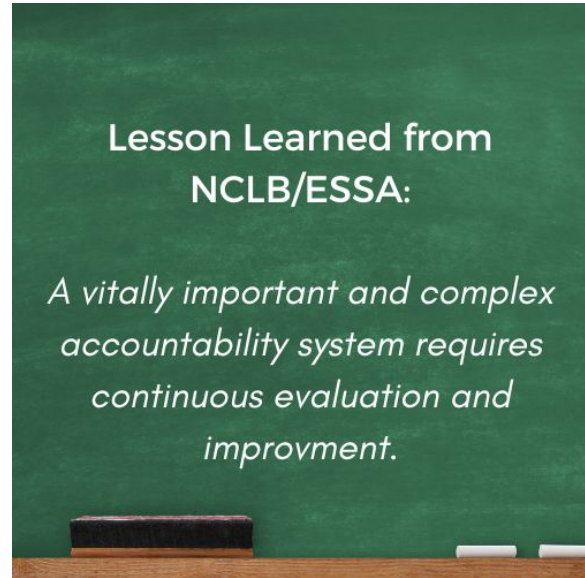
Discussion and Conclusion

This report reflects the insights of two dozen leading scholars of educational assessment and accountability. Notably, it does not reflect the insights of the school, district, and state personnel who are charged with implementing assessment and accountability systems. Inasmuch as that is the case, we offer our recommendations with an awareness that they are one piece of a larger whole, which educators and educational leaders will certainly need to have a hand in crafting.

Drawing on findings from across the field, this report synthesizes more than two decades of research to make a clear case for what might be termed Accountability 3.0. This vision is not motivated merely by what we know about how to improve assessment and accountability systems; it is also motivated by a fundamental concern with systemic inequalities. Accordingly, Accountability 3.0 is rooted both in research evidence and in the values-based goal of educational equity.

We are not alone in seeking to shape the future of educational assessment and accountability. As noted at the beginning of this report, a number of other efforts are presently under way—led by a range of different groups and informed by several distinct theories of change. In the coming years, as policy leaders begin to wrestle with that future, their efforts should be guided by evidence and core principles.

The principles guiding this report are congruent with what we have learned from educational research. And they are aligned with the fundamental aim of equity. In our view, those are the two basic requirements of any educational improvement effort. Our recommendations offer ways to enact those principles now, under ESSA, as well as in future iterations of ESEA.

Accountability mechanisms have value. We should hold each other accountable for the honest and efficient use of taxpayer dollars, for adherence to civil rights and anti-discrimination laws, and for providing equitable educational opportunities. Assessments can be a useful part of such efforts. But if the United States hopes to use educational assessment and accountability to substantially improve public schooling, we have significant work ahead of us. Even in the absence of values disagreements and policy disputes, it will be no small challenge to craft legislation that is simultaneously tight and loose—leveraging federal authority, while also devolving power to states and districts.

Yet it is also the case that our current approach is so obviously flawed that the bar for acting is unusually low. We are hopeful that even an imperfect revision of the Elementary and Secondary Education Act will do tremendous good, as long as it is guided by the principles set forth in this report.

In the meantime, school districts and states need not wait for a reauthorization of ESEA. The U.S. Department of Education can, for instance, work with states to leverage the Every Student Succeeds Act in ways that advance the six principles outlined in this report. There is no reason, for instance, that educational leaders cannot presently work to better center student learning, advance a vision of reciprocal accountability, or more meaningfully include community members. There is nothing barring the way from including a broader array of input and output indicators, offering more interpretable and actionable results, or seeking to improve over time.

Ultimately, we do hope that Congress will take up a now overdue reauthorization of ESEA. When lawmakers do so, we hope they will look to research for guidance, and that they will elevate educational equity as their primary objective. An entire community of scholars stands by, ready to assist.

References

Abowitz, K. K. (2015). *Publics for public schools: Legitimacy, democracy, and leadership.* Routledge.

Abowitz, K. K., & Mamlok, D. (2020). #NeverAgainMSD Student Activism: Lessons for agonist political education in an age of democratic crisis. *Educational Theory*, *70*(6), 731-748.

Adams, C.M., Forsyth, P.B., Ford, T.G., Ware, J.K., Barnes, L.B., Khojasteh, J., Mwavita, M., Olsen, J.J., & Lepine, J.A. (2015). *Next Generation School Accountability.* A Report Commissioned by the Oklahoma State Department of Education. Oklahoma Center for Education Policy (The University of Oklahoma) and The Center for Educational Research and Evaluation (Oklahoma State University).

Adams, J., & Grand, D. D. (2019). *Driving systems change and equity through project-based learning* (p. 66). Learning Policy Institute.

Aguilar, S. J., Holman, C., & Fishman, B. J. (2015). Game-inspired design: Empirical evidence in support of gameful learning environments. *Games and Culture*, *13* (1), 44-70.

AERA, APA, & NCME. (2014). *Standards for educational and psychological testing.* American Educational Research Association.

Anderson, G. L. (1998). Toward authentic participation: Deconstructing the discourses of participatory reforms in education. *American Educational Research Journal*, *35*(4), 571–603.

Andrade, H. L. (2010). Students as the definitive source of formative assessment: Academic self-assessment and the self-regulation of learning. In *Handbook of formative assessment* (pp. 90–105). Routledge.

Asen, R. (2021). Democracy, deliberation, and education. In *Democracy, Deliberation, and Education.* Penn State University Press.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 7–74.

Booher-Jennings, J. (2005). Below the bubble: "Educational triage" and the Texas accountability system. *American Educational Research Journal*, *42*(2), 231–268.

Bryk, A. S. (2010). Organizing schools for improvement. *Phi Delta Kappan*, *91*(7), 23–30.

Campbell, D. T. (1979). Assessing the impact of planned social change. *Evaluation and Program Planning*, *2*(1), 67–90.

Carter, P. L., & Welner, K. G. (2013). *Closing the opportunity gap: What America must do to give every child an even chance.* Oxford University Press.

Collinson, V., & Cook, T. F. (2006). *Organizational learning: Improving learning, teaching, and leading in school systems.* Sage.

Council of Chief State School Officers (CCSSO) (2022). An IMPD Policy Roadmap for States. https://753a0706.flowpaper.com/CCSSOIMPDPolicyRoadmap/

Davis, E. A., Palincsar, A. S., Smith, P. S., Arias, A. M., & Kademian, S. M. (2017). Educative curriculum materials: Uptake, impact, and implications for research and design. *Educational Researcher, 46*(6), 293-304.

DeBray, E., Finnigan, K. S., George, J., & Scott, J. (2022). A civil rights framework for the reauthorization of ESEA. National Education Policy Center.

Dee, T. S., Jacob, B., & Schwartz, N. L. (2013). The effects of NCLB on school resources and practices. *Educational Evaluation and Policy Analysis, 35*(2), 252-279.

Elmore, R. (2002). *Bridging the Gap between Standards and Achievement: The Imperative for Professional Development in Education.* Washington, DC: Albert Shanker Institute.

Elmore, R. (2005). Agency, reciprocity and accountability in democratic education. In *The institutions of American democracy: The public schools*, 277-301.

Elmore, R. F., & Fuhrman, S. H. (1995). Opportunity-to-learn standards and the state role in education. *Teachers College Record, 96*(3), 1–26.

Espinoza, M. L., Vossoughi, S., Rose, M., & Poza, L. E. (2020). *Matters of participation: Notes on the study of dignity and learning. 27*(4), 325–347.

Feldman, J. (2018). School grading policies are failing children: A call to action for equitable grading. Oakland, CA: Crescendo Education Group.

Fishkin, J. S. (2002). Deliberative democracy. *The Blackwell guide to social and political philosophy*, 221-238.

Fullan, M. (1994). Coordinating top-down and bottom-up strategies for educational reform. In R. Anson (Ed.) *Systemic Reform: Perspectives on Personalizing Education*, 7-24.

Fullan, M. (1999). *Change Forces: The Sequel.* Philadelphia, PA: Falmer.

Gagnon, D. J., & Schneider, J. (2019). Holistic school quality measurement and the future of accountability: Pilot-test results. *Educational Policy, 33*(5), 734-760.

Gambell, T., & Hunter, D. (2004). Teacher scoring of large-scale assessment: Professional development or debilitation? *Journal of Curriculum Studies, 36*(6), 697–724.

Gamson, D. A., McDermott, K. A., & Reed, D. S. (2015). The Elementary and Secondary Education Act at fifty: Aspirations, effects, and limitations. *RSF: The Russell Sage Foundation Journal of the Social Sciences, 1*(3), 1-29.

Gearhart, M., & Osmundson, E. (2009). Assessment portfolios as opportunities for teacher learning. *Educational Assessment, 14*(1), 1–24.

Golson, M. E., Roanhorse, T. T., McClain, M. B., Galliher, R. V., & Domenech Rodríguez, M. M. (2022). School-based ADHD services: Perspectives from racially and ethnically minoritized students. *Psychology in the Schools, 59*(4), 726-743.

Gottlieb, D., & Schneider, J. (2018). Putting the public back into public accountability. *Phi Delta Kappan, 100*(3), 29-32.

Hargreaves, A., & Braun, H. (2013). *Data-driven improvement and accountability* (p. 47). National Education Policy Center.

Hartney, M. (2022). *How Policies Make Interest Groups: Governments, Unions, and American Education.* Chicago: University of Chicago Press.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112.

Heilig, J. V., & Darling-Hammond, L. (2008). Accountability Texas-style: The progress and learning of urban minority students in a high-stakes testing context. *Educational Evaluation and Policy Analysis*, *30*(2), 75–110.

Heilig, J. V., Khalifa, M., & Tillman, L. (2013). Why have NCLB and high stakes reforms failed? Reframing the discourse with a post-colonial lens. *Handbook of urban education*, 523-537.

Heilig, J. V., Young, M., & Williams, A. (2012). At-risk student averse: Risk management and accountability. *Journal of Educational Administration*.

Horsford, S. D., & Heilig, J. (2014). Community-based education reform in urban contexts: Implications for leadership, policy, and accountability. *Urban Education*, *49*(8), 867–870.

Ingersoll, R. (2007). Short on power, long on responsibility. *GSE Publications*, 129.

Ingersoll, R. M. (2009). *Who controls teachers' work? Power and accountability in America's schools*. Harvard University Press.

Jackson, C. K. (2020). *Does school spending matter? The new literature on an old question.* American Psychological Association.

Jacob, B. A., & Levitt, S. D. (2003). Rotten apples: An investigation of the prevalence and predictors of teacher cheating. *The Quarterly Journal of Economics*, *118*(3), 843–877.

Jennings, J. L., & Bearak, J. M. (2014). "Teaching to the test" in the NCLB era: How test predictability affects our understanding of student performance. *Educational Researcher, 43*(8), 381-389.

Jennings, J., & Sohn, H. (2014). Measure for measure: How proficiency-based accountability systems affect inequality in academic achievement. *Sociology of Education, 87*(2), 125-141.

Kirp, D. L. (2013). Improbable scholars: The rebirth of a great American school system and a strategy for America's schools. Oxford University Press.

Kirsch, I., & Braun, H. (2016). *The dynamics of opportunity in America: Evidence and perspectives*. Springer Nature.

Koretz, D. (2008). *Measuring up: What educational testing really tells us*. Harvard University Press.

Koretz, D. (2017). The testing charade. In *The Testing Charade*. University of Chicago Press.

Krajcik, J. S., McNeill, K. L., & Reiser, B. J. (2008). Learning-goals-driven design model: Developing curriculum materials that align with national standards and incorporate project-based pedagogy. *Science Education*, *92*(1), 1-32.

Ladson-Billings, G. (2006). From the achievement gap to the education debt: Understanding achievement in U.S. schools. *Educational Researcher*, *35*(7), 3-12. https://doi.org/10.3102/0013189X035007003

Lee, J., & Lee, M. (2020). Is "whole child" education obsolete? Public school principals' educational goal priorities in the era of accountability. *Educational Administration Quarterly*, *56*(5), 856–884.

Lee, J., & Reeves, T. (2012). Revisiting the Impact of NCLB High-stakes School Accountability, Capacity and Resources: State NAEP 1990-2009 Reading and Math Achievement Gaps and Trends. *Educational Evaluation and Policy Analysis, 34*(2), 209-231.

Lee, J., & Wong, K. K. (2004). The Impact of Accountability on Racial and Socioeconomic Equity: Considering both School Resources and Achievement Outcomes. *American Educational Research Journal, 41*, 797-832.

Loeb, S. & Byun, E. (2019). Testing, accountability, and school improvement. *The ANNALS of the American Academy of Political and Social Science*, *683*(1), 94–109.

Marion, S. F., Vander Els, J., & Leather, P. (2017). Reciprocal accountability for transformative change: New Hampshire's performance assessment of competency education. *Voices in Urban Education*, *46*, 20–25.

Marsh, J. A., & Hall, M. (2018). Challenges and choices: A multidistrict analysis of statewide mandated democratic engagement. *American Educational Research Journal*, *55*(2), 243–286.

McDonnell, L. M. (1995). Opportunity to learn as a research concept and a policy instrument. *Educational Evaluation and Policy Analysis*, *17*(3), 305–322.

More Teaching Less Testing Act, H.R. 1741, 118th Cong. (2023). https://www.congress.gov/bill/118th-congress/house-bill/1741?q=%7B%22search%22%3A%5B%22jamaal+bowman%22%5D%7D&s=3&r=10

Murray, B., Domina, T., Renzulli, L., & Boylan, R. (2019). Civil Society Goes to School: Parent-teacher Organizations and the Equality of Educational Opportunity. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, 5(3), 41-63.

National Academies of Sciences, Engineering, and Medicine (NASEM). (2019). Monitoring educational equity. National Academies Press.

National Academies of Sciences, Engineering, and Medicine (NASEM). (2018). How people learn II: Learners, contexts, and cultures. National Academies Press.

National Council on Measurement in Education (NCME). (2023). *Foundational competencies in educational measurement: Presidential taskforce report of the National Council on Measurement in Education.* https://higherlogicdownload.s3.amazonaws.com/NCME/4b7590fc-3903-444d-b89d-c45b7fa3da3f/UploadedImages/NCME_Task_Force_on_Foundational_Competencies_-_Final_Report_v08.pdf

National Research Council. (2011). *Incentives and test-based accountability in public education.* Washington, DC: The National Academies Press.

Neri, R. C., Lozano, M., & Gomez, L. M. (2019). (Re) framing resistance to culturally relevant education as a multilevel learning problem. *Review of Research in Education*, *43*(1), 197-226.

Oakes, J. & Welner, K. (2023). "From Schools of Opportunity to Systems of Opportunity," in Adam York, Kevin Welner, and Linda Molner-Kelley (eds.), *Schools of Opportunity: 10 Research-Based Models for Creating Equity in Action.* (pp. 193-217). New York: Teachers College Press.

Ong-Dean, C., Daly, A. J., & Park, V. (2011). Privileged advocates: Disability and education policy in the USA. *Policy Futures in Education, 9*(3), 392-405.

Palermo, C., & Thomson, M. M. (2019). Large-scale assessment as professional development: Teachers' motivations, ability beliefs, and values. *Teacher Development*, *23*(2), 192–212.

Partnership for the Future of Learning. (2022). *Racially just assessment and accountability initiative.* Bill & Melinda Gates Foundation.

Penuel, W. R. (2021). Possible futures for equitable educational assessment. *Phi Delta Kappan*, *103*(4), 54-57. https://doi.org/10.1177/00317217211065829

Pinder, K. A. (2010). Federal demand and local choice: Safeguarding the notion of federalism in education law and policy. *Journal of Law & Education, 39*, 1-36.

Polikoff, M. (2021). *Beyond standards: The fragmentation of education governance and the promise of curriculum reform.* Harvard Education Press.

Polikoff, M. S., Porter, A. C., & Smithson, J. (2011). How well aligned are state assessments of student achievement with state content standards? *American Educational Research Journal*, *48*(4), 965–995.

PRRAC. (2011, February 2). NCSD issue brief 1: Key principles for ESEA reauthorization (February 2011). *PRRAC — Connecting Research to Advocacy.*

Rolstad, K. & MacSwan, J. (2023). Bilingual language assessment: The persistence of a special case of bias. In A. K. Kibler, C. J. Faltis, A. Walqui & G. Bunch (eds.) *Courageous Contributions: Guadalupe Valdés' Enduring Work Toward Equity in Multilingual Schools and Communities.* Clevedon: Multilingual Matters.

Rothstein, R., & Jacobsen, R. (2006). The goals of education. *Phi Delta Kappan*, *88*(4), 264–272.

Sabin, P. (2021). *Public Citizens: The Attack on Big Government and the Remaking of American Liberalism.* New York: W.W. Norton & Co.

Schneider, J. (2017). *Beyond test scores: A better way to measure school quality.* Harvard University Press.

Schneider, J. & Hutt, E. L. (2023). *Off the Mark: How Grades, Ratings, and Rankings Undermine Learning (But Don't Have To).* Cambridge: Harvard University Press.

Schneider, J., & Saultz, A. (2020). Authority and control: The tension at the heart of standards-based accountability. *Harvard Educational Review*, *90*(3), 419-445.

Schneider, J., Jacobsen, R., White, R. S., & Gehlbach, H. (2018). The (mis)measure of schools: How data affect stakeholder knowledge and perceptions of quality. *Teachers College Record*, *120*(5), 1–40.

Schneider, J., Jacobsen, R., White, R., & Gehlbach, H. (2017). Building a better measure of school quality. *Phi Delta Kappan*, *98*(7), 43–48.

Schneider, J., Noonan, J., White, R. S., Gagnon, D., & Carey, A. (2021). Adding "student voice" to the mix: Perception surveys and state accountability systems. *AERA Open*, *7*, 2332858421990729.

Schneider, J., Piazza, P., White, R. S., & Carey, A. (2022). Student experience outcomes in racially integrated schools: Looking beyond test scores in six districts. *Education and Urban Society*, *54*(3), 330–360.

Sebring, P. B., Allensworth, E., Bryk, A. S., Easton, J. Q., & Luppescu, S. (2006). *The essential supports for school improvement.* Consortium on Chicago School Research.

Shepard, L. A., Marion, S. F., Saldaña C. M. (in press). Standards-based reform and school accountability. In L. Cohen-Vogel, J. Scott, & P. Youngs (Eds.), *Handbook of education policy research (2nd Vol.).* Washington, DC: American Educational Research Association.

The Future of Assessment and Accountability Project. (2022). *Emerging areas of agreement for the future of assessment and accountability.* The National Urban League and UnidosUS.

Thomas, J. Y., & Brady, K. P. (2005). Chapter 3: The Elementary and Secondary Education Act at 40: Equity, Accountability, and the Evolving Federal Role in Public Education. In *Review of Research in Education* (pp. 51–67). Sage Publications.

Tyler, R. W. (1950). Basic principles of curriculum development. *Chicago: University of Chicago Press.* *Performance Improvement Quarterly, 3*(1), 65-70.

Vasquez Heilig, J., Ward, D. R., Weisman, E., & Cole, H. (2014). Community-based school finance and accountability: A new era for local control in education policy? *Urban Education, 49*(8), 871-894.

Waitoller, F. R., & Super, G. (2017). School Choice or the Politics of Desperation? Black and Latinx Parents of Students with Dis/Abilities Selecting Charter Schools in Chicago. Education Policy Analysis Archives, 25(55).

William, M. K. (2010). The implementation of school based continuous assessment (CA) in Zambia. *Educational Research and Reviews, 5*(3), 099–107.

Zúñiga, X., Naagda, B. R. A., & Sevig, T. D. (2002). Intergroup dialogues: An educational model for cultivating engagement across differences. *Equity & Excellence in Education, 35*(1), 7-17.

Appendix A
Description of Spring 2022 Convening

This report reflects the collective ideas of two dozen national experts in assessment, learning and accountability. All but three of these participants convened in person in Boulder, Colorado for a day-and-a-half in the spring of 2022. Costs of the convening were covered by funding from the National Education Association. Prior to the in-person convening, we asked the experts two framing and generative questions: "What are the key topics we should be sure to address in this work?" and "What is the purpose of educational accountability?"

Based on the responses to those questions, we created eight main areas of focus, around which groups of experts worked together to craft their recommendations. Most of the convening was then devoted to discussing research, deliberating about larger questions, considering how the future of testing and accountability can be effective and equitable, and exploring policy recommendations that followed from the research—and then refining those ideas. In the months following the in-person meeting, the organizers assembled the ideas and then sent a written draft to the working-group experts for written feedback. That feedback was then used to develop this final report.

Appendix B
Working Group Mission Statement

The Elementary and Secondary Education Act (ESEA) has been reauthorized eight times since 1965, at a rate of once every seven years. The last time it was reauthorized, as the Every Student Succeeds Act (ESSA), it preserved most of what No Child Left Behind (NCLB) had put in place. Looking ahead to 2022, seven years after the passage of ESSA and 20 years after NCLB was signed into law, it seems unlikely that ESEA reauthorization will be taken up by Congress. Yet whatever the delay this time around, we see a value in developing a coherent vision for what reauthorization might achieve. Particularly given the mounting dissatisfaction with test-based accountability and a greater readiness, at least in some quarters, to confront racism and racialized inequality, we want to consider new directions and policy approaches for school improvement.

Thinking specifically about the measurement and accountability requirements of ESSA, we seek to outline a more useful and humane approach, rooted in research evidence. What elements of current federal measurement and accountability requirements should be preserved? What elements should be peeled away? What new elements should be introduced? How can measurement and accountability best support anti-racist school improvement efforts? We intend to answer these questions by thinking specifically about what states are currently required by federal law to do, as well as what they are currently prevented from doing in light of ESSA's mandates. What should be measured at the state level? How should those measures be used in service of accountability? Other than formal accountability, what are the other uses for measurement? What kinds of flexibility should be built into federal law?

Our fundamental concern is with access to educational opportunity. Insofar as that is the case, we also seek to consider how measurement and accountability systems might be augmented to better address systemic inequalities. Specifically, we will consider how opportunity-to-learn standards and indicators might advance a system of reciprocal accountability. Since the passage of NCLB, national and state policy has been grounded in outcome measures used to hold schools accountable for results. Yet if schools do not have the resources required to create equality of opportunity for young people, even the fairest and most complete measurement system will fail to produce that desired outcome. In order for schools to succeed at their work, another form of accountability is required—one in which the functional capacity of schools is assured by state and federal agencies.

Appendix C
Working Group Bios

Alfredo J. Artiles

Alfredo J. Artiles is the Lee L. Jacks Professor of Education at Stanford University. He is the Director of the Stanford Center for Opportunity Policy in Education and Director of the Research Institute at the Center for Comparative Studies in Race and Ethnicity. He edits the book series *Disability, Culture, & Equity* (Teachers College Press). Artiles has been appointed to three consensus panels of NASEM. He is an elected member of the National Academy of Education, Fellow of AERA, and Senior Research Fellow at the Learning Policy Institute. He was a resident fellow at CASBS.

Derek Briggs

Derek Briggs is a professor in the School of Education at the University of Colorado Boulder where he also directs the Center for Assessment Design Research and Evaluation. Dr. Briggs's research focuses upon advancing methods for the measurement of student learning. He works with states and other entities to provide technical advice on the design and use of large-scale student assessments. He is past editor of *Educational Measurement: Issues and Practice*, president of the National Council on Measurement in Education for 2021-22, and author of the book *Historical and Conceptual Foundations of Measurement in the Human Sciences*.

Elena Diaz-Bilello

Elena Diaz-Bilello is the Associate Director at the Center for Assessment, Design, Research, and Evaluation (CADRE) at CU Boulder's School of Education, and collaborates with state agencies, school districts, and educational organizations to develop practical and sound approaches for addressing assessment and educational policy challenges. Prior to joining CADRE, Elena was a Senior Associate at the National Center for the Improvement of Educational Assessment providing technical assistance and guidance to state agencies, the U.S. Department of Education, national organizations, and school districts in the areas of designing validity and program evaluation studies and improving upon accountability and assessment practices.

Edward García Fierros

Dr. Edward Garcia Fierros is Associate Dean for Diversity, Equity, and Inclusion in the College of Liberal Arts and Sciences at Villanova University. Dr. Fierros is an associate professor in the Department of Education and Counseling. His areas of study are aimed at providing equitable opportunities for all learners. His expertise includes testing and measurement, diversity and equity in assessment, placement patterns of students with special needs and students that are English Language Learners, as well as educational policy related to underrepresented students.

David R. Garcia

David R. Garcia is an associate professor in the Mary Lou Fulton Teachers College at Arizona State University. Dr. Garcia's professional experience includes extensive work in education policy development and implementation. His academic work centers on school choice, accountability, and research utilization. In 2018, he published *School Choice* (MIT Press). His current book, *Teach Truth to Power* (2021), also by MIT Press, is on the intersection between research, policy, and politics.

Derek Gottlieb

Derek Gottlieb is an associate professor of Educational Foundations and Curriculum Studies at the University of Northern Colorado. His research interests center on the role of public institutions, especially schools, in shaping and responding to democratic practices. He is the author of *A Democratic Theory of Educational Accountability* (2020) and *Education Reform and the Concept of Good Teaching* (2015). His current book project is tentatively entitled *No Way to Reform Schools: Recalibrating Expectations in the 21st Century*.

Julian Vasquez Heilig

Julian Vasquez Heilig is Provost and Vice President for Academic Affairs at Western Michigan University. His research and practice primarily focus on K-12 education policy and leadership that impacts community-based equity and innovation. Dr. Vasquez Heilig has conveyed invited testimony in state and national legislative bodies and volunteered expertise and provided empirically-based input focused on community-based education reform. For this work, he was honored with the passage of California Assembly Resolution 1459, which commended his state and national impact in education policy. He has also served as a volunteer education policy advisor and provided input for presidential and gubernatorial campaigns.

Ethan L. Hutt

Ethan Hutt is an associate professor in the School of Education at the University of North Carolina - Chapel Hill. His research focuses on the historical development and current use of school metrics, paying particular attention to the numbers we use to describe, define, and evaluate American schools. He is co-author (with Jack Schneider) of a forthcoming book with Harvard University Press exploring the use of grades, standardized tests, and transcripts in American education. Dr. Hutt is also an associate editor of *Educational Researcher* and co-editor of *High School Journal*.

Nathan Jones

Nathan Jones is an associate professor in the Special Education Program at Boston University Wheelock College of Education & Human Development. He is affiliated faculty with the Wheelock Educational Policy Center (WEPC) and is a founding member of the BU Faculty of Computing & Data Sciences. Dr. Jones's research has been supported through grants from the National Science Foundation, the Institute for Education Sciences, the William T. Grant Foundation, and the Spencer Foundation, among others. His research has appeared in the *Journal of Human Resources*, *Educational Evaluation and Policy Analysis*, *Educational Researcher*, *AERA Open*, *Exceptional Children*, *Remedial and Special Education*, and *Educational Assessment*.

Jaekyung Lee

Jaekyung Lee is a professor of education at the University at Buffalo, SUNY. A fellow of the American Educational Research Association (AERA) and a Fulbright Global Scholar, Dr. Lee specializes in educational policy research and evaluation, particularly on the issues of educational equity and accountability in the U.S. and across the world. He has authored and edited several books on this topic, including *The Testing Gap* (Information Age Publishing, 1997), *The Anatomy of Achievement Gaps* (Oxford University Press, 2016), and *Centering Whole-Child Development in Global Education Reform* (Routledge, 2022).

Jeff MacSwan

Jeff MacSwan is Professor of Applied Linguistics and Language Education at the University of Maryland. He is also Professor of Neuroscience and Cognitive Science, and affiliate Professor in the Department of Linguistics and the Maryland Language Science Center. His research focuses on the linguistic study of bilingualism and codeswitching (or language alternation), and its implications for theories about the role of language in educational settings for multilingual students. Dr. MacSwan is a Fellow of the American Educational Research Association (AERA) and of the National Education Policy Center. He received the Lifetime Achievement Award from AERA's Bilingual Education Research SIG and the Leadership through Scholarship Award from AERA's Second Language Research SIG, both in 2021.

Scott Marion

Scott Marion is Executive Director of the National Center for the Improvement of Educational Assessment. He is a national leader in conceptualizing and designing innovative and balanced assessment systems and accountability reform to support instructional and other critical uses. Dr. Marion serves on the National Assessment Governing Board, overseeing the National Assessment of Educational Progress and coordinates and/or serves on ten state or district Technical Advisory Committees for assessment and accountability. In addition, Dr. Marion just completed nine years of service on his local (Rye, NH) School Board.

Katie McDermott

Kathryn A. McDermott is a political scientist by training who has been on the faculty of both the College of Education and the School of Public Policy at the University of Massachusetts Amherst since 1999. Her research focuses on the origins and implementation of policies intended to make education more equitable, including the standards-based accountability movement and student-assignment plans to support diverse schools. She is currently a co-PI on a study of educational antidiscrimination policy as a means of sustaining multiracial democracy in the U.S, funded by the Spencer Foundation.

Bill Penuel

Bill Penuel is a Distinguished Professor of Learning Sciences and Human Development in the School of Education and Institute of Cognitive Science at the University of Colorado Boulder. He designs and studies curriculum materials, assessments, and professional learning experiences for teachers in science. He works in partnership with school districts and state departments of education, and the research he conducts is in support of educational equity in three dimensions: (1) equitable implementation of new science standards; (2) creating inclusive classroom cultures that attend to students' affective experiences and where all students have authority for constructing knowledge together; and (3) connecting teaching to the interests, experiences, and identities of learners.

Jennifer Randall

Jennifer Randall is the Dunn Family Chair of Psychometrics and Test Development at the University of Michigan and the founding President of the Center for Measurement Justice. Her work seeks to disrupt white supremacist, racist logics in assessment through culturally sustaining practices that are explicitly and unapologetically antiracist. She is committed to working with minoritized communities and our co-conspirators to explore the ways in which we can create a justice-oriented assessment system culture in which the sociocultural identities of students are deliberately considered and valued, not as an afterthought, but rather in the planning and development phases of assessment.

Ricardo D. Rosa
Ricardo Rosa is the Director of the Center for Education Policy & Practice at the Massachusetts Teachers Association. He is a former associate professor and current part-time lecturer in the Department of Public Policy at the University of Massachusetts Dartmouth where he specializes in equity and educational policy. He is a community-engaged scholar and organizer who works with educators and youth. He situates his work within the field of critical pedagogy and critical educational policy analysis.

Christopher Saldaña
Christopher Saldaña is an assistant professor in the department of Educational Leadership and Policy Analysis at the University of Wisconsin-Madison. His research examines the relationship between K-12 school finance and educational opportunity. Chris received his Ph.D. in Educational Foundations, Policy, and Practice from the University of Colorado, Boulder.

Andrew Saultz
Andy Saultz is a professor and Interim Dean in the College of Education at Pacific University. His scholarship focuses on improving equity in schools, focusing on how educational accountability policy impacts and is interpreted by a broad range of actors. The primary aim of his research is to understand: 1) how policymakers across levels of government develop school accountability policy; 2) how teachers and educational leaders implement and react to new accountability policies; and 3) what data and information parents and citizens use to evaluate school quality. Dr. Saultz enjoys coaching his two sons in tee-ball, hiking, and exploring Oregon wine country.

Jack Schneider
Jack Schneider is Dwight W. Allen Distinguished Professor at the University of Massachusetts Amherst, where he is the Director of the Center for Education Policy. In addition to his scholarly work on assessment and accountability, Dr. Schneider is also the co-founder of a statewide consortium of districts piloting an alternative approach in Massachusetts; that work has been supported by the Massachusetts State Legislature and the Spencer Foundation. The author of five books, he writes frequently about education in outlets like the *New York Times* and the *Washington Post*; he is also the co-host of the educational policy podcast "Have You Heard."

Benjamin Shear
Dr. Shear is an assistant professor in the Research and Evaluation Methodology program at the CU Boulder School of Education, and Faculty Partner of the Center for Assessment, Design, Research & Evaluation. His primary research interests address topics in psychometrics and applied statistics, including validity theory, differential item functioning, and the use of educational tests.

Lorrie A. Shepard
Lorrie Shepard is University Distinguished Professor in the School of Education at the University of Colorado Boulder. Her research in educational measurement has addressed the use and misuse of tests in educational settings. Most cited are her contributions to validity theory, standard setting, bias detection, the effects of high-stakes testing, and the integration of learning theory with classroom formative assessment. In recent years Dr. Shepard's work has focused on drawing deeper connections between sociocultural learning theory and culturally responsive pedagogy and assessment practices.

Stephen Sireci
Stephen Sireci is Distinguished University Professor at the University of Massachusetts Amherst, where serves as Executive Director of the Center for Educational Assessment. He is a Fellow of the American Educational

Research Association, a Fellow of Division 5 of the American Psychological Association, and past Co-Editor of the *International Journal of Testing* and the *Journal of Applied Testing Technology*. Dr. Sireci is past president of the National Council on Measurement in Education, current President of the International Test Commission, and architect of the Massachusetts Adult Proficiency Tests

David C. Webb

David C. Webb is an Associate Professor of Mathematics Education at the University of Colorado Boulder. Dr. Webb's research interests include teachers' design and use of classroom assessment and the design of professional development activities to develop teacher expertise in classroom assessment. He currently focuses on the design of instructional resources that promote active learning, computational thinking, and executive function in K-16 STEM education. Webb was a secondary mathematics and CSEd teacher in Southern California. He teaches courses for prospective mathematics and science teachers, including courses that focus on assessment design and practice and the development of student-centered instruction.

Kevin G. Welner

Kevin Welner is professor of education policy at the University of Colorado Boulder School of Education and at the School of Law (by courtesy). He's also the director of the National Education Policy Center, housed at CU Boulder, which works to build bridges between the research world and the broader public. He's a Senior Research Fellow at the Learning Policy Institute and an AERA Fellow. He's the author of a dozen books, including 2013's Closing the Opportunity Gap (co-edited with Prudence Carter), which is the foundation for his recent work about the importance of improving children's opportunities to learn inside and outside of school.

Rachel S. White

Rachel S. White is an assistant professor at the University of Tennessee. Her research agenda centers around 1) issues of power, voice, diversity, and inclusion in education policy making and implementation processes, and 2) examining structures and policies that contribute to or counteract equitable and socially just K-12 education systems. As such, Dr. White examines the politics of education policy making and implementation, focusing on whose voices are heard in the policy making and implementation processes, and how decisions made by political and educational leaders at the school, district, and state level impact educational experiences of students and teachers.