



*Time to steer systems  
toward better balance  
and coherence.*

**Chris Domaleski**

## **Breakthrough or Breakdown? School Accountability in Flux**

Newton's second law of motion states that when force is applied to an object at rest, the rate of change is directly proportional to the force applied. There is little doubt that tremendous force has been applied to education recently, including assessment and accountability systems. First, COVID-19 prompted rapid shifts to distributed learning, suspension of state assessments, and waivers for school accountability. Then the collective anguish over George Floyd's murder in Minneapolis led to a sense of urgency for advancing equity in all institutions, including education.

Already under fire before spring 2020, the dominance of federally mandated test-based accountability in American education had led many educators and policymakers to decry the system as largely out of balance and to suggest that this imbalance has stifled productive local efforts toward meaningful, lasting improvements in student learning. It is not too early to conclude that all these cumulative factors will—and should—change assessment and accountability systems. But what kind of change is appropriate, and how can state boards of education support such changes?

## What Is the Promise of School Accountability?

Before state boards can help nudge their systems back into balance, they should examine the purpose of accountability systems in the first place. Broadly, school accountability can be thought of as a system that 1) signals what outcomes are valued, 2) provides information about school performance with respect to those outcomes, and 3) prescribes a system of supports and interventions based on performance. Improvement is thought to occur by incentivizing the right kinds of behaviors and actions, shining a light on areas where improvement is needed, and providing targeted supports to those areas. That may sound straightforward enough, but this portrayal is built on scores of assumptions and a vast network of actions and interactions that are best addressed in a strong theory of action.

Accountability systems may highlight goals and benchmarks and provide some useful information to guide actions, but real educational progress always has been pegged to the practice of teaching and learning that occurs daily in classrooms. In order to promote school improvement, systems must activate the conditions and supports that provide students an opportunity to learn. School improvement requires attention to the “instructional core,” with these principles in mind:

There are only three ways to improve student learning at scale: You can raise the level of the content that students are taught. You can increase the skill and knowledge that teachers bring to the teaching of that content. And you can increase the level of students’ active learning of the content. That’s it. Everything else is incidental.<sup>1</sup>

Ultimately, I argue that accountability systems can play a role in an overall plan to promote student success, but they are not a “treatment” and far from a holistic prescription for education reform. In fact, too often contemporary accountability systems are built on an impoverished theory of action—that suggests putting data in the hands of policymakers or educators will lead to strategic actions to improve schools.

What then are the key factors essential for leveraging the promise of accountability to improve schools? I will describe three:

principled design, balanced and coherent structures, and reciprocal support.

## Principled Design

There is no gold standard, no single correct approach, for developing and implementing systems that monitor and support school improvement. A system should take into account policy priorities, local conditions, and context. Only then can education leaders design a solution that supports guiding principles and values.

Engaging in a principled design process is best accomplished through a well-explicated theory of action.<sup>2</sup> A theory of action is a tool for both designing and evaluating accountability systems by clarifying goals, assumptions, and the hypothesized mechanisms to bring about the intended changes. Randy Bennett’s recommendations for an assessment system theory of action can be extended to accountability. Bennett’s conceptualization includes

- the components of the system and the rationale, grounded in research and theory, for these components;
- the claims that will be made from results or outcomes;
- the intended effects of the system;
- the mechanisms thought to cause the intended effects; and
- potential unintended or negative effects and the plan to mitigate them.<sup>3</sup>

Developing a strong theory of action starts with clarifying the highest priority goals. For example, the system can be designed to privilege closing achievement gaps or preparing students for postsecondary success. While a system can have multiple goals, focus is important. If an initiative is thought to do everything well, there is a good chance it will not do anything very well.

The value of a theory of action comes from documenting the hypothesized connections between and among the actions and the outcomes they are thought to promote. At a high level, Erika Hall and colleagues recommend portraying these elements with respect to inputs, outcomes, and evidence.<sup>4</sup>

Inputs describe the resources, actions, and conditions that will be necessary to support improvement. It includes the source—federal,

---

**Developing a strong theory of action starts with clarifying the highest priority goals.**

---

**Developing the theory of action is a shared responsibility, which should be led by policymakers and leaders.**

state, district—and nature of the support. For example, inputs might include financial resources and wraparound student support services.

With respect to outcomes, the theory of action should specify the proximal and distal effects the system will promote. For example, instructional practices will improve, family and community engagement will grow, or students will encounter and meet higher academic expectations. Importantly, the supporting conditions and rationale thought to promote the outcome should be made explicit.

Finally, the theory of action specifies the evidence to support connections among inputs and outcomes. For example, improved school climate is documented by survey data, instructional practices are appraised via interviews and observations, and tracking disaggregated growth rates provides insight on academic progress for students in traditionally underserved groups. A systematic collection of evidence can help leaders engage in ongoing monitoring, review, and refinement of the system.

Only by investing upfront in the hard work of creating a well-specified theory of action can the role of assessment and accountability in supporting improved outcomes be understood and evaluated. Developing the theory of action is a shared responsibility, which should be led by policymakers and leaders, including chief state school officers and state boards of education, in collaboration with practitioners, technical advisors, and a broad-based group of stakeholders.

## **Balance and Coherence**

Currently, nearly all the attention on school improvement is connected to state accountability systems that were designed to meet federal requirements as specified in the Every Student Succeeds Act (ESSA). These systems are overwhelmingly influenced by academic indicators in the form of performance on end-of-year summative tests in English language arts and mathematics. This represents a woefully imbalanced, incoherent system.

Creating more balance and coherence involves attending to the emphasis and interrelationships within and among systems at the federal, state, and local levels. To start, it is useful to briefly distinguish the roles at each of level.

**The Federal, State, and Local Role.** Equity is at the core of the Elementary and Secondary Education Act (ESEA) of 1965, an initiative

designed to improve educational opportunities for disadvantaged students. Early federal accountability provisions typically focused on compliance or inputs.<sup>5</sup> However, the scope has grown over the over the years, with perhaps the most pronounced pivot occurring with No Child Left Behind (NCLB), which mandated annual statewide achievement testing and school accountability systems that predominately linked consequences to attainment of proficiency on these tests.<sup>6</sup> Accountability has not been the same since.

State educational authority is drawn from each state's constitution and requirements from the legislature, the state board, and other governing bodies. Chief among the states' roles are to establish the content and rigor of the academic standards, implement the state assessment and accountability systems, and provide support and resources to help districts and schools meet performance expectations.<sup>7</sup> Naturally, federal requirements heavily influence much of a state's efforts in this endeavor, as states must meet the detailed strictures of ESSA and federal peer review before implementing state assessment and accountability systems.

Despite the prominent influence of federal and state authority, education remains primarily a local responsibility. Local school boards, responding to community priorities, and district leaders are charged with the essential front-line responsibilities: hiring and supporting educators and staff, establishing and implementing the curriculum, and managing day-to-day operations such as transportation, facilities, and food service.

**Balance.** Balance primarily refers to the development of systems that are well specified at each level and pegged to the appropriate areas of emphasis. Unfortunately, contemporary accountability systems at the district level, if they exist at all, often treat districts like a “super school” by simply aggregating all the school metrics to the district. Moreover, these school metrics typically mirror the state's standardized model. As previously noted, this model primarily comprises performance on end-of-year tests.

Districts can achieve more balance in at least two ways. First, they can design their systems to reflect the specific goals and priorities of the district and community, especially with respect to the areas under the districts' direct influence. Second, the systems can be specified at a much finer grain size that takes advantage

### Box 1. Promising Initiatives

While much work remains to improve school accountability, there are some promising initiatives. One is the Massachusetts Consortium for Innovative Education Assessment (MCIEA), a collection of school districts and partners that are reimagining assessment and accountability. Their framework explicitly links essential inputs (e.g., educator and leader development, school culture, resources to support teaching and learning) and key outcomes such as academic learning, community building, and well-being. The model puts classroom performance assessment in the foreground in lieu of high-stakes state summative testing and includes a variety of broad school quality measures that are informed by research and community input and linked to prioritized outcomes. By so doing, the model moves toward more balanced, coherent accountability.

Innovation at the state level can be more difficult, not least because of federal constraints. Some states have overcome this by pursuing initiatives outside of ESSA. Recognizing that a one-size-fits-all system does not work well for alternative high schools, the Wyoming Department of Education engaged in a multiyear initiative to produce a supplemental school accountability system.<sup>a</sup> The resulting framework was built on a distinct theory of action that recognizes the unique mission of alternative schools and incorporates a broad, flexible set of indicators selected to support prioritized outcomes. For example, the system promotes student engagement and the holistic development of skills associated with postsecondary success via mechanisms such as individualized student success plans. While this system addresses a relatively small number of schools, the process and resulting framework are instructive.

Numerous states have signaled a move toward more improved school accountability. For example, New Hampshire's Performance Assessment of Competency Education (PACE) could serve as a mechanism to more tightly link curriculum, instruction, and assessment (see article, page 39). Georgia's College and Career Ready Performance Index encourages schools to promote readiness in areas such as fine arts and languages and rewards attainment of a variety of college- and career-ready credentials. This flexibility supports balance, allows for appropriate differentiation, and ultimately provides more useful, actionable outcomes. The Louisiana Department of Education helps districts implement strong practices by 1) reviewing extant curriculum and assessment resources, results of which are publicly available, 2) providing training and instructional resources, and 3) curating a bank of model resources and instructional tools.

<sup>a</sup>Chris Domaleski and Erika Hall, "Wyoming Alternative School Accountability Framework: Recommendations from the Alternative Accountability Advisory Committee" (Dover, NH: National Center for the Improvement of Educational Assessment, 2015).

of local initiatives and inputs. For example, because the district has auspices over personnel and professional development, they are best positioned to monitor and evaluate the efficacy of programs such as new teacher induction, professional development initiatives, and mentoring programs. As another example, schools and districts may have access to data on student performance apart from statewide tests. More focused information provided in a more timely manner is more likely to influence changes in practice.

States play a part in promoting balance as well. Although federal requirements constrain state systems, states can create supplemental initiatives that need not be high stakes, insofar as they are focused on producing performance classifications. For example, states can curate model resources and research-based improvement practices or help collect and report meaningful data that goes beyond summative assessments. As another example, some states have developed differentiated accountability and support systems for alternative schools outside ESSA. Taken together, there are multiple ways for states to partner with districts to promote balance (see box 1).

**Coherence.** As Ben Forman, Charles DePascale, and I detail elsewhere, coherence in accountability addresses at least three dimensions:

- **External coherence:** Are connections among multiple accountability systems logically consistent?
- **Internal coherence:** Are comparisons within the local system logically related?
- **K-12 coherence:** As students advance from early grades to graduation, are the different levels of the system logically connected?<sup>8</sup>

---

**More focused information provided in a more timely manner is more likely to influence changes in practice.**



---

**State systems can be poorly specified, such that they thwart innovation at the local level.**

External coherence is evident when systems at the state, federal, and local levels are mutually supportive. The responsibility goes both ways. Unfortunately, state systems can be poorly specified, such that they thwart innovation at the local level. For example, a state system that places too much emphasis on participation or performance on Advanced Placement courses as a pathway to postsecondary success may create a perverse incentive for districts and schools to eschew a broader range of academic and cocurricular experiences that prepare students for success in college and careers.

In systems with internal coherence, the components are mutually supportive and aligned to the overarching system goals. Incoherence can manifest in several ways. It may occur when systems of support are not connected to primary outcomes of interest. For example, if a central objective of the system is to close achievement gaps but there is no plan to provide supports to students in historically lower performing groups, the system is not internally coherent. There are technical features that can contribute to incoherence, as well. A common example is the disproportionate focus on proficiency rates in most systems, which is ill suited to gauge progress for students below standard, in contrast to academic growth measures.

Finally, K-12 coherence, which could also be termed vertical coherence, refers to a system that is thoughtfully designed to support students' success throughout their educational pathways. For example, Scott Marion and colleagues emphasize the importance of learning progressions as the "organizing framework for connecting various assessments and learning activities in a vertically coherent system."<sup>9</sup> Unfortunately, it is all too common for schools to administer large-scale commercial assessments at regular intervals to gauge student progress with little or no information about the degree to which they are providing useful feedback on the skills most important for students to demonstrate success as they progress through the curriculum.

## **Reciprocal Support**

The idea of reciprocity as a key factor in developing effective accountability systems is not new.

Reciprocity refers to the shared responsibility to support attainment of performance expectations. As Richard Elmore explains, "For every increment of performance I demand from you,

I have an equal responsibility to provide you with the capacity to meet that expectation."<sup>10</sup>

Arguably, contemporary accountability systems have overstated classifications and underspecified the mechanisms for building capacity and providing support to educators and students. In fact, it might be more appropriate to shift the language from "school accountability systems" to "school support systems" to reflect the proper emphasis on support as the key element in any well-designed system.

Given the central importance of reciprocity in promoting improved systems, it is useful to further develop two essential elements to reciprocal support: differentiation and utility.

**Differentiation.** One might challenge the notion that schools are the sole locus of accountability. Brian Gong explores this concept using the term "differentiated" accountability:

Advocates of differentiated accountability may argue that just as it isn't reasonable to hold students accountable for meeting standards until they have been given a fair opportunity to learn, so it is not reasonable to hold schools accountable until schools have been given a fair opportunity to provide the opportunity to learn.<sup>11</sup>

He suggests that an accountability system can be considered incomplete if it lacks details about the obligations and consequences associated with all entities responsible for establishing the conditions for success. This information should be included in the guiding theory of action. It stands to reason that developing such a system will also promote more balanced, coherent structures.

**Utility.** Utility refers to the extent to which information generated in the system is appropriate and actionable. As noted previously, this requires having indicators represent the outcomes of interest and are appropriately specified to detect the desired effects.

Moreover, utility is supported when the timing and manner of reporting are sufficient to inform a helpful response. With little exception, current accountability practice provides information on distal outcomes at the end of the academic year or later. Improved systems will be characterized by data collection and reporting systems that provide more signals along the way to indicate if the desired outcomes are on track. By providing information at regular

intervals throughout the year (e.g., attendance, performance on interim assessments), personnel will be better equipped to detect and prevent a negative outcome.

It should be obvious that such data and support systems must be decoupled from high-stakes classifications if they are to be useful. To the extent that data intended for ongoing monitoring also inform summative classifications, the value of that information will almost certainly diminish due to Campbell's Law. Named for social psychologist Donald Campbell, this principle holds that the likelihood of an indicator being corrupted increases in proportion to the degree to which that indicator is used for consequential decision making.<sup>12</sup>

## Final Thoughts

It is not easy to change the status quo in American public education. But perhaps the moment is right to consider a new path forward for school accountability that is more credibly linked to improved outcomes for all students.

What are some specific actions that state boards can take to better leverage the promise of accountability?

- Work to develop a comprehensive theory of action for school improvement informed by a broad-based and diverse group of leaders and stakeholders. Regard this theory of action as a dynamic document, returning to it often to refine assumptions and guide monitoring and support.
- Use the theory of action as the foundation for reexamining the coherence and balance of the state's accountability system. Refine the system as needed to ensure that it helps the state measure and promote what matters most. Consider including a broad set of indicators that go beyond summative assessments and measures of performance that can provide useful signals of student progress during the academic year. Such indicators may be supplemental—that is, decoupled from the federal system.
- Advocate for state practices that support districts and schools to stand up strong and complementary practices for monitoring and support. These systems should reflect local values, goals, and responsibilities.

- Promote the curation and dissemination of research and resources to help districts and schools. For example, provide model assessment and instructional resources or evaluate the quality of commercial products with respect to the state's academic expectations.

To be fair, it will take more than “a better system” or sheer force of will to realize long overdue reform that promotes better outcomes. Wraparound services and supports to address the needs of historically underserved students must be vigorously pursued. State and federal laws must be friendlier to innovation. And sustained efforts to build capacity to improve teaching and learning are critical. Now is the time to shift focus away from counterproductive practices and toward more promising alternatives. ■

<sup>1</sup>Elizabeth A. City et al., *Instructional Rounds in Education: A Network Approach to Improving Teaching and Learning* (Cambridge, MA: Harvard Educational Press, 2003).

<sup>2</sup>Marianne Perie, “Key Elements for Educational Accountability Models” (Washington, DC: Council of Chief State School Officers, 2007); Erika Hall et al., “A Framework to Support Accountability Evaluation” (Washington, DC: Council of Chief State School Officers, 2016).

<sup>3</sup>Randy Elliot Bennett, “Cognitively Based Assessment of, for, and as Learning: A Preliminary Theory of Action for Summative and Formative Assessment,” *Measurement: Interdisciplinary Research and Perspectives* 8 (2010): 70–91.

<sup>4</sup>Hall et al., “Framework to Support Accountability Evaluation.”

<sup>5</sup>Susan H. Fuhrman and Richard F. Elmore, eds., *Redesigning Accountability Systems for Education* (New York: Teachers College Press, 2004).

<sup>6</sup>Margaret E. Goertz, “Implementing the No Child Left Behind Act: Challenges for the States,” *Peabody Journal of Education* 80 (2005): 73–89.

<sup>7</sup>Chris Domaleski, Damian Betebenner, and Susan Lyons, “Promoting More Coherent and Balanced Accountability Systems” (Dover, NH: National Center for the Improvement of Educational Assessment, 2018); Joanne Weiss and Patrick McGuinn, “The Evolving Role of the State Education Agency in the Era of ESSA and Trump: Past, Present, and Uncertain Future,” *CPRE Working Papers* (2017).

<sup>8</sup>Ben Foreman, Charles DePascale, and Chris Domaleski, “Local Accountability: The Forgotten Element in Education Reform” (Boston: Massachusetts Institute for a New Commonwealth, 2018).

<sup>9</sup>Scott Marion et al., “A Tricky Balance: The Challenges and Opportunities of Balanced Systems of Assessment” (Dover, NH: National Center for the Improvement of Educational Assessment, 2018), p. 3.

<sup>10</sup>Richard Elmore, *Bridging the Gap between Standards and Achievement: The Imperative for Professional Development in Education* (Washington, DC: Albert Shanker Institute, 2002), p. 5.

<sup>11</sup>Brian Gong, “The Next Generation of State Assessment and Accountability” blog post (Dover, NH: National Center for the Improvement of Educational Assessment, April 2, 2020).

<sup>12</sup>Donald T. Campbell, “Assessing the Impact of Planned Social Change,” *Evaluation and Program Planning* 2, no. 1 (1979): 67–90.

**To the extent that data intended for ongoing monitoring also inform summative classifications, the value of that information will almost certainly diminish.**

**Chris Domaleski is associate director at the National Center for the Improvement of Educational Assessment.**