



Curriculum-Embedded Performance Assessment

An Update for Leaders

Table of Contents



01 [Foreword](#)

02 [Introduction](#)

04 [Curriculum-Embedded Performance Assessment](#)
[Research & Evidence](#)
[Best Practices](#)
[Pathways to Scale](#)
[Policy and Accountability](#)
[Trends in Education](#)

09 [Conclusion](#)

11 [APPENDIX A: PARTICIPANTS](#)

12 [End Notes](#)

Foreword

The experience of human life is a series of performance assessments. From daily routines, like showing up on time for work, to high-pressure moments, such as delivering a closing argument or sale, we are constantly tested in various ways. Yet, for many, the assessments we experienced in school were often not well aligned to the real-world challenges we face as adults.

The K–12 education experience—and its associated assessments—has struggled to mirror the types of complex, integrated, and practical knowledge applications we need to succeed in life. Many of us have witnessed individuals who excelled within the structured environment of school but struggled with the ambiguity and multifaceted nature of life beyond the classroom. Conversely, some were told by the education system that they weren't capable, only to go on to achieve significant success in the real world. Perhaps you know such a person—or perhaps you are one.

Despite our lived experiences telling us that traditional assessments fall short, the educational system of tests and measures has struggled to evolve at the scale and speed necessary to meet the demands of our rapidly changing future. This gap is especially evident when we consider the complex, real-world skills that students need to develop.

This PBLWorks white paper seeks to raise the issues associated with a shift to curriculum-embedded performance assessments that not only better prepare young people for the complexities of life but also provide more meaningful insights into the effectiveness of instruction and the quality of our schools. This white paper draws on the insights of leading experts and practitioners who demonstrate that incorporating performance assessments into the DNA of education is neither a novel experiment nor a leap of faith. On the contrary, a strong research base supports the efficacy and impact of these assessments, which are already in use in thousands of classrooms in the United States and around the world.

Performance assessments offer an authentic, time-tested method of evaluating not just what students know but what they can do with that knowledge. As the discussions in this brief emphasize, this approach moves beyond merely measuring rote memorization. Instead, it aligns with the deeper learning objectives that prepare students with the knowledge and skills that are essential in today's fast-paced, increasingly automated, and globally interconnected world.

As educational leaders and policymakers look for ways to ensure that students are truly prepared for the future, curriculum-embedded performance assessments provide a proven path forward. These assessments represent not only a way to enhance student learning but also a tool to help educators and administrators improve the effectiveness of teaching and the overall performance of schools. In a time when education is evolving rapidly, the need for such transformative approaches has never been more urgent.

Jason E. Glass, Ed.D.

Associate Vice President – Teaching & Learning
Western Michigan University
Former Director of the Iowa Department of Education
Former Commissioner of the Kentucky Department of Education



Introduction

For decades now, educational innovators across the United States have recommitted to setting 21st-century learning goals and standards bolstered by related improvements including: shifting to high-quality student-centered instruction aimed at those goals; developing more rigorous and relevant instructional materials to support such teaching; and implementing better assessments that provide more useful information for instruction than is typical of annual, standardized testing.

These innovations include but are not limited to:

- **Deeper Learning** – the higher-order thinking skills, learning dispositions, and collaboration skills needed for students to succeed in 21st-century work and civic life. Deeper Learning competencies promote the ability to transfer learning and apply learning to new and complex situations in an ever-changing global environment. (As defined by the Deeper Learning Hub)¹
- **Portrait of a Graduate** – a school or district’s collective vision for student success. Also known as a Graduate Profile, Portrait of a Learner, Habits of Success, Lifelong Learning Standards, and other expressions, this vision describes the set of skills and competencies that a community agrees its young people need to thrive in learning, work, and life. (As defined by Next Generation Learning Challenges)²
- **Project Based Learning (PBL)** – a teaching method in which students learn by actively engaging in real-world and personally meaningful projects. Students work on a project over an extended period of time—from a week up to a semester—that engages them in solving a real-world problem or answering a complex question. They demonstrate their knowledge and skills by creating a public product or presentation for a real audience. (As defined by PBLWorks)³

There are other important approaches as well, many of which emphasize student-centered learning. This kind of work has been conducted in many settings: at purpose-built charter schools, at high-achieving public districts, at struggling schools—indeed, some of this innovation is to be found in many communities across the nation. Although it is far from the norm everywhere, there has been growing participation in these movements by states, districts, and communities, fueling real expansion in such approaches.

As these innovations reach an even greater scale, supportive system leaders are faced with the vital question: How will we know what’s working and for whom, and how can evidence be used to support further student learning? How will we know that our new goals are being taught, our materials are well-aligned, and our students are achieving? Existing assessments—particularly standardized tests that rely heavily on selected-response items—are often poor measures of these more ambitious goals. Further, as instruction increasingly values the diversity of students’ ideas and experiences, there is a need for assessments to shift in parallel. As instruction cultivates more complex, nuanced thinking, it is increasingly difficult to find assessments that are well matched to these new learning opportunities. An emerging answer to the need for better measures is curriculum-embedded performance assessment:

- **Performance Assessment** – an approach to educational assessment that requires students to directly demonstrate what they know and are able to do through open-ended tasks such as constructing an answer, producing a project, or performing an activity. This demonstration can include generating a short written response, writing an analytical essay, conducting a science investigation, creating a curated portfolio of work, or developing an original research paper. (As defined by the Learning Policy Institute)⁴

In February 2024, a group of experts (Appendix A) gathered together to explore what we know today about performance assessment that is tightly aligned with curriculum, what we still need to know more about, and the implications for education tomorrow. They discussed the state of the research and evidence base, best practices in implementation, pathways to greater scale, related policy and accountability implications, and future education trends. Their discussions have been distilled here in each area so that you can benefit from their expertise as you think about these issues in your state or district, classroom or community.

Their reflections also inform an emerging set of cross-cutting principles that can guide leaders as they begin to pursue curriculum-embedded performance assessment reform initiatives—or expand the work they have already begun. State agency and district leaders should be intentional about their assessment innovation efforts and role—forming a coherent point of view, committing to co-creating with locals, and aiming at eventual widespread adoption. They should pursue, as well as contribute to, evidence-based work—soliciting broader school and teacher participation in research initiatives and embracing local action research. Initiatives should be equity focused—affording all students the opportunity to demonstrate both progress toward and mastery of content as well as use of skills across the state’s richer range of academic subjects and educational goals (not just a small set of federally-mandated subject tests); and local assessment results should be actionable for improving instruction, directing learning, and closing achievement gaps, not just for documenting them. Work should be coherent—creating a holistic approach to creating a balanced assessment system that is transparently supported by policymakers and communities alike. Lastly, the capacity of educators and leaders should be supported with professional learning to tie assessment to curriculum and to use the results of better measures for instruction, learning, and systemic accountability.

The aim of the meeting was to raise important issues in the design and use of performance assessment and to foster a robust back-and-forth discussion among a range of experts from across the educational ecosystem, not to arrive at consensus. Accordingly, the following summation of the conversations is not intended to imply endorsement from any specific individual, agency, or institution.

We thank the William and Flora Hewlett Foundation for sponsoring this initial set of discussions, and we invite you to continue them among your own colleagues in your own context.



Curriculum-Embedded Performance Assessment

Below is a summary of the discussions held by 16 experts—ranging from teachers to state education agency leaders, researchers to intermediaries, academics to practitioners.

Research & Evidence

A strong research base undergirds the design and use of high-quality performance assessment. Research tells us that performance assessment has an effect on the enacted curriculum—that when performance assessments are part of the mix, teachers focus less with students on rote facts and more on higher-order thinking. Researchers can also demonstrate that existing performance assessments are predictive of key future student outcomes. However, we don't have robust enough research that enables advocates to make strong causal inferences about how and when performance assessment works, and under what conditions, as distinct from interrelated, more holistic improvements in teacher capacity and curriculum quality and relevance.

Performance assessments are deliberately designed not solely to assess specific academic knowledge but to assess students' application of knowledge and skills to new or novel contexts—vital learning that is not covered in a state-wide annual standardized test. Curriculum-embedded performance assessments are necessary in order to measure the depth of what a state's standards and graduate profiles currently require. Therefore, performance assessment evidence should count for grades, high school graduation, college admissions and placement, and for local contributions to statewide accountability measures; indeed, in many places it already is used in these kinds of ways. Future research should concern itself with these core purposes and measures of education, and the evidence and methodologies best used to support them.

The educational ecosystem—researchers, policy makers, state and local agencies, and others—need to accept more forms of deep, systematic research that tells us something about what is and isn't working, why, and for whom. We need an aggressive research agenda that includes local, short-cycle research projects that include both empirical data and the case studies and stories that provide additional context. State Education Agencies (SEAs) should connect more with other agencies, higher education, and other sectors to build multiple measures for—and richer ways of—evaluating students' college, career, and civic readiness and to establish a research agenda about what's working. Advocates and school leaders might call upon the many and varied learning communities within their own innovation networks to recruit their participation in the next set of necessary research on curriculum-embedded performance assessment.

Best Practices

Curriculum, instruction, and performance-based assessments should be well aligned, and all should be supported by professional learning. Curriculum and testing designers should work backwards from the learning goals of a unit—the standards-based content and success skills—to determine assessment priorities. One integrated approach is Project Based Learning. For example, in the PBLWorks curriculum, high-quality assessment is curriculum embedded and used primarily for the purposes of understanding, informing, and improving teaching and learning; assessment is designed to inform instruction and not just to provide a grade. Assessment includes knowledge and skill acquisition and use as well as application. The speed of results, who sees the results, and how they are used matters to engaging students and mitigating disconnection from learning and providing information that is relevant and meaningful both in and outside of the classroom.

High quality PBL that delivers more rigorous academic content standards and disciplinary mindsets as well as success skills⁵—along with embedded performance assessment—is more complex than traditional instruction. There needs to be associated professional learning for teachers and others to help ensure high-quality implementation. This should include professional development on how to use performance assessment results to improve instruction and determine student accomplishment and achievement, including both academic outcomes and competencies. Teacher support needs to be offered through and with curriculum-embedded performance assessment products.

Assessments should strive for cultural relevance but should not be labeled as such unless it has truly been intentionally and meaningfully incorporated into the design of curriculum, instruction, and assessments. Through what they include in assessment and rubrics, designers can place value on certain behaviors, actions, and perspectives that might not otherwise be reflected.



Pathways to Scale

It is important to address the extent to which transformational change moves away from being a limited experiment to being the way everything works, and to attend to the pathways toward that scale. A key question is, “What is the opportunity we want to meet with performance assessment?” We must then design with great urgency, but we need to execute with great intentionality.

States and districts can have a different starting point for their assessment reform discussions—state-initiated, district-initiated, policy-driven, grant-funded, building on longtime discussions, a new initiative, etc. But all leaders have control over how they do their work, which is often what matters most. For their part, states can encourage and make space for local innovation. New initiatives in Hawaii⁶, Kentucky⁷, and New York⁸ for performance assessment—while different from one another—each have modest federal support.

C.E. Coburn's re-definition of getting to scale may be a useful framework for performance assessment reformers: Scaling incorporates *depth of implementation* (the nature of the change in classroom instruction), *sustainability* (reinforcement and ongoing learning with connections to other reformers, and alignment with district policy), *spread* (widely shared norms, principles, and beliefs), and *shift in ownership* (authority for the reform moves to districts, schools, and teachers who themselves have the capacity to sustain, spread, and deepen the reform).⁹

It is vital to create coherence in statewide reform. District and state alignment is key so that elements of the system mesh in ways that make sense for local leaders, teachers, and students. Signals must be sent that performance assessment matters so that we do not continue to face a separated “two systems” model—one of mandatory state standardized tests that are comprised of mainly selected and constructed response items, and the other a nascent and tenuous performance assessment system.

With the idea of ensuring such symmetry across the system in mind, states can encourage, make space for, and even directly support local innovation. If they move too quickly, and without local engagement, they risk perpetuating harm to communities, especially those with high numbers of Black and Brown students and low-income families. We need to move, therefore, “as fast as we can but as slow as we must.”

Additionally, when a state mandates a set of directions to a place that is way ahead of where the field is, the result is often very performative local responses; and that doesn't change teaching and learning. There is an important role for community engagement in local, regional, and statewide policy conversations; such engagement is already commonly employed by states and districts when developing a portrait of a graduate.

Policy & Accountability

Some current assessment and accountability policies—particularly those that rely primarily on standardized testing scores in a limited number of academic subjects—are viewed by many as a barrier to achieving growth in performance assessment implementation at scale. With standardized assessment there is both a restrictive set of outcomes being measured and a restrictive way of measuring those outcomes. There will need to be a shift in federal assessment and accountability policies if advocates want to see a significant change in the use of performance assessment at scale.

A driving question is: How do we better mirror the learning process in the assessment and accountability process? Beginning policy engagement with questions of what we want from schooling is a powerful starting point. Leaders should ask communities what is important to them as co-creators, then build a coalition to answer that question, and only then start a formal policy process. It is important to have ongoing meaningful roles for the community—and for students—to define what is important (as is often done with development of district and state portraits of a graduate) so that the right things are being measured. Ongoing care should also be taken to ensure that the data reported is accessible and aligned to this shared view so that families, students, and community members can also participate in the sense-making of that data. These processes must attend to racial and economic disparities in community involvement and help make sure a new system of accountability doesn't reproduce the harms of the past.

The state approach should be about supporting assessment flexibility so that districts can innovate towards a shared vision. How can states generate inspiring opportunities for locals? An important emerging concept is that of “Reciprocal Accountability”¹⁰—a policy frame at the state level (and federal level) that limits the negative stakes attached to state testing to make room for districts and schools to measure other things that are important to their communities.

Performance assessments are not a simple swap-out for existing standardized testing, and reciprocal accountability offers room for additional efficient and effective measures, rather than simply seeking to take all prior measures away. Again, such new measures of knowledge and skills might stem from a local district's Portrait of a Graduate or statewide learning goals.

The time is ripe for this policy approach. There is today significant flexibility for states to start this work under the current system, but it means that the state must make the decision to begin, to engage locals, and to help interested/affected parties get on the same page. Indeed, U.S. Secretary of Education Miguel Cardona authored a 2022 letter¹¹ asking for more state leadership around assessment innovation. As states look to innovate, federal funding is also available as demonstrated by the Competitive Grants for State Assessments (CGSA) support that has helped new initiatives in Hawaii, Kentucky, and New York as well as other states. Some experts now believe that an accountability system that incorporates robust performance assessments (not just short performance tasks) could make it through federal assessment peer review if a state eventually decided to use it for federal reporting.

Trends in Education

We must remind and caution ourselves that, very often, innovations benefit those schools and families that are already the best resourced. As we look ahead to the future trends in education that will affect performance assessment we can ask a range of provocative questions: What needs to change in policy and practice if we are to take a more student-centered, human-centered design approach to assessment for all schools? How will we shift instructional frameworks, adopt culturally responsive practices, and change policy frameworks if we become more student-centered? How can all this promote equity?





Key questions that inform the future include:

What do culturally-responsive models of performance assessment mean for us, and what are the models of it? For reforms to take off in the future, we need proof points—and it is critical to keep “local” local; we are not trying to create a standardized model of local assessment. The movement needs broader evidence of what this work looks like and its impact on change, which would allow leaders to focus on evidence-based reform.

What are the best practices to support teachers in using curriculum-embedded performance assessment data effectively and efficiently to answer questions about student learning, support instruction, and enable improved learning opportunities to be provided to students? We are lacking sufficient teacher professional learning for performance assessment—especially training in how to use the new richer and more complex data that comes from the performance assessments to support improvements in teaching and learning.

How do we address the fact that the U.S. educational system is historically very resistant to change and often slow to innovate or accept innovations? As we work on a new generation of instruction and assessment, we may simply need to focus initially on building what we want and not try to make it retro-fit into the present system or waste time and energy engaging in systems reform debates per se. We must attend to change management as assessment reforms grow in scale; we must resist the simplistic urge to back-map performance assessment innovation onto dominant state annual testing paradigms; and we must foster substantive engagement by local educators and leaders—not satisfying ourselves with performative adherence to a one-size-fits-all mandatory approach without first attending to Coburn’s scale concerns.

What will be the effects of technology, including Artificial Intelligence (AI) on performance assessment design, implementation, evaluation, and scoring/feedback? We need to make sure technology is a natural, embedded part of the assessment process, especially as technology offers the promise of providing students with faster feedback on complex projects that often take a long time to score and providing teachers with broad insights into patterns and trends across student responses that could inform future instruction and curriculum revision efforts. There are two vital categories of tech innovation: better addressing problems we can't solve right now with our current strategies alone, and using technological innovation to help us improve what we're already doing. AI poses some real threats, but perhaps AI could be trained to actively disrupt bias and pay attention to students on the margins. Students should be brought into the conversation about AI use. Those investing in educational technology should act in partnership with educators and other experts who really understand the nature of learning and child development, and folks who care about the social-cultural context of learning.

Conclusion

The group of experts who convened in February 2024 explored a number of areas of vital interest to leaders who are considering curriculum-embedded performance assessment—or who are contemplating expanding their current work with the intent of going to greater scale. While not implying consensus or endorsement by any specific individual, agency, or institution, their discussions inform an emerging set of cross-cutting principles:

Intentional

SEA and LEA leaders can encourage and support efforts to begin—or expand—use of curriculum-embedded performance assessment right now. There are many examples of fellow states and districts currently engaged in this work. The federal government has been increasingly supportive. But system leaders must take the time to develop a point of view co-created with schools and communities about the value of local innovation. They must help convene the relevant actors in their state and support connections between innovators. Regardless of the pace at which they intend to proceed to scale, they must ultimately aim toward fully incorporating performance assessments as a meaningful part of what is normal in education.

Evidence-Based

The growing field of performance assessment is grounded in a strong evidence base with the potential to expand even more quickly if it can find appropriate financial support and engage school participants more broadly. Reformers must remain committed to evidence and balance expensive and time-intensive randomized controlled trials (RCTs) with faster, more applied local action research.

Equity-Focused

More robust assessment is needed that measures the full range and depth of a state or district's desired educational outcomes for all students, not merely shallow acquisition of knowledge in a limited number of core academic subjects. Aligning those assessments closely to the curriculum offers results that will be useful to teachers and learners as well as for auditing purposes. In a balanced assessment system that incorporates multiple measures, testing results should not stop at the level of documenting basic achievement gaps—they should reflect community values, be actionable for improving instruction, and be a meaningful part of accountability systems.

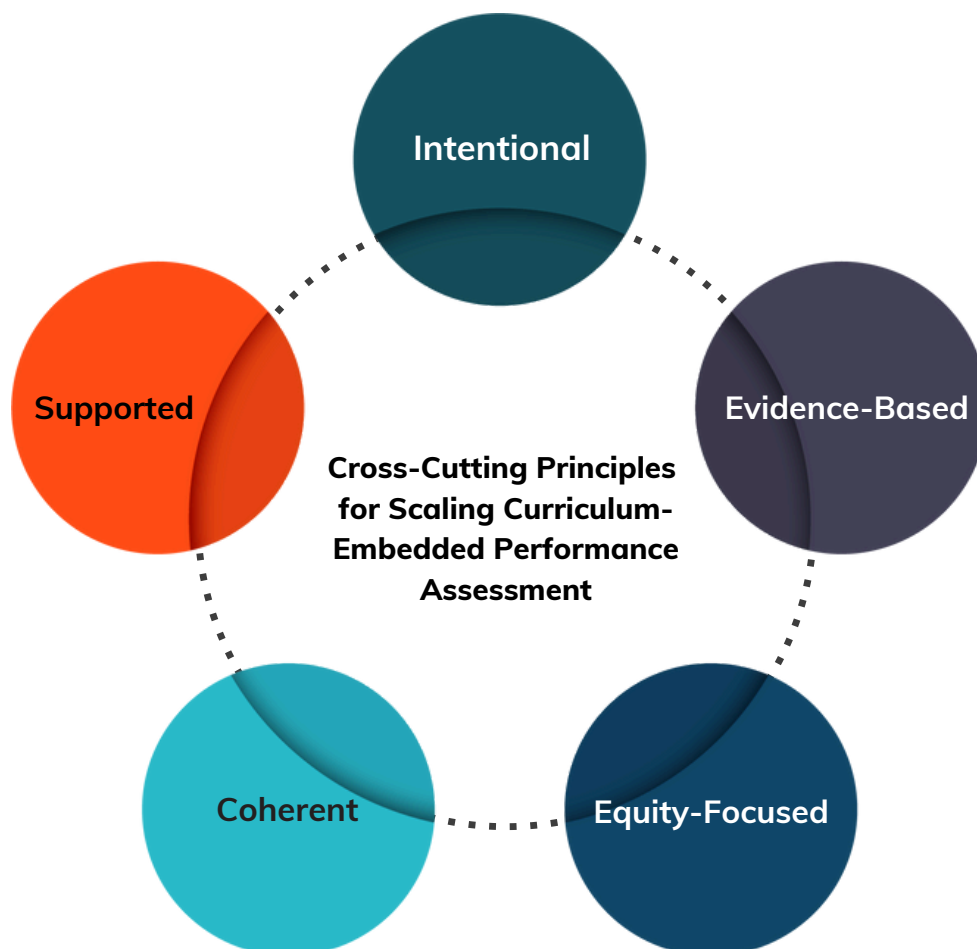
Coherent

Assessment reform should make sense up and down the system. Different states and districts will start their journey from different places, engage a varying range of local and national partners, and support their work in myriad ways from federal awards, foundation grants, their own budgets, or other sources...but each should take deliberate steps to support local innovation within the framework of a coherent approach that is transparently supported by policymakers. Often, vanguard districts and states begin by encouraging robust discussions of what local communities want the graduate profile of their learners to look like and then seek to make room in their assessment system for local contributions that reflect the resulting knowledge, skills, and dispositions.

Supported

Richer education goals, student-centered instruction, more robust materials, and embedded performance assessments are, frankly, more demanding of leaders, educators, and students. States and locals must support educators to plan for and use the results of better measures for instruction, learning, and accountability. There is no one-size-fits-all approach, but as systems adopt materials and assessments, leaders should ensure that publishers and partners are committed to necessary professional learning in design and implementation.

With these principles in mind, states, districts, and schools can take the next step in curriculum-embedded performance assessment. Experts are optimistic: it is safe to get going; there is no need to strive for a single mandated solution; multiple models exist for instructional materials, performance assessments, starting points, and pathways to scale; and federal support for state and local action is increasing. But they were also clear: We urgently need better measures of what matters to local communities; there are much more robust ways to assess what is reflected across the whole range of knowledge and skills in a state's standards and local educational goals; and assessment must be made useful to the student as a learner, the teacher as an instructor, and the system as an audit of progress and equity. One powerful solution is performance assessments tightly connected to curriculum.



Appendix A: Participants

Attendees

Allison Armour-Garb, Senior Advisor, Policy Research & Development, New York State Education Department

Aneesha Badrinarayan, Director of State Performance Assessment Initiatives, Learning Policy Institute

Tyler Belanga, Education Officer, Assessment and Accountability, Hawaii Department of Education

David Cook, Director of Innovative Learning, Kentucky Department of Education

Carla Evans, Senior Associate, National Center for the Improvement of Educational Assessment

Loretta Goodwin, Chief Operating Officer, Aurora Institute

Virgel Hammonds, Chief Executive Officer, Aurora Institute

Bob Lenz, Chief Executive Officer, PBLWorks

Alprentice McCutchen, Social Studies Teacher, New Rochelle High School, New Rochelle, NY

Gretchen Morgan, Managing Partner, Center for Innovation in Education

Alcine Mumby, Vice President of Program / Director, Deeper Learning Leadership Forum, Envision Learning Partners

Ray Pecheone, Co-Executive Director, Stanford School Redesign Network, Stanford University

Anna Saavedra, Co-Director, Center for Applied Research in Education, Dornsife College of Letters, Arts, and Sciences, University of Southern California

Kathleen Schwille, Executive Director, Teaching and Assessment, PBLWorks

Lauren Stoll, Research Associate, SCALE Science at WestEd

Ruth Chung Wei, Director of Research and Innovation, Envision Learning Partners

Jill Wertheim, Director, SCALE Science at WestEd

Support

PBLWorks staff:

Ari Dolid, Director of Curriculum Design

Telannia Norfar, Director of Math Curriculum Design

Cris Waldfoegel, Executive Director of Professional Services

Kristi Wagner, Director of Research and Evidence

Christopher Shearer, Founder & Principal, Third Sector Strategy LLC

End Notes

¹ <https://deeper-learning.org/>

² <https://www.nextgenlearning.org/portrait-of-a-graduate-in-practice#what-is>

³ <https://www.pblworks.org/what-is-pbl>

⁴ <https://learningpolicyinstitute.org/product/cpac-performance-assessments-support-student-learning-brief>

⁵ <https://www.pblworks.org/research/success-skills-rubrics>

⁶ <https://www.hawaiipublicschools.org/TeachingAndLearning/Testing/PADI/Pages/default.aspx>

⁷ <https://www.education.ky.gov/school/innov/Pages/Current-Projects-and-Initiatives.aspx>

⁸ <https://www.nysed.gov/plan-pilot>

⁹ <https://www.gettingsmart.com/2019/09/09/rethinking-scaling/>

¹⁰ <https://www.nciea.org/blog/moving-education-from-accountability-to-shared-responsibility/>

¹¹ <https://www2.ed.gov/policy/gen/guid/secletter/231120.html>