



A BRIDGE TO THE FUTURE OF STATE ASSESSMENT

April 2026

Dr. David Nitkin, Dr. Deborah Gist, Leslie Colwell

Table of Contents

Section 1: Why Reimagine State Assessment? Why Now?	3
The Unique Opportunity Created by Artificial Intelligence	6
Why State Leaders Can—and Must—Lead	7
How This Paper is Organized	8
Section 2: A Long-Term Vision for State Assessment	9
2.1. Continue to Serve Essential Functions—and Set the Guardrails	10
2.2. Integrate Instruction, Assessment, and Real-World Application	12
2.3. Assess a Broader and More Rigorous Set of Outcomes	15
2.4. Encompass New Governance Models	17
Section 3: What Can State Policy Leaders Do Now to Move Towards that Long-Term Vision?	20
3.1 Preserve and Strengthen What Works	21
3.2 Avoid Over-Promising What the Current System Can Deliver	23
3.3 Build R&D Capacity to Create What Comes Next	25
3.4 Create Space for Innovation	29
3.5 Protect Public Trust at Every Step	32
The Leadership We Need	34
Acknowledgments	35

About Transcend

Transcend is a national nonprofit that supports communities to create and spread extraordinary learning for all. We do this by partnering closely with schools and systems, sharing tools and resources, and cultivating a surrounding ecosystem conducive to change. Scan the QR code to learn more.





Section 1:

Why Reimagine State Assessment? Why Now?

Over the past 30 years, the U.S. has built one of the most comprehensive systems of educational assessment and accountability in the world. This system brought long-overdue attention to student achievement gaps and sought to hold states, districts, and schools responsible for the success of all students. From the mid-1990s through the early 2010s, American students experienced [improved educational achievement and attainment, particularly among historically underserved groups](#). These are meaningful accomplishments that deserve recognition.

The goals of this system are timeless and essential: illuminating student achievement and generating evidence to support improvement. Yet the available technology for implementation has always had significant limitations. To make learning visible at scale, standardized tests compress complex human development into limited content areas, narrow formats, and standardized pacing. The narrowness is a feature, not a flaw; it was a practical solution to the difficult problem of measuring learning consistently and efficiently across diverse schools and students. But while the demands of the world and workforce have changed, K-12 learning models and the systems designed to assess them have remained stubbornly consistent.

In our work at [Transcend](#), we have helped hundreds of schools and districts design and implement new learning models that maximize achievement and improve student experiences. We have also surveyed hundreds more through our co-stewardship of the [Canopy Project](#). Through this work, we have observed how current assessment systems both support and constrain educational improvement, and particularly how they inhibit new models designed to create the kinds of customized, relevant, and rigorous learning experiences that [research shows](#) are necessary for learning.

A changing world requires new models for learning that improve upon the existing approach used by most schools, and these models will require new assessment approaches that are coherent with their goals and design. This creates an essential question: what vision for assessment should guide us in a world where we aim to transform the student experience while also preserving the core goals—rigor, transparency, and fairness—of our current assessment systems?

This paper is designed to help state policymakers answer that question.

One response might be that the current assessment system works, we just need to double down on implementing it effectively. But the answer cannot be to recommit with increased ferocity to an early 2000s playbook which, while associated with some meaningful gains, also assumes and reinforces one-size-fits-all models that [learning science is showing](#) to be increasingly misaligned with what students need, including the need to accelerate reading and math learning.

The Hidden Logistical Costs of Standardized Testing

While debates on the burden of statewide assessment typically focus on the instructional time lost when students take exams, that view misses the full logistical costs these tests impose on adults. Standardized testing requires months of preparation, technology coordination, scheduling, compliance documentation, training, and related planning. In practice, this pulls time, attention, and human resources away from daily teaching and learning.



To give one illustrative example from the Canopy Project’s most [recent research report](#), one school leader describes how “for the two-month window of testing, every single special educator, 75 percent of their job is just test administration, test logistics, data entry. All the [testing] accommodations have to be listed in the state platform, so now I’m pulling my valuable special education teachers out of working with kids to input [data] for a minimum of 30 minutes per kid. ... A lot of our [students with disabilities] are doing one-on-one testing, so when teachers are doing a four-hour test with them, that means the other twelve kids on their caseload are getting absolutely no support during the testing window.”

This anecdote aligns with our experience as school and system leaders. We watched as talented administrators were forced to redirect their time from January to April away from coaching teachers to instead create byzantine testing schedules and complete mandatory state paperwork. We also saw multiple weeks of teacher professional learning canceled to make time for mandatory test security training and hours spent stripping all student work and instructional guides from classroom walls to make them “state-test ready.” The time an individual student spends completing an exam is only the tip of the iceberg for the indirect costs—hidden in calendars, staffing, and leader time—that are largely unaccounted for in typical policy discussions.

Even if policymakers want to double down on the statewide assessment systems of the early 2000s, the political conditions no longer exist to do so. For more than a decade, standardized testing has faced sustained opposition from teachers, families, and advocates across the political spectrum. [Opt-out movements](#), the [steady elimination of high school exit exams](#), and [repeated legislative efforts to reduce testing requirements](#) reflect a [broad erosion of public support](#). As [Andrew Rice recently argued in *New York magazine*](#), coalitions of teachers' unions and dismayed parents succeeded in dismantling regular testing regimes. This shift helped usher in today's period of stagnation, but is unlikely to be fully reversed without creative new approaches and messaging.

This creates a dilemma for state leaders. The original moral core of the accountability movement—rigor, transparency, and fairness—still matters deeply. But defending a legacy strategy built on an outdated technology is both tactically unsustainable and strategically insufficient. Principled defense must be coupled with a proactive agenda that addresses critics and attracts new allies.

What's required is innovation: recommitting to the core goals of statewide assessment in a changed world while taking advantage of advances in learning science, evolving school models, and new technologies that make new approaches feasible.

The Unique Opportunity Created by Artificial Intelligence

The idea of richer, more meaningful assessments is not new. For decades, many educators have utilized performance assessments, portfolio assessments, and other approaches designed to capture deeper learning. Examples like New Hampshire's Performance Assessment of Competency Education (PACE) and the New York Performance Standards Consortium demonstrate that such approaches can be implemented across sites while ensuring comparability. However, their complexity compared to traditional standardized assessments poses challenges to the prospect of scaling them efficiently and reliably across millions of classrooms.

Artificial intelligence (AI) offers the prospect of unlocking this bottleneck. New AI models may one day be able to assess new content in new ways and at much lower costs. The ability of AI to make inferences about student learning based on passive observation of learning process and products such as video, audio, and written work products may also mean that the era of "sitting down to take a test" may be ending, replaced with a new one

in which assessment is interwoven with instruction and real-world application in ways that are seamless and unobservable.

These capacities are not, in themselves, ready to drop into statewide policy tomorrow. They remain at the stage of possibility and piloting, not yet proven tools, and moving from promise to practice will require commitment to technical infrastructure, policy support, iterative testing, and careful governance to establish, apply, and validate AI-powered approaches.

But just as the Wright brothers' first flight at Kitty Hawk signaled an exciting new future without immediately ushering in the golden age of air travel, new AI tools may open the door to a new golden age of assessment that was inconceivable only a few years ago. For the first time in decades, we can see a realistic path toward these possibilities emerging within a timeframe we can plan around. Realizing that potential will require deliberate, sustained research and development (R&D), paired with thoughtful policy design. States are the logical engines for that work.



Why State Leaders Can—and Must—Lead

State policymakers are essential to this work because assessment systems are built—and constrained—by state policy. States determine what assessments are required, how growth is calculated, which indicators matter, how data are reported, and what consequences follow. Those decisions shape district behavior far more than any single innovation at the school level. Local leaders can pilot new models, but if those models run counter to the state's assessment architecture, they rarely scale. Only states can redesign the rules of the game, redefining what counts, what is visible, and what earns attention.

Meaningful shifts in statewide assessment also require federal flexibility, and states are the only actors positioned to seek waivers under Every Student Succeeds Act (ESSA) or to apply for pilots through the Innovative Assessment Demonstration Authority (IADA). The current administration's stated emphasis on "returning education to the states" places renewed responsibility on states as the primary drivers of America's strategy for educational improvement (some may argue that they always have been). With the federal government [expressing openness](#) to considering waivers from ESSA testing provisions,

states also play a key role in establishing guardrails that ensure that any changes to existing systems retain and strengthen core principles of rigor, transparency, and fairness.

What states cannot do alone is finance the full scale of the R&D and infrastructure required to modernize assessment. State budgets are constrained, and most agencies are already operating at capacity. The larger share of investment will need to come from the federal government, business coalitions that see workforce alignment at stake, and philanthropy willing to fund early-stage experimentation and capacity building. But states play a catalytic role. They can articulate a coherent vision, define guardrails, set research agendas, and signal seriousness through policy action, as we'll discuss in-depth in Section 3.3. By creating clarity and momentum, governors and chiefs can build the political groundswell that unlocks outside capital. If state leaders want outside investment, they need to articulate a vision worth investing in.

How This Paper is Organized

The remainder of this paper is structured to support state leaders in navigating this transition with both vision and practical guidance. In Section 2, we begin by articulating an aspirational framework for 21st-century assessment—a North Star grounded in enduring goals but unconstrained by legacy technology. We then turn in Section 3 to what's possible today and what states can begin implementing now, with a focus on maintaining essential guardrails while accelerating innovation to transformative new approaches.

The aim is to help leaders move forward without sacrificing rigor, transparency, or fairness, building a bridge to the future that preserves what matters and releases what no longer serves.





Section 2:

A Long-Term Vision for State Assessment

Every lasting change begins with a vision of what could be. By anchoring in a long-term aspiration while taking deliberate steps toward it, leaders can ensure that near-term decisions build toward a more durable and effective system.

At the same time, transforming state assessment is not an overnight project. The work is complex, the stakes are high, and many of the most promising innovations are still emerging. Moving too quickly to replace existing systems would risk eroding public trust and destabilizing schools, even if the long-term vision is sound. Moving too slowly risks ceding leadership to critics or to actors who would apply new technologies without appropriate safeguards. What is required instead is disciplined progress in service of a clear and compelling long-term direction.

Below, we offer four guiding principles for a long-term vision for state-level assessment. We believe these systems should:

2.1. Continue to Serve Essential Functions—and Set the Guardrails

State assessment systems exist to serve specific public purposes. At their core, statewide systems are designed to:

- Provide transparency and build public trust in how public education dollars are spent
- Signal learning expectations so schools do not quietly lower standards, especially for [historically underserved students](#)
- Monitor performance across schools, districts, and student groups, with a particular focus on the most vulnerable groups
- Support improvement by providing consistent, clear information to evaluate where interventions are working and where they are not

Traditional assessment systems have a mixed track record in accomplishing these tasks. For example, they are much better at illuminating disparities than helping us solve them. But any vision for the future must not sacrifice—and indeed, must improve—our ability to accomplish these four essential tasks.

Another way of framing these essential functions is through the lens of the groups they serve. As we reimagine statewide assessment, these systems must:

- Provide the **public** with transparent information to build trust and engagement
- Signal to **educators and students** a high bar for rich, rigorous learning expectations
- Enable **policymakers** to monitor and improve statewide educational growth and achievement
- Give information to **families** to help them make decisions for their children. This includes:
 - Performance of schools so they can make choices about where to enroll their child

- Performance of their child in particular, so they have an **independent check on information provided by their school.**

Just as important is being explicit about what roles state assessments should **NOT** serve:

- Providing actionable instructional guidance for teachers
- Functioning as real-time diagnostic tools for classroom decision-making

These two roles are important, but state assessments are poorly suited to serve them compared to the full marketplace of formative and diagnostic assessments, including assessments built directly into high-quality instructional materials (see Section 3.2).

As states consider new approaches, certain guardrails must remain firmly in place to avoid repeating past mistakes or drifting into incoherence. In particular, any future assessment system should publicly report disaggregated data by student group, set minimum technical quality standards for any measure used in public. Accountability, require bias audits and differential impact analyses, preserve comparability across student groups, even when formats for demonstrating learning vary, protect student data and privacy through a clear assessment governance framework, and

Assessment Reform Is Political Work

Education policy is sometimes discussed as though it were simply a matter of getting the technical details right. Build a better assessment. Improve the statistical model. Deploy a new technology. If we do those things well enough, the thinking goes, political resistance will melt away.

However, anyone who has spent time in public service knows that policy does not move forward on technical merit alone. Education is deeply personal to families, educators, and communities. It is also a major public enterprise involving billions of dollars and millions of lives. Results shape reputations, influence funding decisions, and affect how the public judges schools and elected officials. Unsurprisingly, the groups affected by those signals often have strong views about how assessment should work and how results should be used.

Leaders who want to improve assessment systems should expect disagreement and criticism. But the purpose of state assessment policy is not to avoid controversy; it is to provide the public with honest information about whether children are learning. State leaders must remain committed to that transparency while driving to improve how that information is gathered and shared.

ensure new assessment approaches are only rolled out with sufficient preparation and implementation resources

Recent history offers cautionary examples of what can happen when shared guardrails around rigor, transparency, and fairness are removed before credible alternatives are in place. When the University of California, San Diego eliminated the use of standardized test scores in admissions without replacing them with comparably validated measures, the system lost a common signal for student readiness. As a result, the [share of admitted students requiring remedial math increased sharply](#), from 32 students in the fall of 2020 to 1,000 students in 2025. While these numbers may be partially influenced by pandemic-era learning disruptions, they underscore the risks of removing widely understood measures before new approaches are sufficient to serve similar public purposes.

In addition to preserving these essential functions and guardrails, we believe state assessment systems must evolve in three additional ways, which we describe below.

2.2. Integrate Instruction, Assessment, and Real-World Application

Some day we will shake our heads in bewilderment that it ever felt normal to shut down learning for weeks so students could sit silently in a room filling in bubbles. The problem is not just that this experience is unpleasant—although rare is the educator who was called to the craft of teaching out of fervent desire to proctor exams—but that it encodes an outdated vision of what school is supposed to look like: compliance over agency, sameness over personalization, silence over collaboration.

Standardized testing both reflects and reinforces an industrial-era model of schooling, rewarding uniform pacing, isolated work, and passive demonstration of knowledge. Over time, this creates a systems-level contradiction: we ask schools to personalize learning, foster creativity, and prepare students for complex, real-world problem-solving, while measuring their learning through processes that assume the opposite. When the form of assessment is misaligned with the vision for learning, the

...we ask schools to personalize learning, foster creativity, and prepare students for complex, real-world problem-solving, while measuring their learning through processes that assume the opposite.

results may be technically defensible but feel increasingly illegitimate. Public trust erodes accordingly.

A future-ready assessment system requires evidence that emerges from instruction and real-world application over time, not isolated testing events. Many established assessment systems already rely on this type of approach. The International Baccalaureate uses [performance-based assessments](#) to evaluate complex reasoning and application, Montessori education relies on [structured observation](#) to understand student development, and Big Picture Learning schools center [portfolio presentations and exhibitions of learning](#). Many schools utilize competency-based approaches in which students advance when they demonstrate mastery rather than according to an arbitrary calendar. What limits the spread of these approaches today is not their credibility, but their cost and complexity. They require resources that most systems struggle to provide at scale, including intense educator training, calibration across classrooms, and significant time devoted to reviewing and moderating student work.

This is where technological innovation begins to matter—not because AI replaces professional judgment, but because it can lower the cost of doing things many educators already know how to do well. Observation, portfolio review, and performance assessment are powerful ways to understand learning, but they are difficult to sustain precisely because they demand time, coordination, and trained expertise. Over time, AI could support these practices by helping to surface patterns and organize evidence across student work. The opportunity is to make proven approaches to assessment workable at a scale that has historically been out of reach.

In practice, this type of assessment can take at least three complementary forms:

Formal curation:

Students assemble evidence of learning over time through portfolios that culminate in capstones or exhibitions.

Consider a science performance task in which students design an experiment, collect and analyze data, revise their methods in response to unexpected results, and present

their findings to peers. Appropriately trained AI models could support the assessment of portfolios that include lab notes, data visualizations, written analysis, and a final presentation by helping educators track evidence of students demonstrating content knowledge, scientific reasoning, and grade-level-appropriate writing skills.



Passive observation:

Evidence of learning is gathered continuously and unobtrusively as students engage in authentic work. Consider a structured Socratic seminar: students discussing a shared text, listening to one another, building on peers' ideas, and grounding claims in evidence. Traditionally, capturing the student demonstration of those skills requires an adult observer to take notes, score against a rubric, and later reconcile judgments with colleagues. Emerging technologies raise the possibility that some of this work could be supported in the background by tracking participation patterns, identifying moments where students build on one another's thinking, or even noting growth in each student's engagement across months of discussions. As we discuss in [section 3.5](#), proactively addressing privacy and safety concerns is especially critical for this type of assessment.



Paired approaches that combine the two:

Curated portfolios are paired with on-demand demonstrations or defenses. This mirrors existing systems such as AP Studio Art, AP Seminar, or AP Research, where curated portfolios are complemented by live presentation, with both evaluated against shared criteria. The live performance element provides a check to ensure that the work students submit in their portfolios is truly their own.



None of these approaches is close to ready for use in state summative assessment. As with many emerging technologies, their early testing should occur in low-stakes contexts, such as classroom- or school-level formative assessment, before it eventually “ladders up” to higher-stakes uses. The absence of a clear, fully articulated vision for how these approaches could function as state summative assessments is therefore not a failure of imagination, but a predictable consequence of being early in the learning cycle. States should expect to learn their way toward that vision through disciplined experimentation over time.

Designing Learning for the Age of AI

Assessment is one of many areas where AI can be transformative. Transcend recently published a new report that outlines how AI can catalyze the learning young people need. [Read more here.](#)



What we can say with confidence is that applying these kinds of assessments for statewide summative purposes will require both new technologies and new state capacity. States may need dedicated teams responsible for calibration, moderation, and validity studies, as well as updated conceptions of comparability that remain psychometrically defensible while moving beyond the [narrow forms currently emphasized in federal peer review](#). It may also require moving from “content agnostic” assessment to a state vision that more intentionally intertwines state summative assessment with statewide content standards or content-rich curricula, similar to the [approach Louisiana has been testing](#) with its ELA assessments.

The application of AI offers the prospect of educators and policymakers having access to more evidence of learning in more accessible formats, with less time spent testing. And perhaps one day, a world in which students prove their skills by doing meaningful work rather than filling in bubbles.

2.3. Assess a Broader and More Rigorous Set of Outcomes

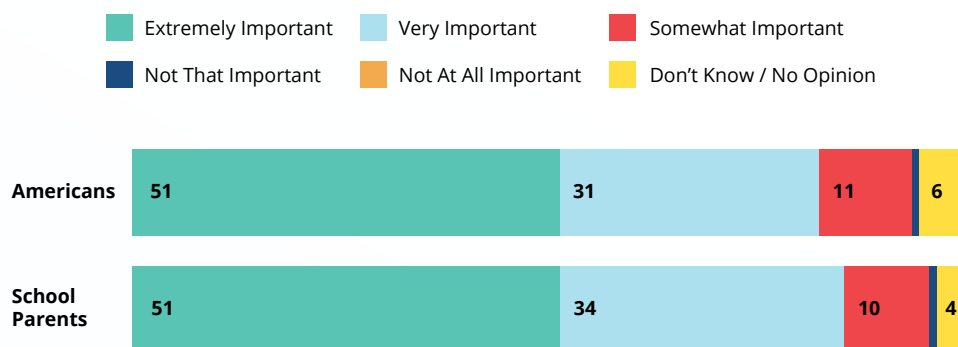
Once assessment is integrated with instruction and real-world application, the question of what we can assess also expands. Skills like critical thinking, communication, and collaboration have always been essential, but have been difficult to measure for structural reasons; a comprehensive assessment of most 21st Century skills requires [diverse evidence collected over time](#), and complex skills do not exist in the abstract. [A student cannot demonstrate critical thinking skills without meaningful content to think critically about.](#)

This is one of the central advantages of integrating assessment with learning and real-world application. Rather than forcing states to choose between measuring core academics or broader competencies, it allows both to be observed simultaneously from the same activities. A student’s written analysis can demonstrate both reading comprehension and complex reasoning. A math project can reveal procedural fluency alongside creativity. A science investigation and lab report can surface content knowledge, collaboration, and communication. In this way, broadening what is assessed strengthens, rather than dilutes, the rigor of core academic skills.

Public demand for this type of expansion is clear. [When asked](#), “How important is it that your child’s/children’s school prioritizes developing students’ durable skills (such as communication, critical thinking, teamwork, adaptability, and responsibility)?” 85 percent of school parents said these skills were very important (34 percent) or extremely important

(51 percent). This is true among many different types of families. For example, the [National Urban League and Unidos recently did listening sessions](#) with young people, their families and communities, educators, and youth development workers from historically excluded groups. Among their findings was a prioritization of “success beyond academics to broader competencies, attitudes, and aptitudes that are social, emotional, and cognitive.” If parents and students believe in the importance of these skills, then assessment leaders should aspire to measure them in ways that are valid and trustworthy.

Figure 1. How important is it that your child's/children's school prioritizes developing students' durable skills (such as communication, critical thinking, teamwork, adaptability, and responsibility)?



Source: America Succeeds

Integrated assessment also makes it possible to measure something that traditional systems have largely ignored: the quality of students’ learning experiences. When assessment is integrated with learning and real-world application rather than confined to annual testing events, systems can gather evidence on factors like [engagement, intellectual challenge, and belonging](#) that are both predictors of achievement and valuable in their own right. An assessment system that renders those experiences invisible sends an implicit signal that they do not matter; making them visible aligns statewide assessment with a more humane and more honest vision of what constitutes a high-quality education.

It is also important to acknowledge the real challenges in translating this vision into practice. There is not yet full agreement on precise definitions, progressions, or standards for many complex skills, and states and organizations continue to use different language and frameworks. But across these efforts, meaningful convergence is emerging. [State graduate profiles](#), the [durable skills framework](#) developed by America Succeeds, and skills research from [the Carnegie Foundation and ETS](#) all point to a shared core: critical thinking, communication, collaboration, creativity, and similar skills. The task ahead is to build on this growing alignment by convening states, educators, and measurement experts around shared expectations that are broadly understood, publicly legitimate, and that we aspire to one day include in state summative assessments.

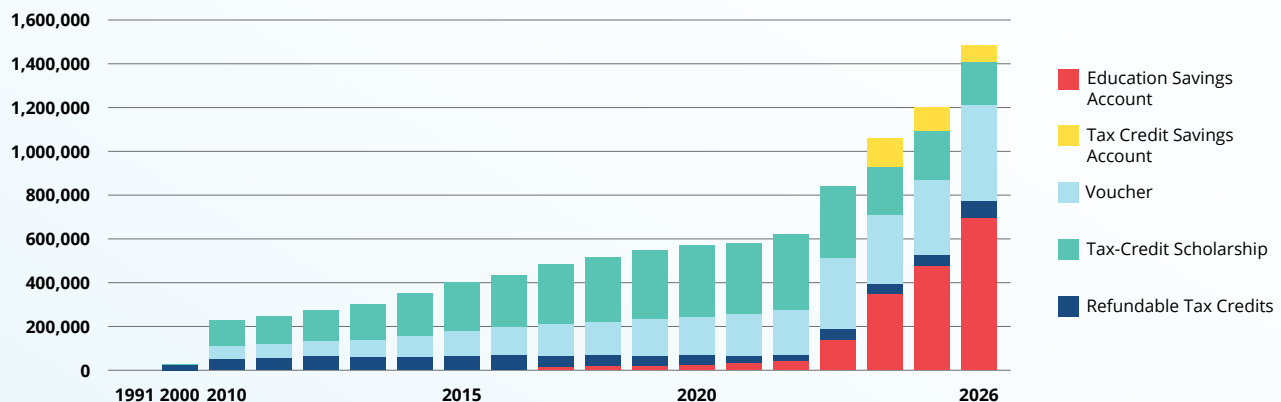
2.4. Encompass New Governance Models

Taken together, the shifts described above open up new possibilities for how states can ensure that all public education dollars lead to real learning. Historically, that has meant encompassing all public district or charter schools, with the assumption that each child is assigned to only one school at a time. In the future, however, this assumption will likely no longer be a good fit for the reality of American K-12 education. We increasingly see evidence that it is no longer even a good fit for today.

There have always been gaps in this model. Private schools and homeschools have historically fallen outside of public assessment systems. Even within public schooling, many innovative programs have utilized cross-campus or flexible models that do not neatly match the “one school-of-record” assumptions of assessment systems. For example, career and technical education (CTE) models like [CART](#) or the [Brooklyn STEAM Center](#) enroll students from multiple “schools of record” and in some cases from multiple districts. Similarly, the growing hybrid homeschool movement sees students enrolling in public districts only part-time, with those districts receiving per-pupil funding proportional to the amount of services the family opts into.

These gaps are likely to accelerate, given the expansion of [state education savings account \(ESA\) programs](#) and [new federal tax credits](#). Nearly 1.5 million students utilize some form of public funding to pay for private educational options. Nineteen states have at least one private school choice program with universal eligibility, meaning they are open to all K-12 students. [Approximately 27 million students](#), or about half of the K-12 school-aged population, are now eligible for some form of publicly funded private school choice program.

Figure 2. Number of Students Participating in ESA, Tax-Credit ESA, Voucher, Tax-Credit Scholarship, and Refundable Tax Credit Programs



Source: EdChoice

These state-level programs will soon be supplemented by a stream of federal funding unlocked by a new federal scholarship tax credit. [This program](#) will allow any taxpayer to claim a tax credit of up to \$1,700 for cash contributions to scholarship-granting organizations (SGOs) that provide money for students' educational expenses. [Georgetown University's Edunomics Lab](#) has estimated that this could unlock \$28 billion per year in new funding, potentially a massive accelerant for student enrollment in homeschooled, private schools, microschools, and other out-of-system options like “unbundled” educational models. Planning a response to these new policies does not require agreement on the merits of ESAs or school choice—only recognition that student learning now occurs across more varied environments than our current systems were built to monitor.

Traditional assessment systems, which assume that all students can be comprehensively linked to a single public school, are increasingly out of step with this reality. Given existing legislation and funding, it is plausible to imagine students evenly spread across traditional public district or charter schools, private schools, or homeschooled that they pay for with public funds. In this paradigm, which Bellwether has described as “[Assembly](#)” education and LearnerStudio and Getting Smart have described as “[Horizon 3](#)” learning, families stitch together their ideal educational experience from a marketplace of public and private offerings.

We believe that it's essential for policymakers, families, and the public to have access to information about student learning across these diverse types of learning environments. Generating this type of information will require policymakers to balance the importance of enabling local innovation with the necessity of accounting for whether public dollars are being spent on programs that actually help young people learn. This will almost certainly require a more nuanced assessment approach than simply expanding traditional standardized testing to new school sites.

What could this look like in practice? Following are three illustrative approaches that show how states could modernize assessment to encompass new governance models without defaulting to one-size-fits-all standardized testing.

States can establish a technical backbone that makes learning legible across settings without requiring uniform experiences.

Rather than mandating a single assessment or format, states can define what counts as valid evidence of learning, how different forms of evidence are evaluated, and how results are reported on common scales. Performance tasks, portfolios, and exams can be linked through shared standards and statistical methods that preserve comparability

and trend monitoring (as referenced earlier, this will require a different understanding of “comparability” than what is embodied in the [current federal peer review process](#)). In this model, rigor comes not from forcing sameness but from clear expectations about validity, bias, and transparency—allowing innovation while maintaining coherence at the center.

States can empower a diverse set of agencies to gather validated evidence of learning.

This can include not only existing schools and districts, but also regional consortia or independent third parties—such as universities or nonprofits—with the capacity to design and administer assessments, similar to how many states license and monitor charter school authorizers. Students would demonstrate mastery when ready—whether learning occurred in a district classroom, a CTE center, a tutoring program, or an out-of-system model—and earn portable, state-recognized credentials recorded in a learner educational record. What matters is not where learning happened or who administered the assessment, but whether the pathway adds up to demonstrated mastery under publicly credible rules.

States do not need to impose this system all at once.

A voluntary-first, incentive-based approach offers a more realistic path forward. Participation in new assessment and credentialing systems can begin as optional but tied to concrete benefits such as course credit, credentials with labor-market value, eligibility for funding streams, or access to advanced learning opportunities. This allows technical challenges to be resolved, standards to stabilize, and public trust to be built before stakes rise. Over time, as participation grows and legitimacy solidifies, states can decide where stronger requirements are warranted—creating momentum for innovation without triggering the backlash that often accompanies rapid, top-down change.

In a future in which “unbundled” learning models are increasingly common, modernizing statewide assessment systems may mean moving toward a model in which assessment is tied directly to the learner, not the institution. This type of system is far from ready for implementation today. But now is the time for state leaders and their allies to begin sketching their outlines and launching low-stakes pilots for learning.



Section 3:

What Can State Policy Leaders Do Now to Move Toward that Long-Term Vision?

The preceding section outlines a long-term vision for statewide assessment. The question now is how to begin moving toward that future. What practical steps can state-level policy actors take today to lay the groundwork?

We are entering a **bridge period**—one where state leaders must preserve the accountability and transparency that families and the public rightly expect and need, while simultaneously creating space to design what comes next. We believe the next decade of assessment reform will be defined by whether state leaders can successfully build this bridge—time-bound, policy-sanctioned spaces where new models can mature before carrying full stakes.

The challenge is balance. States cannot afford a free-for-all in which schools operate with no oversight, nor can they afford to lock themselves into rigid structures that discourage experimentation. The task is to chart a middle course—one that protects the best of what works while empowering schools and districts to try new models and approaches that point toward the future.

To support policymakers ready to lead through this moment, we outline five practical next steps designed to safeguard what matters now and make room for what's next.

- 1. Preserve and Strengthen What Works**
- 2. Avoid Over-Promising What the Current System Can Deliver**
- 3. Build R&D Capacity to Create What Comes Next**
- 4. Create Space for Innovation**
- 5. Protect Public Trust at Every Step**



3.1 Preserve and Strengthen What Works

A bridge to the future must rest on a stable foundation. Even as states plan for new approaches to assessment, there are core elements of the current system that deliver real value and should be preserved—and, in some cases, strengthened—during the transition period.

3.1.A. Maintain statewide, comparable assessment and public reporting as non-negotiable guardrails.

Statewide assessments that generate comparable data across schools, districts, and student groups, and are publicly reported in disaggregated form, remain essential to protecting students furthest from opportunity and public trust. These guardrails are protections against quiet backsliding, especially for historically underserved students.

States should be wary of prematurely moving away from statewide, comparable testing before credible replacements exist. Recent efforts in [some states](#) to abandon statewide assessments without a sufficiently rigorous alternative offer a [cautionary lesson](#): pulling away from comparability too quickly risks creating opacity rather than enabling innovation, weakening public trust at precisely the moment it is most needed.

3.1.B. Align cut scores to an honest, high bar for rigor—ideally anchored to NAEP.

Preserving annual testing is not enough; states must also confront whether their definitions of “proficient” are truly rigorous. Many states set proficiency cut scores that fall [well below](#) what national [benchmarks, such as NAEP](#), define as meaningful readiness. [Recent work in Virginia](#) to revise its standards illustrates that states can recalibrate expectations transparently and responsibly, pairing higher standards with clearer communication rather than using low cut scores to mask underperformance.

3.1.C. Release assessment data faster so it can actually inform decisions.

Assessment data loses much of its value when it arrives too late to matter. In many states, families do not receive spring test results until well into the next school year, long after enrollment decisions have been made. Faster reporting is both [feasible and necessary](#). In a world of expanding school choice and cross-district enrollment, states should aim to release preliminary results within days of test completion and full results within weeks so families can use them when deciding where and how their children will learn. States can achieve speed without sacrificing accuracy by modernizing data systems and prioritizing timeliness as a core design goal rather than a secondary convenience.

3.1.D. Preserve rigor while offering students multiple opportunities to demonstrate mastery.

High standards are most defensible when paired with multiple chances to meet them. [Much has been written](#) about the “southern surge,” in which states like Mississippi, Louisiana, Alabama, and Tennessee paired high-quality instructional materials, professional learning, and high-stakes testing to drive impressive improvements in early literacy. An often overlooked element of these states’ approaches is that they offer students [up to three opportunities](#) to take the 3rd-grade state test and demonstrate

their readiness to move forward. In [Mississippi](#), a total of 32,839 third graders took the assessment in 2023-24. 76 percent of third graders passed the initial administration of the reading assessment, but after the final retest that number climbed to 84 percent. Offering students multiple opportunities to demonstrate mastery is an established and manageable way for states to boost achievement and begin moving towards more customized systems.

3.1.E. Recommit to using assessment data to support improvement—not just compliance.

Finally, preserving what works requires states to take seriously the “accountability” side of assessment and accountability. [Federal reviews](#) have consistently found that while states collect vast amounts of data, they often fail to use that information to meaningfully intervene with districts and schools that need improvement. Recommitting to data use means going beyond compliance monitoring or lackluster technical assistance. The details of how states redesign accountability systems to drive improvement is a critical question that extends beyond the scope of this paper. What matters is the principle: assessment without meaningful follow-through erodes credibility, while assessment paired with intelligent support strengthens both trust and results.

3.2 Avoid Over-Promising What the Current System Can Deliver

As states navigate this bridge period, one of the most important—and often overlooked—tasks is to be honest about what current statewide assessments are and are not designed to do. As covered in Section 2.1, state assessments play a vital role when they are used for the purposes they were built to serve: enabling policymakers to monitor statewide performance and growth; providing the public with transparent information that builds trust; signaling a high bar for rigorous learning expectations; and giving families independent information to help them make decisions about their children’s education.

These are consequential functions, and they matter precisely because they operate at the system level. But they are also bounded. Statewide assessments are not designed to provide actionable instructional guidance for teachers, nor to function as real-time diagnostic tools for classroom decision-making. Expecting them to do so misunderstands both their purpose and their design.

This distinction is [widely understood among assessment experts](#), but it is often discussed sloppily in public debate. [Reputable organizations](#) regularly report that teachers do not find state tests useful for day-to-day instruction—treating this as a failure of the assessments themselves rather than a misunderstanding of why state tests exist in the first place. These critiques are akin to criticizing an annual financial audit because it does not help a manager decide how to allocate next week’s travel budget. Audits exist to ensure accuracy, comparability, and accountability at the system level; daily expense tracking serves a different function entirely. When we fault statewide assessments for not behaving like formative classroom tools, we set them up to fail—and in the process, we undermine their legitimacy.

This confusion has led some states to pursue a well-intended, but we believe misguided response: [expanding statewide testing](#) by layering on through-year or interim assessments in the hope of making the system more [instructionally useful](#). Evidence from recent state experiences suggests this approach often backfires; [teachers have reported](#) that these systems are more time-consuming ([one analysis](#) calculated an average of 27 additional hours of testing time), quiz students on concepts they have not been taught yet, and provide results that are difficult to interpret. Meanwhile, [a survey of school leaders](#) working to design more effective learning approaches found that respondents preferred “less frequent state tests” over “interim score reports at multiple times in the year” by nearly a 3-to-1 ratio. These leaders—even those who were strongly in favor of state testing—emphasized that a lower administrative burden from mandatory state testing would increase their opportunities to utilize other forms of assessment that they find more useful for daily instructional improvement.

A more productive path for state leaders is to draw clearer boundaries—and then invest appropriately on either side of them. Rather than stretching statewide assessments beyond their intended role, states should encourage districts to conduct [assessment audits](#) that identify and eliminate redundant, low-quality, or misaligned local measures and streamline the overall assessment landscape. [Many districts](#) have accumulated layers of interim, benchmark, and diagnostic tests that overlap poorly with state standards and with one another; [a recent audit](#) conducted by the Oregon Department of Education found no fewer than 232 distinct test instruments in use across the state, and national data suggest Oregon is not an outlier.

Recent national analysis by [Education First](#) underscores why audit-and-streamline strategies matter: in some districts, students take as many as 88 academic assessments before high school. This burden is not evenly distributed: English learners experience

about 47 additional hours of testing across K–8, largely driven by ELL-specific measures—an example of how more testing can compound inequity even when intentions are good. The same analysis found no statistically significant relationship between testing volume (number of assessments or testing time) and ELA/math proficiency or growth, suggesting many systems are spending time on redundant measures that don't translate into better learning.

At the same time, states can support districts in adopting **high-quality, standards-aligned instructional materials** that include embedded formative assessments designed for classroom use. These tools—built into curriculum and specifically designed for teachers—are far better suited to informing instruction than any statewide exam could be. By pairing leaner statewide assessment systems with stronger local instructional materials, states can improve the quality of information educators receive without asking state tests to do work they were never meant to do.



Avoiding over-promising is a commitment to using the right tools for the right purposes, and to preserving the credibility of statewide assessment to serve the essential functions for which they were created. This will be essential as we bridge towards new approaches that better deliver on these fundamental goals while also offering exciting new possibilities.

3.3 Build R&D Capacity to Create What Comes Next

If states want to realize the possibilities described earlier in this paper, they will need to invest deliberately in assessment R&D that is grounded in what the system is trying to produce. Most states do not yet have a clear, public set of priority questions that link assessment innovation to the core purposes of statewide systems. A research or learning agenda makes those questions explicit and usable—for governors, chiefs, legislators, and agency teams—to guide investment, oversight, and communication. Before expanding or replacing assessment requirements, state leaders should insist on credible answers to questions such as:

- **Comparability:** How can evidence from heterogeneous tasks, formats, and scoring processes be placed on a common scale that supports valid comparison, aggregation, and trend analysis?
- **Bias and fairness:** What methods can reliably detect, quantify, and mitigate bias, including differential task functioning?
- **Human–AI balance:** What division of labor between automated scoring and human judgment maximizes reliability, transparency, and trust—while minimizing burden and unintended bias?
- **Usability and transparency:** How should results be summarized and communicated so that educators, families, and policymakers interpret them accurately and use them appropriately?

A research agenda only matters if the state has the routines and authority to act on it. Capacity is a central constraint for many state education agencies: accountability, reporting, and compliance demands leave limited bandwidth to build new R&D infrastructure, and many “research” functions are funded through short-term grants rather than stable appropriations. Precisely because capacity is limited, states can begin by establishing a small, explicit R&D function, lightly staffed yet empowered to convene partners, define clear evidence thresholds, and move insights into operational decisions.

Below are four specific ideas for types of experimentation that align with the research questions above:

1. Build and adopt a [Portrait of a Graduate framework](#) and explore assessments aligned to **Portrait of a Graduate competencies**, such as [North Carolina’s partnership with ETS](#) to test AI-powered assessments of durable skills and New York State’s [Performance-Based Learning and Assessment Networks \(PLAN\) pilot](#). These efforts test whether complex competencies like collaboration, communication, and critical thinking can be measured rigorously and transparently while remaining aligned to statewide expectations rather than becoming idiosyncratic local measures.
2. Invest in **instructional coherence** by encouraging the development of integrated assessment–instruction–curriculum approaches. Rather than treating assessment as a separate event, these models embed evidence-gathering directly into instruction. Examples include Illustrative Mathematics’ [integration of AI-powered](#)

assessments from Snorkl, and Teaching Lab Studio’s investment in comprehensive instructional models that seamlessly embed assessment tools like Podsie into coherent pedagogical models.

3. Test **modular or checkpoint-style assessments** that travel across learning environments, including unbundled or hybrid models. Initiatives such as XQ’s [math badging work](#) and Chad Aldeman’s suggestion of [phonics or multiplication checkpoints](#) could offer an alternative to single, high-burden summative exams. These approaches could allow students to demonstrate mastery when and where learning occurs, while still producing third-party, comparable signals for families and policymakers.
4. Support **multi-district assessment collaboratives** that allow local systems to experiment together while generating shared learning. Existing examples include [Colorado’s Student-Centered Accountability Program \(S-CAP\)](#), the [New York Performance Standards Consortium](#), the [Massachusetts Consortium for Innovative Education Assessment \(MCIEA\)](#), and [Kentucky’s Local Laboratories of Learning \(L3s\)](#). These collaboratives can offer a way to test new approaches under state oversight, align evidence standards, and reduce duplication—helping states learn faster without fragmenting accountability.

Kentucky offers a useful model for assessment R&D that stays anchored to statewide coherence while making space for local innovation. Through the Kentucky United We Learn (KUWL) effort and its L3s, the state is testing new ways to assess student outcomes and local accountability approaches and intentionally generating learning that can inform a future statewide model. Importantly, Kentucky coupled this design work with federal assessment R&D capacity via a Competitive Grants for State Assessments (CGSA) award to align local and state measures and build a roadmap for policy, technical quality, and scaling. This is a posture worth learning from: create state-level evidence standards, run learning through real district work, and use what is learned to shape the next version of the statewide system rather than letting innovation fragment into incoherent local measures.

States should also collaborate, because no single state is likely to have the resources to develop, validate, and continuously improve next-generation assessments across a wide range of competencies while meeting high standards for cultural responsiveness, bias mitigation, transparency, and comparability. A multi-state R&D consortium can pool investment and share core technical infrastructure (e.g., shared task banks, scoring

rubrics, comparability designs, bias audits, and reporting prototypes), reducing duplication and accelerating learning across pilots. Philanthropy and the federal government can catalyze this shared endeavor through matched funding and common governance so states can opt in, co-invest, and co-develop—building higher-quality “public goods” that any participating state can adapt without starting from scratch. Intermediary organizations can help synthesize findings and translate technical insights into policy-relevant guidance.

Even as states use multi-state partnerships as a force multiplier, they must treat R&D as core public infrastructure, not a series of one-off pilots. Serious R&D requires serious, sustained investment: dedicated staffing; a modern longitudinal data system that can link K–12 learning, postsecondary transitions, and workforce outcomes; shared public rubrics and technical criteria for evaluating performance evidence; and statewide professional learning to build assessment literacy and scoring reliability. Philanthropic and federal support can help, but it should supplement, not substitute for, state commitment. When states invest their own resources, they reinforce accountability to the public and ensure that R&D priorities remain aligned to state needs. Over time, this positions states not as passive adopters of externally developed solutions, but as active stewards of an assessment system that can evolve responsibly—strengthening legitimacy while preserving the essential purposes those systems exist to serve.

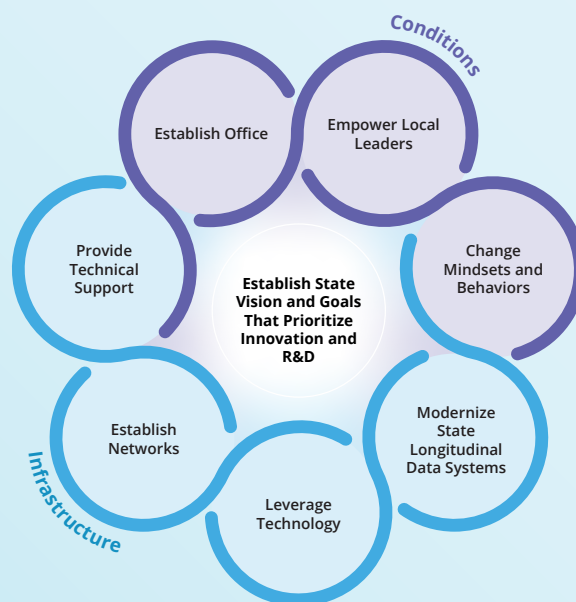
New York State Portrait of a Graduate

In July 2025, the NY Board of Regents adopted the NYS Portrait of a Graduate and required that all students demonstrate mastery of its six competencies to graduate high school. This will require new forms of assessment.



State R&D Playbook

Transcend, working in partnership with the Alliance for Learning Innovation (ALI) and Education Reimagined, recently published a [state-level R&D brief](#) and [Playbook](#) that outlines how state leaders can build the type of infrastructure described above. Informed by 100+ stakeholders across 20+ states, the Playbook is an interactive tool offering concrete strategies, real-world examples, and practical tools to help leaders identify, test, and spread what works for learners.



3.4 Create Space for Innovation

Investing in R&D is necessary, but it is not sufficient. If states want new assessment approaches to flourish, they must create room for them to take root. Too often, innovative models are asked to prove themselves while still being constrained by rules designed for a very different system. During this bridge period, state leaders should look for targeted, time-bound ways to remove constraints while preserving essential guardrails.

In some cases, this may include states seeking flexibility from federal requirements via federal waivers under Section 8401 of ESSA or applications to the IADA. In other cases, it may simply require rethinking long-standing state rules that inadvertently lock schools into one-size-fits-all assessment practices. The goal is to hold high expectations while giving promising approaches enough room to demonstrate whether they can meet or exceed a high bar for rigor, transparency, and fairness.

3.4.A. Pilot alternate-year testing to free capacity for new approaches.

Reducing the footprint of existing statewide standardized tests would create space, instructional time, and administrative capacity that states could redirect toward newer and potentially more instructionally useful assessment models. Moving to alternate-year

testing would involve real tradeoffs. But every policy choice reflects tradeoffs, including sticking with the status quo. The relevant question is not whether drawbacks exist, but whether the balance of drawbacks and benefits justifies disciplined experimentation.

The concerns are serious. Testing less frequently could lower the effective “minimum n” for some student groups, increasing the risk that performance gaps are obscured or that students are not identified for timely support. Parents whose children are performing below grade level

could lose access to timely data that empowers them to demand the support their children need to be successful. Less frequent testing could also reduce the precision of longitudinal growth estimates, which serve as an important counterweight to status measures that are closely correlated with students’ economic background. These risks may be particularly salient in schools serving large numbers of highly mobile students.

Still, emerging evidence suggests these drawbacks may be manageable. The Center for Assessment [examined five testing-reduction scenarios and found](#) that both alternate-year testing and half-length assessments could generate school support designations similar to those produced under current systems, particularly when paired with mitigation strategies such as adjusted minimum-n thresholds. [Researchers at the University of Missouri](#) similarly found that growth models based on two-year testing intervals conveyed information comparable to traditional one-year growth models, including for student subgroups. And work from [Arizona State University](#) shows year-over-year school score correlations between 0.9 and 0.95, suggesting that annual testing often yields limited new information relative to its cost.

States need not scale this approach immediately. A prudent first step would be to authorize a small set of districts to pilot alternate-year testing while rigorously tracking impacts: instructional time reclaimed, educator workload reduced, parent and student perceptions, student identification rates, and the stability of growth estimates. A lower-risk option would be to replicate the Missouri analysis using historical state data to simulate alternate-year growth calculations before implementing any policy change. Either path would ground the debate in real evidence of benefits and drawbacks.

...every policy choice reflects tradeoffs, including sticking with the status quo. The relevant question is not whether drawbacks exist, but whether the balance of drawbacks and benefits justifies disciplined experimentation.

3.4.B. Allow testing above and below grade level, with strong safeguards against tracking.

One of the most significant constraints in current systems is the requirement that every student be tested only on grade-level standards, regardless of where they are in their learning trajectory. This design choice prioritizes comparability and uniform standards but can also obscure real growth and discourage instructional acceleration or targeted support. [Chad Aldeman suggests](#) a creative alternative approach: "...instead of the current system, where every student is tested on grade-level standards regardless of their performance level, states should be allowed to test students above and below their grade level (versions of this for above-grade-level tests [have already been approved by federal waivers](#))."

We would suggest additional guards against tracking or lowered expectations. For example, states could require that families provide active consent prior to their children being offered off-grade-level assessments. This could include requirements for plain-language explanations, translated materials, and families' right to opt back into grade-level assessment in future years. Districts could also be required to pair below-grade-level assessment with instructional supports such as small-group tutoring or additional learning time, perhaps tied to the specific funding or resourcing requirements.

3.4.C. "Bank" demonstrated mastery while preserving end-of-year accountability.

Relatedly, states could allow students to "bank" demonstrated mastery once they have met grade-level expectations. Under this approach, states would continue to report the percentage of students in each school who meet grade-level benchmarks by the end of the year—a critical function for transparency. But students who have already demonstrated mastery could move on to more advanced content without repeatedly retesting on material they have already learned.

This recommendation, as well as the one above on assessing above- and below-grade level content, would also enable competency-based models in which students advance at customized rates based on their demonstrated mastery of content. [In a recent survey of nearly 200 school leaders](#) working to design more effective learning models, nearly half reported that current state assessment and accountability systems make it harder to personalize learning. Piloting these types of policies could change that.

3.4.D. Set clear timelines for approving innovation requests.

Even when flexibility exists on paper, innovation can stall in practice due to slow or opaque approval processes. States should establish clear timelines such as a 90-day window, for approving or denying district applications for assessment pilots or flexibility. Predictable timelines signal seriousness, reduce uncertainty for districts, and prevent bureaucratic delay from becoming a de facto veto on innovation.

3.4.E. Protect innovators from sanctions during pilot periods.

Finally, states should explicitly protect participating districts from punitive consequences while pilots are underway. This requires careful selection of participants and clear expectations, but it is essential for learning. Districts will not take responsible risks if a single imperfect year can trigger sanctions or reputational harm. Temporary protection does not mean the absence of oversight. It means judging pilots on whether they generate credible evidence and lessons as part of a bounded, carefully planned strategy designed to improve long-term student learning.

3.5 Protect Public Trust at Every Step

In moments of transition, legitimacy is not a procedural detail; it's the mechanism that allows institutions to move without breaking. State education agencies operate in a high-stakes environment where assessment systems carry legal, political, and moral weight. When trust erodes, state education agencies lose both public confidence and their ability to govern effectively: legislative support weakens, implementation slows, litigation risk increases, and even well-designed reforms stall.

For that reason, any effort to modernize assessment must begin by reaffirming and operationalizing the guardrails outlined in Section 2.1, including minimum technical quality standards for any measure used in accountability; bias audits and differential impact analyses; protections for student data and privacy; and comparability across student groups and schools, even when learning is demonstrated in different ways.

During pilots and innovation periods, states should also explicitly protect longitudinal trend lines, ensuring that experimentation does not sever the public's ability to understand progress over time. Innovation that compromises transparency, even temporarily, risks undermining the very legitimacy it seeks to build.

Data privacy deserves particular attention, both in substance and in public perception. History offers a clear warning. [The collapse of inBloom](#)—a well-funded, technically sophisticated student data platform—was not driven by a failure of engineering, but by a failure of trust. Parents and advocates perceived unclear governance, opaque data use, and insufficient safeguards, and the backlash was swift and decisive. As assessment systems become more data-rich and AI-enabled, state leaders must be explicit about what data is collected, how it is used, who has access to it, and how long it is retained. Just as importantly, they must communicate these protections clearly and repeatedly. In an environment where [fear about AI is growing](#), silence or ambiguity amplifies concern.

How states talk about technology—especially AI—will matter as much as how they deploy it. Parents are unlikely to accept “AI grading their children.” States should therefore be precise and honest in framing AI as a tool that supports educators by organizing evidence, flagging patterns, or assisting with moderation, not as a replacement for human judgment. This framing must also be substantively true. If AI systems are embedded in assessment workflows, states should be prepared to explain where human review occurs, how decisions are made, and how errors are identified and corrected. AI models must also exhibit “[explainability](#)” in which educators, families, and parents have transparent access to the criteria and process by which they make inferences about student competencies.

Trust also depends on proof. New assessment approaches should not carry public stakes until they clearly demonstrate that they meet or exceed existing standards for reliability, validity, fairness, and transparency. Where evidence falls short, states should be prepared to pause, revise, or abandon approaches—even if they are innovative, well-funded, or politically appealing.

Finally, the process by which change occurs matters. States that have navigated contentious assessment reforms most successfully have done so through inclusive, transparent engagement. [Kentucky’s multi-year assessment redesign process](#) and [Chicago’s inclusive approach](#) to redefining school quality indicators both demonstrate the value of early and sustained engagement with educators, families, civil rights



organizations, and community leaders. These processes created shared understanding, surfaced legitimate concerns early, and strengthened the durability of the resulting systems.

The Leadership We Need

Technological advances have created unprecedented possibilities for the future of assessment, but the future will not be determined by technology alone. It will be determined by how deliberately states govern a period of transition—one in which old systems are losing legitimacy, new approaches are still emerging, and the costs of getting it wrong are real. The danger in this moment lies at both extremes: clinging to outdated systems that no longer command public confidence and are misaligned with the skills and experiences students need for the future, or abandoning public assessment without building something credible to replace it.

But state leaders can take a different path. One that refines what has worked, acknowledges what has not, and invests thoughtfully in what could come next. A path that treats assessment not as a single test, but as public infrastructure designed to serve families, protect students, and provide honest signals about whether public education is delivering on its promises. The bridge period actions outlined here are intentionally modest in scope but serious in intent: guardrails that protect trust, R&D that builds evidence, and flexibility that creates space for learning without sacrificing transparency.

At its best, assessment is not merely a measurement tool, but also a moral signal about what a society values. The task before state leaders is to safeguard that signal and deliberately calibrate it for what comes next. This means resisting both the comfort of the familiar and the temptation of the untested. In a moment defined by polarization and uncertainty, the country is hungry for state policymakers with the courage, discipline, and vision to step into leadership of this work.



Acknowledgments

We would like to express our gratitude to the many individuals who generously shared their time, expertise, and thoughtful feedback during the development of this paper. Throughout the drafting process, colleagues and partners across the field engaged with early ideas, reviewed multiple iterations of the work, offered detailed written comments, and in some cases took additional time to discuss ideas with us directly. These conversations and critiques challenged us to sharpen our thinking. We are truly grateful to the following individuals for their thoughtful feedback:

Jennifer Alexander, PIE Network

Jonathan Alfuth, KnowledgeWorks

Ivett Arellano, Learning Policy Institute

Thomas Arnett, Clayton Christensen Institute

Madi Ashour, Colorado Children's Campaign

Sarah Bishop-Root, Education Reimagined

Jessica Cardichon, Learning Policy Institute

Kirsten Carr, Foresight Law + Policy

Liz Cohen, 50CAN

Juan D'Brot, Center for Assessment

Sydnee Dickson, University of Utah (formerly Utah State Superintendent)

Chris Domaleski, Center for Assessment

Keith Dysarz, XQ Institute

Ben Erwin, CCSSO

Michael Horn, Harvard Graduate School of Education

Rachael Ledwidge, VT Agency of Education

Susan Lyons, Lyons Assessment Consulting

Michael Maher, NC Department of Public Instruction

Tiffany Miller, Learning Policy Institute

Gretchen Morgan, Linked Learning Alliance

Babak Mostaghimi, LearnerStudio

Nicholas Munyan-Penney, EdTrust

Daniel Oscar, Realign Education

Lillian Pace, KnowledgeWorks

Diana Perdomo, Democrats for Education Reform

Michael Petrilli, Thomas B. Fordham Institute

Nicole Pollock, Democrats for Education Reform

Andy Rotherham, Bellwether

Rachel Safferstone, XQ Institute

Patrick Sims, PIE Network

Juliet Squire, Bellwether

Ned Stanley, 50CAN

Eric Tucker, The Study Group

Christy Wolfe, Bipartisan Policy Institute

About the Authors

Dr. David Nitkin leads Transcend's work related to the future of assessment and accountability

Dr. Deborah Gist is a superintendent-in-residence at Transcend. She was previously superintendent of Tulsa Public Schools, commissioner of education in Rhode Island, and state superintendent of education for the District of Columbia.

Leslie Colwell is Transcend's Director of Policy.