The Campaign for GRADE-LEVEL READING

CHAPTER 2

THE NATION'S REPORT CARD

A CALL TO ACTION FOR RAISING ACHIEVEMENT AND CLOSING GAPS

CHAPTER 2

HOW TO RESPOND TO NAEP RESULTS IN A WAY THAT HELPS LIFT STUDENT ACHIEVEMENT FOR ALL

Handled wisely, each release of NAEP results can be a lever for communities and educators to drive achievement through high-quality data, evidence-based responses, and well-executed implementation. Handled poorly, each can be a missed opportunity — or even worse, cause for despair and new fodder to feed a fatalistic narrative. This section is a guide to preparing for all future NAEP releases.

The declines revealed by the 2022 and 2024 NAEP reports have created an almost irresistible urge to jump to convenient conclusions about why the bottom keeps falling out and how to fix (or abandon) public schools. For those long focused on achievement, **the temptation to leap past careful analysis and dive straight into remedies is strong**.

But some panelists stressed that urgency without nuance doesn't work. Racing to reactions and solutions without deeper analysis doesn't lead to lasting results. The NAEP dialectic requires real understanding, informed action, and the determination to make the action truly effective.

Several experts exemplified how to be curious about each NAEP rather than react with preset conclusions. Munro Richardson, Executive Director of Read Charlotte, pressed for deeper analysis of the states with excitingly successful results. "Did the state do something different — or were there only some districts that did something different there?" Sara Randazzo, the influential Wall Street Journal education reporter, was eager to "dig more into the research around why the lower-performing students have gone down more. There's a lot of theories out there," none of them simple or proven or necessarily correct.

"I just exercise caution in trying to go from NAEP results in a particular jurisdiction to my preferred policy solution," said Morgan Polikoff, professor at USC's Rossier School of Education. "That's the trend a lot of us fall victim to."

Panelists called on one real-world example after another to describe the lessons they've learned to bring to NAEP every two years. If we could sum that mindset up, it would include this insistence on painstaking analysis rather than leaping to conclusions; an openness to learn and share information; and the gritty resolve to catalyze those around you to act.

LESSON 1

NAEP IS ONLY A BEGINNING OF INSIGHT, ANALYSIS, AND CHANGE

For all its value, NAEP data doesn't provide detailed enough disaggregation to drive serious education policy on its own. When it releases the NAEP results, IES includes few insights about what the findings mean for educators or policymakers. "When we think about those instructional changes that need to happen in the classroom, we need very targeted data to do that," said Kristen Huff of Curriculum Associates, a 50+ year educational consulting firm known for sharpening school practices. "**That's not going to come from NAEP.**"

To take full advantage of NAEP to drive student achievement, far more data work, from many different sources, is required.

One example: The Education Recovery Scorecard (ERS) — a mid-pandemic project of Thomas Kane, Professor of Economics at Harvard Graduate School of Education, and Sean Reardon of Stanford Graduate School of Education — has marshalled local test results from roughly 11,000 school districts across 43 states[1]. By making everything comparable to the NAEP threshold, ERS has given educators the first-ever view of "achievement across state lines but also for a given district over time," said Kane, a webinar panelist.

Data dives post-NAEP are critical because NAEP only provides results by state and for 26 major cities. **While NAEPs over time identify macro trends (only two states have improved in both math and reading since 2019), ERS reveals micro trends (over 100 districts outperformed pre-pandemic levels in both subjects).** Though it has yet to capture scores for all third to eighth graders, ERS has already spurred an important new line of inquiry by detailing, said Kane, "the magnitude of the improvements in districts like Birmingham and Los Angeles, Compton, and DC."

The **Trial Urban District Assessment (TUDA)** also provides a more useful and granular view than the state-by-state NAEP. An expansion of NAEP, TUDA tracks student performance in 26 urban school districts, including those in Chicago, New York City, and Los Angeles, allowing researchers and policymakers to dig locally into achievement gaps and tailor responses to the district's particulars, avoiding one-size-fits-all solutions. TUDA provides the data that ERS turns into some of its findings and into key questions like: **Which practices from high-performing urban districts offer the most promising models for communities still struggling with learning loss?**

[1] ERS researchers excluded data from states that did not post sufficiently detailed data, changed their state tests between 2015 and 2023, or had fewer than 95% of students take the state assessment. As a result, a total of eight states were excluded from the analysis leaving a total of 42 states and the District of Columbia.

LESSON 2

MOVE QUICKLY BEYOND NATIONAL OR STATE AVERAGES

Not only are NAEP results insufficient for policy decisions by themselves, but overemphasis on the assessment's broad national and state averages can lead to genuinely damaging outcomes. Said Virginia Secretary of Education Aimee Rogstad Guidera:

> "When I arrived in this role and said we're going to restore excellence to education for every child in Virginia, I was met with, 'How dare you? This is Virginia. We have excellence everywhere.' And I said, 'Have you disaggregated your data? Good luck telling some of these families that their kids are getting a great education.' And it's because broad data masks everything."

The problem is that surface-level NAEP numbers mask the realities that must guide *effective* action. As Maryland State Superintendent Carey Wright, who led Mississippi's remarkable reading transformation, put it: "Behind every data point is a face. And if you don't know who that face is, you're not going to be able to figure out what that child needs." Guidera warned:

> "We all fall into the lull of averages. **If I could do anything, I'd blow up that we just keep using averages when we talk about NAEP.** Until we start really talking about all kids in a way that we see every child, it's too easy to ignore the kids who need the most. We don't actually see what's happening."

Effective insight and solutions require deeper disaggregation by race, income, geography, English learner status, disability, and many other factors. The difference between surface-level and deeper analysis is stark. Munro Richardson from Read Charlotte revealed what would have happened to the ubiquitous 2025 headlines proclaiming "40% of Fourth Graders Score Below Basic on Reading" if we had simply factored in socioeconomic data (SES).

In fact, Richardson reported, **among low SES students, almost 60% were "Below Basic."** Meanwhile, among high-SES students, 60% were "Proficient." This rudimentary disaggregation exposes what Richardson called "two entirely different worlds" within the averages that have been the basis of news coverage. With even this minor disaggregation, the national narrative shifts from "American students are struggling" to "low-SES students are being failed while high-SES students maintain proficiency."

The stakes of deeper analysis extend beyond statistical precision — they shape whether we identify the right problems for the right students. Going back to 2018, for example, Brookings analyzed NAEP achievement gaps. NAEP data back then still showed gaps narrowing or static by race and also the standard measure of poverty — free/reduced-price lunch qualifiers. But Brookings revealed how the addition of better data, in this case Sean Reardon's far more detailed family income-based data, found that achievement gaps between students in the 10th and 90th percentiles of family income were actually exploding. This critical finding, the authors noted, was "outside of NAEP's view."

This anecdote is not an outlier. **It's independent research based on additional investigations that reveals what's truly happening beneath NAEP's big-picture gaze.** In this case, the lowest-income children were falling behind much faster at just the moment that growing U.S. income inequality was leading more children to be categorized as very low income. Based on Reardon's data, it was impossible not to see that responses tailored to America's youngest and most economically challenged students were critically needed. (Examples: national pre-K, wraparound services, high-dosage tutoring, afterschool programs.) Relying only on NAEP metrics and broad averages had failed to specify the correct takeaway.

Charles Barone, Senior Director of the Center of Innovation at the National Parents Union, worried during one webinar that the same situation is occurring today in New Jersey, where he lives. New Jersey scored in the top five states across all age groups in the 2024 NAEP. Yet the state, he said, is "resting on its laurels" and failing to respond to the extreme needs under the surface of the broad brush of averages.

> "There's lots of kids who you saw dipping. We're not helping them. So if you care about equity, you want to pay attention not just to the average achieving student, but where those already at the bottom are losing ground."

Webinar panelist Kevin Mahnken, reporter from The 74, the education-focused website, demonstrated the power of deeper disaggregation to raise key questions. Most racial groups in NAEP 2024 tracked the same pattern in which top-performing cohorts held steady while low performers fell dramatically. But he discovered that the pattern among one group, Latino eighth graders, was starkly different and would require deeper analysis.

> "The Hispanic eighth grader reading drop [over two years] was such a thoroughgoing decline. You saw declines for reading at the 25th percentile, so the low performers. But also at the 50th percentile, the average Hispanic eighth grader dropped significantly. And also at the 75th percentile. So even the comparatively high-performing students did fall."

Granular disaggregation by sub-subgroups produces findings that are less easy to sum up and require more exploration. Katharine Stevens, Founder and President of the Center on Child and Family Policy, gave an example:

> "When you disaggregate Black students by not only income but maternal education and family structure, there's gigantic variation by those subgroups — ranging from 17% scoring 'Below Basic' to 64% scoring 'Below Basic.'"

A 47-percentage-point spread within a single broad demographic group fundamentally challenges simplistic narratives about student achievement. It reveals that race alone explains far less than we might assume, and that family structure, educational background, and economic stability are correlated with dramatic differences in outcomes even within the same racial category. These insights are invisible when we only examine broad racial categories or even income cohorts, said Ian Rowe, Senior Fellow at the American Enterprise Institute:

> "NAEP as everyone knows is the scorekeeper, not the entity that's trying to say why these results are coming forth. But **because NAEP reports data through race, class, and gender, what I call the usual suspects, those tend to suck the energy out of the room for any other explanations that could be transcendent."**

The sub-subgroup approach forces us beyond comfortable generalizations toward analysis that can actually drive improvement. It reveals bright spots within demographics and communities often written off as uniformly struggling. This complexity shouldn't paralyze action — it can focus it. Rather than implementing broad programs based on surface-level demographic patterns, deep disaggregation allows for tailored responses that speak to the specific combinations of factors that drive strong or poor performance.

LESSON 3

THE COMMUNITY CONTEXTS THAT NAEP CAN'T CAPTURE MUST BE UNDERSTOOD

Applying the NAEP report card successfully to help any school requires factoring in the context outside its boundary. Because look at what that boundary might be: a chain-link fence, a packed city sidewalk, a rustic stone wall, or a grassy sports field. All so different, all reflecting different community environments. That was a message repeated by many speakers: NAEP analysis and resulting responses must make sense for the specific families, communities, and kids they're meant to help.

Almost 15 years ago, just as NAEP averages stopped rising, Matthew Di Carlo, a researcher at the Albert Shanker Institute in Washington, D.C., made **two points about the "answer" to declining NAEP scores: 1. It's not one answer; 2. A lot of the answers lie outside of school.** He wrote:

> "Discernible trends in NAEP test score data are almost certainly due to a combination of factors, and it's unlikely that one policy or set of policies is dominant enough to be identified as 'the one.'
>
> [Yet] people involved with these speculative arguments seem a bit too willing to assume that schooling factors — rather than changes in cohorts' **circumstances outside of school** — are the primary driver of NAEP trends."

Experience has certainly re-taught us again and again that it is a **myth to think of schools as simple instruction centers.** Those interpreting NAEP's results and acting to lift them in historically marginalized communities must consider the community context that helped produce the scores — or their responses will risk ineffectiveness.

Three such contexts came up repeatedly for panelists. Each profoundly impacts how well children learn, think, and score collectively on NAEP. Yet improving them is mostly beyond the control of individual classroom tactics and entails shifting the entire ecosystem in which children live and learn — a worthy goal but one that requires community-based responses.

ATTENDANCE: Getting from home to the school

Attendance came up repeatedly as the central blocker of academic success. Hedy Chang, Founder and Executive Director of Attendance Works, set the national context: Chronic absenteeism has doubled since the pandemic, with two-thirds of schools now reporting rates above 20%. Lori Masseur, Early Learning Director at Read On Arizona, connected the dots between those rates and achievement: "For every 1% we improved attendance, proficiency rose 1.5 points."

The frequently made point: **Schools can only establish instruction at a level that lifts achievement if students are predictably present.** Negative outcomes, said Chang, speed up "every year that you're chronically absent, particularly in these early grades. That's when you're building the foundational supports for student success, when learning is scaffolded and habit forming." Poverty in turn appears closely correlated with attendance slides. **Historical research shows that children from economically challenged families are four times more likely to be chronically absent in kindergarten than their peers from higher-income families.**

TEACHER MORALE

Kevin Mahnken of The 74 identified a pattern in data released by NAEP: Teachers of high-performing students were more likely than teachers of low-performing students to be satisfied working in their school. (NAEP conducts a survey parallel to the exam in which students and teachers answer questions covering everything from who lives at home to reading interest to instructional practices — another source with which to disaggregate NAEP scores.)

Evidence is strong that **discouraged teachers disproportionately impact schools serving historically marginalized communities** where poverty, absenteeism, low achievement, and teacher turnover are highest. Arthur Everett, a high school social studies teacher in Brooklyn, New York, and winner of the Empire State Excellence in Teaching Award, testified that insufficient support for reforms and funding contribute to teacher discouragement, **but that low morale at its core is due to runaway community expectations of teachers.**

> "Teachers are maybe locked into some performative or administrative things — old ways of being. **But there needs to be more recognition that the system is not working and that it should not be on the back of individual teachers to uphold it.** We romanticize in movies and human interest stories the superhuman teacher who did above and beyond for their 30 kids, but the system shouldn't be built on the idea that you have to break your back in order to be considered a good teacher."

Evan Stone, CEO of Educators for Excellence (E4E), which represents 40,000 teacher-advocates across the country, reported that E4E's survey of teachers found that only 19% of them recommend the profession to others and only 16% say the profession is sustainable. **"NAEP tells us that our student achievement is in crisis. I would say the teaching profession is also in a state of crisis, and we need transformation for both."**

CRISIS AND TRAUMA

Community realities often create learning obstacles that register in NAEP data as academic deficits rather than what they actually represent: how stress, fear, and instability harm developing minds. Panelists testified that when students arrive at school carrying the weight of housing or food insecurity, family trauma, or community violence, their academic performance can reflect their mental state rather than their teachers' effectiveness or the school's instruction.

Jean-Claude Brizard, CEO of Digital Promise, urged policymakers to recognize **that students may score "Below Basic" because trauma literally drains cognitive resources and rewires developing brains.** "Psychological stress, such as fear among undocumented families, impacts how children learn," Brizard said.

LESSON 4

STATE AND DISTRICT DATA IS WHERE ACCOUNTABILITY MUST BEGIN

If a state or city is surprised by NAEP results, something is fundamentally wrong with their data systems. NAEP's perpetually looming benchmark should motivate school systems to track student achievement much more frequently — ideally in real-time. The message to educators from many panelists: Don't wait for the next warning, get ahead of it.

The tools exist to create educational systems that operate like the National Hurricane Center's Cone of Uncertainty, alerting vulnerable populations in time to escape from harm's way. The question isn't whether we can build such systems, but whether we will.

At the state level, NAEP's biennial results, particularly the percentage of "Below Basic," must be viewed as a recurring opportunity for recalibration of state tests and standards. States that continue to let their own proficiency standard lag far behind NAEP's should close the differences. Carey Wright, Maryland State Superintendent of Schools, revealed that while NAEP 2024 showed Maryland moving from 40th to 20th in fourth-grade reading over just two years, her state's assessment was failing to capture the shifts with enough depth. "We just issued an RFP yesterday for **a brand new statewide assessment that we're aligning more to the rigor of NAEP.** That's critical."

Alabama State Superintendent Eric Mackey also stressed commitment to assessment integrity: "Fifteen years ago, we were showing 85% to 90% of our kids proficient on our state test, and NAEP did not show the same. So we adopted a test that's much more rigorous — because it gives parents real results about what's honestly happening with their students." Five years ago, the state adapted further by remaking its assessment again. "**We want the results we're seeing on our state test verified by the NAEP,**" Mackey explained.

Multiple panelists, particularly state school chiefs, described how achievement data can be used to fight for state reforms, and encouraged all states to build a data-to-difference-making pipeline. Indiana Secretary of Education Katie Jenner described how her state used data to trigger improvements in reading performance. "For four years in a row, since the pandemic, we've had [legislative] bills that just put all the attention on reading. You have got to talk about it. When you're tired of talking about it, you should talk about it more."

Margaret Spellings, who led the federal Department of Education during much of the Bush administration, during which states became more accountable and student scores improved, said:

> "When we focus on every kid, measure (yes, that's testing) and report that data, hold it up and learn from it, and make it a central part of what we invest in, we go in the right direction."

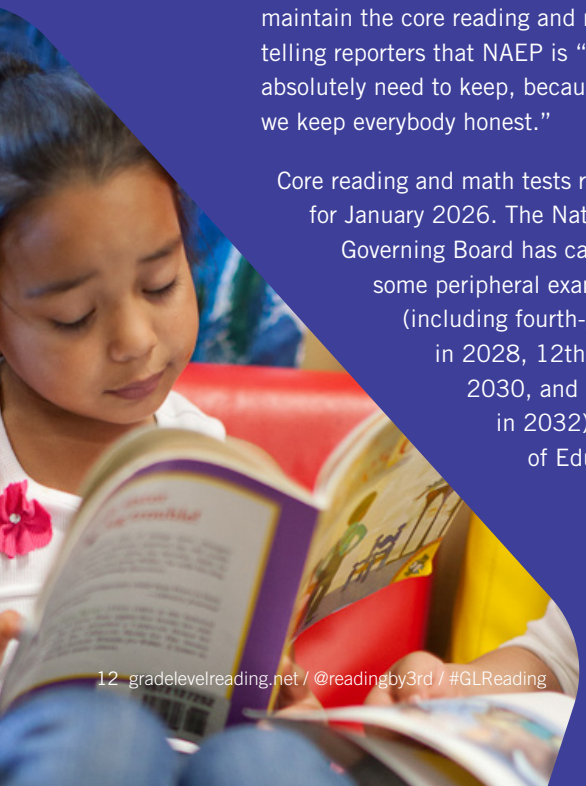
NAEP's Status at the Time of Publication

NAEP remains in place despite significant upheaval at the federal level in 2025. Education Secretary Linda McMahon has pledged to maintain the core reading and math exams, telling reporters that NAEP is "something we absolutely need to keep, because it's a way that we keep everybody honest."

Core reading and math tests remain on schedule for January 2026. The National Assessment Governing Board has canceled or deferred some peripheral exams to save costs (including fourth-grade science in 2028, 12th grade history in 2030, and all writing tests in 2032). The Institute of Education Sciences (IES) and National Center for Education Statistics (NCES), which oversee the program, have few remaining staff members.

"I'm worried about the health of the NAEP," said former U.S. Secretary of Education John King, "the defunding of IES, and our ability to have good information about where we're succeeding and where we're struggling."

Even so, Fordham Institute President Michael Petrilli, who helped create the U.S. Department of Education's Office of Innovation and Improvement and is an enthusiastic consumer of data from NCES, pointed out that the agency recently brought in a respected researcher to help redesign IES. "Just a little bit of optimism there: They seem to understand the value of it."



How Leading Districts Use Data to Accelerate Progress

The highest-achieving systems have learned to make data work for students in real-time. The School District of Indian River County has consistently improved achievement, moving from 38th to fifth among Florida school districts in the past seven years; the district is now ranked second in third-grade reading proficiency. Superintendent David Moore testified that much of the rise stems from the innovative use of data. He summed up Indian River's approach as: **"One clean real-time data system aligned to our district strategic plan."**

Included in Indian River's data system are not only individual student assessments, but also which organizations and agencies are working with which kids, what instructional coaching has occurred, specific challenges faced by principals and teachers, what instructional strategies yield the biggest improvements, how well those who provide each intervention are doing their jobs, and classroom observational research.

> "It's data in real-time that forces the conversation around: What does good instruction look like? How fast can we make decisions that are as close to kids as possible? I think **some of the outcomes we've been able to realize have been solely about making an instructional shift based on our data as quickly as possible.** Rather than waiting three weeks or four weeks, we're able to evaluate the things that matter the most to determine where we need to make an instructional shift or professional development evaluation that improves outcomes — moving as fast as possible."

In truth, most districts struggle to stay ahead of annual state assessments and the biennial NAEP report card. That means parents, who need transparent information about their children's education to make informed decisions, are often not aware of their school or student's assessments. "School performance data," reported Ashara Baker, New York State Director for National Parents Union, "tell it how it is. Whether it hurts or not, we have to know how poorly these schools are performing."

Baker discovered this firsthand when she started having conversations with families across New York in her role. "The biggest thing that stuck out was families really didn't know how students were performing." The scope of this information blackout was worrisome to her.

> "New York's NAEP results show that only 31% of fourth and eighth graders are 'Proficient' in reading. [Using state assessments,] there are 15% of Rochester third graders and 4% of eighth graders who are proficient in reading. But when I give this report to a parent in Rochester, they don't know what that means. They don't know how bad their schools are doing."

John Johnson, a parent leader with the National Parents Union Leadership Council on Assessment and Accountability, described a similar information gap in Colorado. "I don't think that's even close to being adequate," Johnson said about the limited data parents receive at twice-yearly parent-teacher conferences. He emphasized the deeper problem: "They just feel like they're not being informed. I hear, 'oh, they're doing fine. We're not concerned.' But that doesn't really spell out what the issue is or if there is an issue." Johnson noted that by the time concrete data appears in fourth-grade NAEP results, critical learning gaps have already formed in earlier grades when parents received vague reassurances instead of actionable information.

This isn't just a communication problem — it's a barrier to educational equity that rolls up into growing achievement gaps on NAEP. Parents who don't understand how their schools are performing can't advocate for better resources or demand improvements. The consequences ripple through entire communities, trapping families in systems that are failing their children without transparency about how those systems are performing.

Given NAEP's long lead time, state and local assessments must take on the job of charting student achievement trajectories, particularly for at-risk populations. This data needs to reach families, teachers, principals, voters, and legislators with clarity and urgency.

Experts from the National Parents Union (NPU), the network of 1,800+ parent organizations representing more than 1.7 million parents and caregivers, testified to the power and necessity of robust data systems. When data flows effectively to parents and communities, transformation accelerates, said Keri Rodrigues, the organization's founder, describing how transparent information has helped mobilize NPU's parent-led advocacy groups:

> "It's wonderful to watch parents harness this data, make it actionable, and call for change — literally creating their own reports so they can hold systems accountable and demand evidence-supported solutions. This is the power of schools giving parents information instead of being paralyzed by fear that bad news will make parents angry."

Without NAEP, policymakers, educators, and parent groups would lack the external benchmarks needed to make informed decisions about resource allocation, intervention strategies, and systemwide reforms. (See *"NAEP's Status at the Time of Publication" on* page 12.)

LESSON 5

NAEP AS LEVERAGE FOR EVIDENCE-BACKED RESPONSES

The latest NAEP results underscore a reality in the current landscape: States possess unprecedented power and resources to lift achievement, yet outcomes vary dramatically based on how they leverage these capabilities. Observed Kevin Huffman, former Tennessee state education chief and current CEO of Accelerate, a national initiative to embed high-impact tutoring in public schools:

> "For 10 years, states have had a fair amount of autonomy and ability to step up, define their plan, define what they need, and then execute strategies. To the extent people were waiting for central guidance from the federal Department of Education on what's our big plan to help kids — that's not happening. **Some states have really taken advantage of their autonomy. Some have not.**"

Mississippi's fourth-grade reading gains over the past decade produced **an especially compelling set of best practices that state leaders can master, mimic, and spread**, said Indiana State Superintendent Katie Jenner. Beyond sustained political commitment across multiple election cycles, Mississippi used specific literacy reforms, teacher training approaches, and instructional materials adoption to guide the way. "A lot of the solutions that we've deployed in reading, we've copied straight off the playbook of Mississippi," Jenner testified. Indiana fourth-grade reading is nearly back to where it was before the pandemic.

NAEP results can trigger state education leaders to research what their state's outstanding districts and schools are doing that others might emulate. "Take deep dives with these educators," advised Kristen Huff, Vice President of Assessment and Research at Curriculum Associates.

> "There are educators out there knocking it out of the park — getting students who are behind grade level onto grade level, getting students who don't know how to read to learn how, getting students who don't know how to be fluent with their math counting to be fluent."

Every NAEP release, added Colorado Commissioner of Education Susana Córdova, is an opportunity for state leaders to focus more sharply on what works to help children from challenging circumstances. Even as her state scored above the national average overall in 2024, she said, "about a third of our fourth graders are 'Below Basic' on NAEP. We know if kids aren't on grade level in third grade, so many other things are going to be challenging for them."

> "I just left the room where our statewide data, our statewide plans, and our entry points, particularly for the most struggling schools, are posted all over the wall. Our internal Colorado priorities are to move with urgency for the most struggling schools."

Spreading actionable solutions is critical for a state that seeks to make districts accountable to higher student performance. That's because **without models to adopt, a state's top-down demand for improvement will breed resentment rather than results**. As John King emphasized:

> "One of the greatest threats to accountability is people feeling like they don't know how to improve. Because if people feel like they don't know how to improve, then they start to resent the accountability. It feels like a 'Gotcha.'"

This tension between those who establish accountability yardsticks and those on the ground who are responsible for improvements highlights the vital role of research partnerships. Rice University professor Ruth López Turley, Director of the Kinder Institute for Urban Research and former member of the National Board for Education Sciences, emphasized the critical importance of school districts drawing on best-practice researchers when they seek improvement.

As an example, Turley described her work with the Houston Education Research Consortium, in which national research that overwhelmingly points toward the effectiveness of full-day pre-K programs in improving student achievement only gained traction with local policymakers after it was validated by Houston-specific data and experience.

"Connecting national evidence with local evidence — they can really reinforce one another," she said, going on to advise:

> "The use of data [by researchers or advocates] can be off-putting to school districts because it can be weaponized, used as a gotcha moment to embarrass them. But if there's a close partnership with school districts where the [researchers'] goal is 'We're not trying to point fingers at anybody, we're just trying to improve student outcomes,' then **really strong relationships of trust can work on really challenging problems together in a constructive manner. Beautiful things can happen.**"

These evidence-based solutions — literacy reforms, tutoring models, community partnerships — all depend on the kind of innovative data systems discussed in the previous lesson. Take the Mississippi example of dramatic reading improvements. Adam Gamoran, President of the William T. Grant Foundation, described how important federally supported data initiatives were to the state's turnaround, referencing "NAEP scores that showed the dire situation for decades, research on strategies to improve reading, evaluation studies showing us where efforts succeed, fail, and need improvement. Talk to people in the Mississippi state government about how they did it, and they point to IES's Southeast Regional Educational Laboratory (REL) that did a lot of that work."

When states lack that infrastructure or federal help (REL Southeast has been sharply cut back in 2025, Gamoran noted), they cannot track whether their responses work or defend effective programs from cuts. Maryland's experience, as described by Pascale Small, Deputy Director of Impact and Engagement for the National Parents Union, shows what happens when that data foundation is missing.

> "We asked for data about wraparound services in Maryland's 600 Community Schools so we could protect these programs from budget cuts, but we couldn't get it. Without reliable, timely, and actionable data, I can't advocate for children or my community. This creates a cycle where no one gets the resources they need."

The data gap extends beyond timeliness to basic inclusion. Tiara Thomas, a member of the Piscataway Conoy tribe who serves on the National Parents Union Leadership Council on Assessment and Accountability, described how Maryland's Indigenous students often don't appear in achievement data at all. Without data showing which students need services, her district — the only one in Maryland still maintaining an Indian education program — struggles to justify federal funding. The invisibility creates a catch-22: "How do we put that data in the grant if we don't have the data?" When schools cite privacy concerns as reasons not to share information with small populations, Thomas pushed back directly: "I'm always going to be the smallest community. So if I took that answer all the time, I'd never have information to do anything."

LESSON 6

DON'T SLEEP ON IMPLEMENTATION (UNDERSTANDING THAT GOOD POLICIES ALONE ARE NOT ENOUGH TO LIFT STUDENT ACHIEVEMENT)

States and districts regularly announce new reading initiatives, math curricula, or teacher training programs, but the execution — how consistently and effectively these changes reach students — often remains incomplete, unmeasured, and under analyzed. This gap between intention and reality is often why even proven practices can fail to move NAEP scores.

Implementation begins with matching the specific model of a broad response like high-impact tutoring to a specific community. Said Adam Gamoran, President of the William T. Grant Foundation: "We have specific research that details not only that tutoring is effective, but **how to implement it properly — what types of dosages are available, what characteristics of tutors are needed, how much coordination is needed between tutors and teachers.**"

The trick, Gamoran explained, is for local educators and communities to find the correct version to implement within their local context.

> "We have a lot of knowledge. But to put findings into practice and evaluate them, we need partnerships between the researchers in tutoring models and the communities or school districts. Educators and district leaders don't just say, 'I need tutoring. I'll go to the research.' They say, **'What tutoring programs have worked in districts like mine, with communities like mine?'**"

Once a response has been kicked off, implementation tracking requires different metrics than outcome tracking. Rather than waiting for the next state assessment or NAEP result, states and districts must actively monitor whether teachers are using new curricula as designed, whether professional development is changing classroom practice, and whether tutoring programs are reaching the intended students with the right intensity and duration.

Many of the implementation difficulties raised by panelists concern teacher training, particularly in the science of reading. **States that have invested heavily in professional development in science of reading practices now need to make sure the training is effective and reaches kids.** Indiana created an excellent best-practices teacher training program, called Indiana Literacy Cadre, that Secretary of Education Katie Jenner said produced increases in scores that were 12.5 times higher in schools that participated compared to schools that did not. Even this successful intervention, however, with data to back up effectiveness, isn't being fully implemented because schools were given an out. "Those schools that are not participating — that's unacceptable," Jenner said.

Meanwhile, new teachers come into schools unaware of the science of reading because education schools and teacher preparation programs don't teach it. As John King observed: "**It's not necessarily popular to say to every teacher prep program in the state, 'You're going to need to change your syllabi to reflect the science of reading.** You're going to need some faculty to deepen their expertise.'"

The challenge of inconsistent teacher preparation extends beyond individual programs to entire state systems. Maryland was funded by the Elementary and Secondary School Emergency Relief Fund to push high-dosage tutoring that emphasized the science of reading. "But the problem," said Carey Wright, Maryland State Superintendent and former head of Mississippi's schools, "was that **there was no standardization around what professional learning of the science of reading looked like.** And so it varied by district." The state has since stepped in, giving everyone "the same high-quality professional learning."

The implementation challenge becomes even more complex for multilingual learners. Melissa Castillo, former senior advisor to the U.S. Secretary of Education, described how Arizona addressed this: "We have legislation that's made very clear the types of training that teachers have to take, but unfortunately, we didn't explicitly speak to what that means for English learners. So the state had to not only provide guidance, and support around the science of reading, but training that spoke to what that means if you're learning how to read in Spanish, not in English."

Like implementation difficulties, **implementation victories are often hidden away in institutional decision-making.** Yet these must be celebrated and replicated when possible. Joan Dabrowski, Baltimore City Schools Chief Academic Officer, described how real-time sharing of classroom information in recent years helped drive the district's noteworthy student achievement improvements. The sharing came about because the district's Principal Supervision team underwent a shift in department. "We think joining them and my Academics team into one team was a game-changer for us," Dabrowski explained.

> "As our Academics team is developing curriculum and supporting professional learning development, we have Principal Supervisors and math and literacy specialists inside each school getting on-the-ground data about what's working. We come together with them every week. We hear in real-time that, 'Wow, there's this particular curriculum unit that is really tripping people up.' Or, 'it's taking longer than we expected.' Then our team can push out updates and additions. **We're trying to interrupt these silos that are very common across districts.**"

Every NAEP release also brings a wave of what Polikoff long ago coined "Mis-NAEPery," incidents when NAEP results are presented inaccurately, out of context, or for purposes other than lifting achievement in the classroom. Those committed to student achievement must be prepared to counter these uses with the speed and analytical rigor NAEP deserves, Polikoff told panelists:

> "I'm very concerned about people misusing NAEP for political gain. You see folks who are real strong advocates of dismantling public education saying, 'Oh well, public schools aren't working.'"

The most common form of Mis-NAEPery involves cherry-picking statistics. One commentator said the most recent NAEP showed that "an astonishing 69% of fourth-graders and 70% of eighth-graders scored at or below Basic in reading" — slyly adding the Basic cohort into the Below Basic group to gin up what the commentator called "troubling reality" to advance a pre-existing policy preference.

Converting NAEP-driven statistical snapshots into blanket condemnation is unsettling to those inside schools, said Karyn Lewis, Vice President of Research and Policy Partnerships at the education consulting firm NWEA. "The worst aspect I worry about is how it lands with teachers and school leaders that have just been killing themselves over the last four years to try and respond to this crisis and support learners. It must be so heartbreaking."

Mis-NAEPery isn't always about attacking public education — it can also involve premature celebration that masks deeper problems, including rising achievement gaps. As Keri Rodrigues of the National Parents Union observed about Massachusetts: "Boy, was my governor excited to tell everybody about the NAEP scores and how we were number one and King of the Hill. But when you take a closer look at those numbers, it didn't seem like there was a lot to be celebrating there." **Scores among Black children in Massachusetts** continued to fall sharply.

It may seem to be thankless advocacy work to perpetually counter Mis-NAEPery — to publicly demand more rigorous analysis, call statistics-based political theater what it is, and insist that NAEP results be used to strengthen rather than abandon our schools. But there's a battle of narratives afoot and NAEP results are a big part of it. **We can't let this powerful tool of accountability become a sledgehammer assisting a teardown.** Margaret Spellings argued that education advocates must prioritize substantive discussions about student achievement in today's "wildly competitive, AI-innovating world."

> "As we engage with policymakers, talk about the data. How are we doing? Are we going in the right direction? Where do we lag and in what subjects? We have to keep reminding people what's at stake and where we are. We can't do it enough."

gradelevelreading.net / @readingby3rd / #GLReading

The Campaign for
GRADE-LEVEL
READING