

ESSA and State Capacity

Can the States Take Accountability Seriously?

Arnold F. Shober

US Senator Lamar Alexander (R-TN) presented US Secretary of Education John King with a litany of ills in a testy Senate hearing in early 2016: “Local school boards, classroom teachers in states, had gotten tired of the US Department of Education telling them so much about what to do . . . the Department of Education had become a national school board, telling Kansas what their standards must be, telling Tennessee how to fix failing schools, telling Washington State how to evaluate teachers.”¹

Alexander left no doubt that the US Department of Education had become an albatross. In his view, the Every Student Succeeds Act is a good south wind, pushing federal policy away from its suspicions of the past. For fifty years, federal education policy assumed that states had little will and less capacity to improve American education. Federal policy was prescriptive, categorical, and rigid.² The epitome of this distrust toward states came with No Child Left Behind in 2001 as states were handed strict requirements about standards, assessment, and the consequences of both.

For all their suspicion, however, federal policy makers have had little progress to show for their efforts. A racial and ethnic gap in academic performance has remained stubborn and wide. Disparities in quality between teachers, schools, and districts have remained large. Skeptics on the right and left have wondered whether federal policy is now a stumbling block. Perhaps the states could show leadership. Some already have: Tennessee, Massachusetts, California, and other states have innovative assessments and school choice programs. Maybe the states could perform. In this, ESSA is a watershed. It surrenders federal leadership to the states.

Ironically, most states can exhibit such initiative only *because* of earlier federal leadership. Fifty years of federal direction has imbued many state and local officials with a desire for quality equal education. The success of the Common Core State Standards by 2014 spoke to the near-universal desire of governors and state boards of education to give all students some semblance of quality education. Governors have been on board this train for close to twenty years.³ Fifty years of federal tutelage has also endowed the states with the technical capacity to design, measure, and advance achievement. Despite its critics, NCLB dramatically improved states' data-collection and processing abilities.⁴ But a secondary accomplishment of the federal standards regime was to uncover wide disparities among the states. The median students in Massachusetts, New Jersey, and Minnesota have outperformed the best students in Alabama and Mississippi. Latino and African American students have continued to lag behind those in other groups. Further, the dramatic decline in "proficient" test scores from Common Core-linked tests has created an unenviable crucible for state officials. Paradoxically, fifty years of federal compliance has helped give states the capacity to drive student learning. But without federal oversight, it's uncertain whether they can transform compliance into creativity.

ESSA'S POTENTIAL

ESSA's architects have great expectations for the states, but the states will succeed only if they do not squander their inheritance. First, states have a tremendous opportunity to replace NCLB's simplistic performance metrics with multiple measures that are honest about individual students' learning and the role that schools and teachers play in that process. Indeed, rules implementing the law inserted the requirement that the measures be "fair" to participants in the process.⁵ Alternatives including surveys, scale scores, and student growth matrices all hold promise to this end, but any changes will require careful thinking in the statehouse and state bureaucracy. States will be tempted to tinker with existing systems and continue the fairy tale of unbelievably high proficiency rates.⁶

Second, ESSA opens new frontiers for the use and understanding of student data. Student data may sound esoteric, but states have the opportunity to leverage data to benefit teachers' practice. Yet states will be tempted to yield too much of the assessment process to teachers and districts, potentially

undercutting the data's promise. These groups have little incentive to report on failures, and if they're given too much leeway, ESSA could obscure academic inequities more than they often already are.

Third, ESSA allows states to move beyond NCLB's "short list" of school remedies (school choice, tutoring, or reconstitution). NCLB created a significant research base for states on school turnarounds. Combined with new, more flexible turnaround spending in ESSA, the states could significantly improve their services to low performing schools. Still, any aggressive interventions will prompt local backlash, and ESSA will not allow the states to shift blame to the federal government. Real turnarounds will require bold—and risky—leadership from state politicians.

Performance Metrics

The first, and easiest, opportunity for states is the design of better performance metrics. NCLB's Adequate Yearly Progress (AYP) was widely panned as unrealistic, statistically unsound, and subject to political manipulation. And it was. ESSA abolished that system, but retained the commitment to school-level transparency and specified two well-worn measures of academic proficiency: performance on state tests and graduation rates.⁷ States will have little trouble extending these elements. Tony Evers, the president of the Council of Chief State School Officers and Wisconsin state superintendent, suggested that states' existing work created "an emergent research base to show what might work well and how a state could begin to implement these type of innovations."⁸

But *can* the states be innovative? They have shown remarkable reticence. They were given an opportunity, not a mandate, to act when the Obama administration's NCLB waivers offered them the chance to experiment with multiple, alternative measures. Eighteen states took the opportunity, but they were decidedly uncreative. Of these, twelve used ACT or SAT scores or participation rates, ten used attendance rates, and, in a throwback to the 1980s, seven used a count of advanced course taking in high school.⁹

Each of these measures had a real and measurable correlation with school-level or student performance, but none suggested that states were willing to experiment with *better* ways to identify schools (or students) who need the most help. Fortunately, two promising "inside-the-box" reforms—a "transition matrix" of student performance, and an index of social and emotional learning—have been tried in real public schools and so are part of Evers's "emergent research base."

Transition matrices eliminate the unintended consequence of categorizing students' performance into levels. NCLB categorized students' performance as below basic, basic, proficient, or advanced. But NCLB led many school officials astray. Because it required schools to show improvement over time, some school officials directed their best efforts toward students who scored just below the proficient cut score. There was no payoff for schools if higher or lower performing students scored better; those students would not "improve" into a new performance category. Transition matrices thwart this temptation by creating per-student growth plans. Instead of rewarding or sanctioning schools for a snapshot of overall student performance, students receive "points" for raising their test score enough to move to the next proficiency level compared to the previous year's test. To find a school's score, the state finds the mean student growth score for all students in a school and compares that number to state growth targets. This is not a statistically savvy approach, but it is tailored to effective implementation. The matrices are easy to explain to teachers, principals, and parents.

Second, policy analysts and a few school districts have experimented with a variety of tests to measure "social and emotional learning" (SEL). These tests ask students or their teachers a battery of questions about students' social awareness, self-management, and decision-making skills. Not only are they currently available from a variety of universities and vendors, but most of them have already been carefully studied for their relationship to academic performance, school completions, or at-risk behaviors that would impair school performance. The best-known examples are the Devereaux Student Strength Assessment (DESSA), the Comprehensive School Climate Inventory (CSCI), and the Social Skills Improvement System (SSIS).¹⁰ These tests provide individual-level data and could be easily paired with student assessment data, allowing districts to control adequately for different school populations.¹¹ Neither of these measures is revolutionary, but that's central to their appeal. Even when states were offered the opportunity to think differently about assessment by the US Department of Education, none launched truly experimental measures. The benefit of transition matrices and SEL tests is that they've been used at scale in the last decade, so states can look forward to lower startup costs should they adopt them.

Data Education

NCLB and prodding from the Obama administration's Race to the Top program unquestionably improved the states' technical capacity to collect,

report, and analyze student assessment data. While that data generated value for some researchers and forced measurable improvements, it also provoked anger and misunderstanding. No one truly knew how schools or teachers should respond to a 35 percent proficient rating, because no one really knew what that meant; most believed that failing to meet AYP was at least as much a statistical artifact as a genuine failure. Louisiana State Superintendent John White was emphatic that data was central to all successful school leadership: "You have to start with essential facts, a set of essential facts, and tell a story with them to support a vision for change in education. The worst things will happen with ESSA if it becomes an exercise in box checking. You have to have data-driven leadership."¹² Compared with NCLB, ESSA offers states a way to provide meaningful data training, longitudinal reporting, timely reports, and curriculum links to the Common Core. All of these could improve both the acceptance of assessments as useful pedagogical tools and offer practical improvements in data use.

The challenges of state data are substantial given the transient nature of many students, especially in low-income neighborhoods. Even administrative data like attendance, gender, and poverty status is notoriously error-prone, undermining its use for either identification or instructional support. In the waning years of the Obama administration, a handful of states banded together under the banner of "data quality" to improve the situation and to build "a school culture in which data help school staff make informed decisions about instruction and school programs."¹³ The Data Quality Campaign (DQC), an organization encouraging states to boost the availability and integrity of educational data, suggested that forty states had taken substantial steps toward a robust use of data, including generating meaningful educational research, maintaining a statewide data repository, and creating progress reports for students.¹⁴ ESSA comes at a fortuitous time for testing technology. States could employ shorter but more frequent assessments as part of their school quality measures to support teachers in the classroom, and multiple firms have launched scalable tools to emphasize teacher understanding of data.

Finally, ESSA will allow states to complete what has long been a quixotic errand in education circles: tying state assessment to classroom curriculum. States have been chasing curriculum alignment without much success since the 1970s. Local districts used widely divergent curriculum and held students to a variety of graduation standards. Assessment in the 1980s and 1990s turned away from curriculum, testing students against their peers or

against state standards in the hopes that districts would align to those standards. Today, the near-universal adoption of the Common Core by school districts (even in states that never adopted it or rescinded their participation) all but guarantees that teachers will know what their students are supposed to learn. If the states incorporate those standards into their assessments, teachers could receive meaningful feedback for their daily practice. Further, transient students can all expect assessments to test consistent material in any district. Combined with better longitudinal tracking, student growth can follow these students throughout the state.

Fifty years of federal compliance, with its substantial record-keeping requirements, has provided states with a base of technical competence. ESEA built local constituencies for standards and accountability across all states in the last twenty years. The Obama administration's Race to the Top program pushed states into quick adoption of a common curriculum and better-quality assessments. Now ESSA has given states the opportunity to make them their own, but that will require leadership, not compliance.

Turnarounds

ESSA offers states a third opportunity: to remedy the overly broad approaches that NCLB took to improve lagging schools. As with alternate measures of performance and data education, some states and districts have a long history of "turnaround" efforts tied to school finance cases from which to draw lessons. NCLB also provided states with a real research base for studying school improvement. ESSA significantly increases the flexibility of turnaround funding, giving states additional resources to implement creative and research-based school improvement.

The richest source of state capacity comes from four decades of school finance litigation. After the Supreme Court washed its hands of school finances in *San Antonio Independent School District v. Rodriguez* (1973), advocates of higher funding turned to state courts to remedy unequal or inadequate funding.¹⁵ Plaintiffs rarely need to look far to find schools that fail some test of "adequate" educational performance, and they have compelled state legislatures and state bureaucracies to address low scoring or poorly funded schools. Indeed, NCLB bolstered adequacy arguments by giving advocates the federally required "proficient" benchmark. If schools could not meet these standards, they were obviously not adequate. One adequacy proponent noted that "from a litigating point of view, this stuff is dynamite and the more extreme NCLB gets, the better it is for us plaintiffs."¹⁶ In the

aftermath of these cases, a handful of states have taken measures suggesting they *could* implement turnarounds even without a judgment against them.¹⁷

Beyond school finance, states can look to ten years of turnaround research from NCLB. In that time, approximately 4 percent of all Title I money was available for federally specified turnaround efforts through School Improvement Grants (SIGs). SIG funds were meant to boost persistently low performing schools, but the money was tightly restricted. Funds could be used for closing a school (and reopening it), converting it to a charter school, firing and hiring half of a school's personnel, or one of a handful of less intense "transformative" strategies, including performance pay or longer school days. Despite these options, most states, many districts, and major interest groups thought the program straitjacketed local leaders, many of whom wanted more freedom to build school- or district-level collaboration.¹⁸ ESSA addresses this critique in a big way: not only does it boost improvement funding to 7 percent, but it removes many restrictions on what the funds can be used for. They have to be for "innovation," but the states can decide what that innovation looks like. This change will "open up possibilities for real creativity on the spending side," according to Louisiana's John White. "The funds come with enough discretion to pursue turnaround models that we know work."¹⁹

The SIG research generated decidedly mixed results. Data indicated that 54 percent of schools first funded by the SIG in 2011–2012 had gains in mathematics scores in the next academic year, but 40 percent had losses (non-SIG schools' numbers were 45 percent and 47 percent, respectively).²⁰ Nevertheless, proponents noted that for the three years of the program, SIG schools were more likely to post gains than losses and more likely to post gains than non-SIG schools. The modest gains appeared to coincide with consistent elements, including persistent monitoring of student and teacher data and local buy-in. With data and local support on hand, schools could adjust curriculum, pace, and goals.²¹ States' new assessments, especially if they are not once-a-year summative tests, could easily assist this goal.²² Looser controls on turnaround financing can give districts—through the state—incentives to enhance these activities, especially as state-centric reforms typically yield the weakest results.

Although much of ESSA only codifies flexibility already available through waivers, states now have some legislative certainty that their experimentation is not tied to a particular presidential administration. States have a formal opportunity to embark on new measures of success, improve

local communication about results, and tune turnaround efforts to local conditions.

And states and districts *have* embraced innovation.

Ironically, ESEA put state departments of education in a much better position to serve districts and schools than they would have been without federal pressure. Although federal legislation guaranteed that state departments of education would be implementing an odd mix of compliance-driven policy, creative state-level assessment, and district service, some states were able to leverage federal mandates to provide genuine innovation.

Some education departments were able to design accountability measures in a way that was attentive to teachers and parents. Some states built ties to local interest groups—in education and in the business community—to ensure that education work would have local political support. In some cases, local groups helped design state measures that then helped provide field training for schools and districts.

POLITICAL REALITIES OF STATE CAPACITY

All this suggests that state departments of education *can* provide meaningful guidance and oversight of academic quality in schools and states *if* they have politically capable leadership and the technical capacity to do so.²³ ESSA will be a boon to these states, but there are serious technical and political barriers that may impede the law's success in other states.

The first is a subtle shift from accountability to collaboration. Second, federal education policy has spent fifty years fostering compliance-driven state bureaucracies; ESSA is unlikely to shift that mentality. Third, contrary in spirit to much of the bill, ESSA imposes additional requirements to protect student subgroups and special education students. Whatever their benefits, the requirements place limits on what states can do.

Turning Around or Tuning Out?

The most striking shift is in language after a decade of exhaustion with data-driven public accountability. Instead of “accountability,” many state officials have begun to talk about “responsibility” and “collaboration.” Kansas Commissioner of Education Randy Watson was explicit: “Academic skills are important, but not to the exclusion of other things. . . . You don't do test prep for four months for one test that doesn't matter.”²⁴ Wisconsin's Tony Evers told the Senate HELP Committee that his Department of Public Instruction

would begin with “stakeholder engagement” as an “opportunity to hear multiple perspectives of what is working in the state, what needs to be changed, and how people envision flexibility in practice.”²⁵ California's school leaders took this kind of rhetoric to heart and created a dashboard of school measures instead of calculating a single measure of school quality. California State Board President Michael Kirst argued that the state should not prioritize some educational outcomes over others—that should be left to parents and other stakeholders. “If you're smart enough to look at five things on the dashboard of a car and still drive, you should be able to understand a school,” he said.²⁶ It may be simplistic or presumptuous to assign a single A to F grade to a school when stakeholders have many goals for education. An A collapses all of those goals into a single metric, but that can focus attention on the best and worst like nothing else. Multiple measures may allow teachers, staff, administrators, parents, and the public to focus on what they value most, but will also detract from the pursuit of academic performance.

While there were problems with NCLB's accountability metrics, ESSA's softer focus may replace meaningful school improvement with politically palatable local “collaborations” or simply excuse poor performance as a local problem. The danger is especially acute because 95 percent of federal turnaround funding is to be sent to school districts, precluding some bolder reforms, including charter schools and alternate school district arrangements. District-led strategies can provide excellent political buy-in from the stakeholders who are more directly affected—the teachers, support staff, and administrators. But the premise of “turnarounds” is that *these* personnel are somehow part of a school's academic problem. Thus, according to Andy Smarick, a fellow at the Thomas B. Fordham Institute, direct federal spending on districts “muddies authority, undermines good governance, and has the potential to inhibit state experimentation and slow emerging state innovations.”²⁷ If schools are to prepare students for college or career in the future, collaboration now may be a less powerful motivator for change than starkly low student performance. District-based efforts may improve the district, but they are unlikely to address core problems or try substantial innovation.

Data for Decision Making?

A second risk is states' well-honed compliance mind-set. ESSA gives states the chance to change the way they manage and service data. But years of operating in compliance mode makes it more likely that they'll treat data collection as routine paperwork rather than as an opportunity for innovation.

Despite the NCLB-induced improvements in data gathering and processing, states' focus remains thoroughly pragmatic. Kansas's Data Quality Certification is typical. The state's program was highlighted by the Thomas B. Fordham Institute as making "solid advances in [its] education data system," but its program included substantial training on data entry and navigation of state and federal databases—including which school is responsible for entering data for which students—rather than actually thinking about data.²⁸ Pennsylvania's 2015 Department of Education Data Summit acknowledged this natural rut. The conference, themed "Moving Beyond Compliance: Getting Value Out of Data," highlighted "data dashboards," "early warning systems" for dropout prevention, and database basics for retrieving data.²⁹ Both of these programs, and similar ones in other states, provide necessary fundamentals, but they indicate that states have only started to help districts analyze student data and have done little to help parents or teachers understand how psychometric exams, growth models, or cut points work. If data training found a champion in the state legislature or executive mansion, it's possible that some states would offer innovative and robust training. A handful of states did this very thing in the 1980s under performance-minded governors. But it was personality-driven and short-lived. Absent the federal lever, state departments appear unlikely to reinvent how they look at data.

Yet even if they wanted to use data—to track individual progress, for example—other hurdles would remain, especially states' current interpretation of the Family Educational Rights and Privacy Act of 1974 (FERPA). This act was meant to prevent abuses by government officials—especially school officials—in the midst of widespread government abuse in Watergate, including the use of tax records to target political opponents. (FERPA was signed by Gerald Ford twelve days after Richard Nixon resigned.) Since that time, threats to student privacy have morphed from unwitting school secretaries letting something slip to domestic and international hackers looking for identities to steal. Here, the Department of Education has undermined real opportunities for good practice. Regulations and litigation have placed the privacy of student data at the forefront of professional development for practitioners.³⁰ Thus, FERPA has replaced thinking about data with thinking about regulatory compliance. If ESSA's new measures are to yield better benefits than NCLB's simplistic school- and district-based measures, federal policy makers will have to clarify *when* student privacy is actually threatened.

Limits of Universality

A third limitation to state capacity comes from ESSA itself. Despite the act's generally open-handed flexibility, it doesn't offer relief from categorizing students by subgroup or for testing children receiving special education services, both major components of NCLB's testing regimen. These should be the easiest of the reforms from ESSA, but its continuing requirements create practical limits on how innovative states can be.

Civil rights groups and many Democrats in Congress were adamant that states' assessment measures remain in force for "all" students. Their intention was for schools to remain politically (and financially) accountable for on-standard performance even by small demographic groups of students. This was a deliberate choice. Before ESSA, the US Department of Education had allowed states to combine multiple groups of students into larger pools (as "supersubgroups"). ESSA removes that option. As a result, state accountability measures may suffer the high variability characteristic of small populations. That variability was the genesis of a major critique of NCLB—that a school or district could miss AYP purely from statistical artifacts in small subgroups. The performance measures used by the states will have to account for small-group variability carefully.

A potentially greater challenge for states lies in ESSA's treatment of special education students. A core component of federal accountability legislation is that all students in a state will be held to the same standards, regardless of their income, race, or academic ability. Some special education students fit uncomfortably in this framework because their learning trajectories may be substantially different than those of their peers. The law recognizes this and allows states to give alternative tests to students with "severe cognitive disabilities," so long as they comprise no more than 1 percent of students (about one-tenth of all special education students). Learning-disability interest groups lobbied hard to keep this cap low to prevent schools and states from sidelining special-needs children. They succeeded, and in this area, ESSA's regulations are stricter than they were under NCLB; Evers argued that the waiver requirements from these rules were even more "onerous" than NCLB's.³¹ Part of the reason is that states are held to a standard that individual districts are not. The law requires that of all students tested in a state, no more than 1 percent may be alternate assessments, yet districts may exceed the cap. Of course, for a state to meet ESSA's requirements,

every district that gives more than 1 percent of its students alternate assessments will have to be balanced by districts that give fewer alternate assessments. The states can try to mitigate this difficulty by defining what “severe cognitive disabilities” are, but after pressure from civil rights groups, ESSA regulations prohibit states from linking “disability” to school performance, English language proficiency, or (solely) behavior in school. Because “all” students must be given the same tests—even those with less severe cognitive disabilities—state policy makers will be tempted to lower standards across the board to ensure that test passage rates are politically tolerable.

Despite the seemingly open-handed nature of ESSA, these statistical mandates have made *genuine* state leadership difficult. Neither requirement is flexible, and some stakeholders see them as even less flexible than NCLB’s prescriptions. These requirements guarantee that at least part of ESSA’s implementation will be compliance-driven. Although well meaning, these requirements risk setting up the rest of ESSA for an unimaginative, cookie-cutter implementation. Without careful thought and concerted effort, these may undermine ESSA’s broader possibilities for creativity and innovation in the states.

ESSA: PROMISE OR PERIL?

ESSA’s promise is not that every state will suddenly become Massachusetts, Texas, or Delaware. Instead, it is that states will be able to borrow—or pool—measures, data education, and school improvement strategies from similar states and build on other education consortia, whether those are the assessment consortia built into Smarter Balanced, PARCC, ELPA21, ACCESS, or something else. The federal government “borrowed strength” from state capacity to implement NCLB; it relied on state bureaucracy to be able to design or purchase state tests, build up a meaningful assessment framework, and monitor schools for data quality.³² ESSA gave states the opportunity to “borrow strength” from one another. The public is ready, too. In 2015, some two-thirds of the public continued to support the NCLB-era annual testing in math and reading, and states and localities received the lion’s share of support (about 81 percent) for assessing and revitalizing low performing schools.³³ States may have great expectations for newfound freedom from federal oversight, and their success will be reflected in their cooperation.

But will they take their opportunity? Political challenges to ESSA loom large: states will no longer be able to shift blame easily, the work of schools may shift, and testing may reopen political wounds.

When states faced technical trouble or political pushback from NCLB, it was convenient and sometimes true that the fault lay with the federal government. The National Education Association made this argument consistently, and Connecticut (unsuccessfully) sued the federal government on the charge that NCLB was compelling state spending. When states came under pressure in late 2014 over Common Core-aligned tests, state lawmakers made a show over withdrawing from federally funded testing consortia.

There will be no such out with ESSA. ESSA’s potential new measures create opportunities to highlight disparities in the provision of education, and that could generate a new wave of school finance litigation. Under duress, state defendants have occasionally admitted that academic performance should be comparable within their states. When *Rose v. Council for Better Education* (1989) and a handful of school finance cases in other states suggested that states should equalize nontested factors, no one knew how to measure them. Now, defendants would be unhappy to discover that they have created a measure—classroom climate, parental involvement, or student perseverance—for courts to require equity. The push for standards, accountability, and transparency was meant to empower parents, teachers, districts, and the public to ensure that all children were served by public education. An unintended consequence of ESSA might just be that education policy is placed even more firmly in the hands of state judges. The result would not be innovation, creativity, or even local democracy; ESSA could instead paralyze state education policy.

Second, new measurements will reshape how schools do their work and create new political puzzles. Just as NCLB successfully pushed math and reading to the core of schooling and made testing student performance a prominent feature of education policy, ESSA’s codification of flexible measurement may permit schools to retool their approach to career education and work skills. These may be worthy goals, but such a shift would exacerbate a decades-long debate over tracking and schools as “sorting machines.” Critics on the left have already decried business influence in education; prioritizing work skills would only heighten their anger. Common Core was a long-term project of the business community and governors of both parties,

but it reshaped state political alliances, especially on the right. ESSA may do the same, this time potentially on the left.

Third, if lawmakers choose to add new, creative measures, they will face the risk of inflaming opponents of school testing. Although most parents do not support their students opting out of tests—*Education Next* found 67 percent support for accountability testing in schools—legislators caved to pressure in 2015 and 2016 and curtailed tests in Arizona, Indiana, Georgia, Ohio, and South Dakota, and others imposed time limits on academic testing.³⁴ Gov. Terry McAuliffe, fêted for reducing Virginia's exams related to its standards of learning, said, "I think we've heard from the students [that] these tests don't really help us prepare for college, or more importantly, they're really not helping us prepare for life."³⁵ This attitude was widespread as ESSA rolled out, suggesting that *meaningful* performance testing, only recently accomplished, could be ephemeral. "Better" assessment may be better from a psychometric perspective, but without political support, ESSA-inspired metrics are unlikely to succeed.

CONCLUSION

Still, these risks should be tempered by the larger truth that "policy" can only do so much. State leaders have to supply a vision for quality education. NCLB's theory of action was largely that measurement and reporting of academic performance would prompt teachers and schools to change. There is evidence that NCLB got states partway there. But policy makers in the states, the US Department of Education, or Congress cannot depend on policy alone to improve learning in America's schools. States can design new measures, educate teachers and parents about their students' progress, and think creatively about school turnarounds, but ESSA will not win the long game. "We have to stop fetishizing policy," John White said. "We need an authentic vision for students tailored to a mission of improvement. In my experience, policy is neither a tremendous opportunity nor a tremendous barrier. It is a system to be shaped."³⁶ Some states turned in far-reaching, innovative reforms under NCLB because key state leaders dreamed an impossible dream *in spite* of the law. ESSA has expanded the frontier of possibilities, but state leaders must grasp their freedom to take others with them.