

# Accountability Restated<sup>1</sup>

Elliot Regenstein

## Introduction

State accountability systems are the method by which states define what constitutes success in the education system. When *No Child Left Behind* (NCLB) was the law of the land, that definition was grounded in two core indicators required by the federal government: student proficiency, and student graduation rates. The passage of the *Every Student Succeeds Act* (ESSA) has added a new core indicator, student growth, and opened up the opportunity for states to add new indicators.

The federal government has been right to keep a focus on student excellence. But in doing so, both NCLB and ESSA have suffered from a significant flaw: by placing the entire emphasis of accountability on grades 3-12, federal law has shifted the focus of schools away from the years that represent the best opportunity to positively impact student outcomes. Data show that if a cohort of students is a year behind at the end of second grade, only about 15% of districts can get that cohort caught up by the end of high school. Even if every single district in the country was brought up to what is now an elite level, that still wouldn't be enough to help the kids who've already fallen behind in the early years.

These data have significant equity implications, because the kids who fall behind early tend to be low-income students and students of color. Moreover, we've doubled down on the problem by using definitions of school success in grades 3-12 that have functioned largely as proxies for wealth – rather than really looking at what kind of educational experience children are having. That problem has been addressed to some degree through the inclusion of growth measures in accountability. But that still keeps the entire focus of accountability on the tested years, third grade and up.

An accountability system that focuses solely on third grade and up has no realistic hope of achieving excellence. Only by expanding the definition of educational success to include the early years can the system reach its full potential. This proposal builds on the existing system by proposing a new core indicator: school inspections, modeled on a practice already common in Europe – and in United States early childhood programs. These inspections not only provide useful information to schools in the school improvement process, they also represent an indicator that can be used across all age spans.

Accordingly, I propose an accountability system that reflects the following values:

- *Improving student outcomes.* A core goal of the education system is for students to graduation from high school ready to be successful in whatever they do; accordingly, this value should remain at the heart of any accountability system. Indeed, this value is already reflected in existing accountability systems through the use of standardized test scores and graduation rates.
- *Improving student growth.* Students' rate of progress is critical to their ultimate outcomes, and student growth is the best method we have for measuring the impact of schools on that growth. Its addition to federal accountability is welcomed. But it is important for states to treat it as a separate value from student outcomes, and not simply mash the two together into a single composite score.

---

<sup>1</sup> This proposal is based on Part I of [Education Restated](#), my 2022 book published by Rowman & Littlefield. Some data and figures are quoted from that book verbatim.

The true value of growth data comes from its interplay with proficiency data: how do the two data points together allow states to identify the schools where external support is most needed?

- *Improving educator processes.* The use of an inspectorate provides something test scores do not: useful feedback on what a school needs to do to improve. Moreover, it allows accountability systems to place a meaningful focus on the years prior to third grade – which represent the best opportunity for the system to dramatically improve performance.

Accountability systems have had a powerful impact on district and school decision-making; they are an imperfect tool, but an essential one. Accountability systems create incentives that drive district and school behavior – and while districts and schools will react differently to those incentives, as a policy matter states should create incentives that reflect the right values, and collectively work in harmony. This proposal describes an accountability system that would help districts and schools focus on the years – and practices -- that matter most, while improving performance throughout the education system.

### *Key Principles*

The indicators in an accountability system should reflect the values behind that system. The values embedded in NCLB were important ones: student achievement and high school completion. Those values also complement each other. Emphasizing only achievement would create an incentive for schools to push out students who were not achieving at high levels, an incentive counteracted by measuring completion. On the flip side, measuring only completion would create an incentive for schools to water down the rigor of a high school diploma, which is balanced by measuring achievement.

But while those indicators are limited. Numerous educators and scholars have pointed out that the test scores used to measure student achievement are heavily correlated with students’ family income. This meant that while those scores might be useful metrics of how *students* are doing, they are not necessarily an accurate measure of how *schools* are doing. The ESSA corrected for that by requiring states to include student growth measures, which show how much progress students are making; this change was properly seen as providing a more accurate measure of school quality.

At the very least, districts and schools should be expected to produce a year’s worth of growth every year. In fact, most districts fall short:

<b>Annual Rate of Growth</b>	<b>Percentage of Districts</b>
<b>1.1 years or more</b>	15%
<b>1.0-1.09 years</b>	25%
<b>.95-.99 years</b>	13%
<b>Less than .95 years</b>	47%

These results have appropriately sparked conversation about the importance of improving district performance, so that a higher percentage of students are making at least 1.0 years of progress per year. But what they also show is that in 85% of districts students are making less than 1.1 years of progress per year. That means that even if every single district in the country achieved what is now the 85<sup>th</sup> percentile of growth – which would be an extraordinary accomplishment – a cohort of students a year behind at the end of second grade would not be caught up by the end of high school. Accordingly, focusing solely on third grade and up is not likely to be sufficient to help children reach proficiency by the end of high school.

Improvement in the early years also may be easier to achieve than improvement in later years. Students in K-12 are largely already being educated, and the challenge now is to improve the quality of that education – which, for years, has proven to be very difficult. But according to the National Institute for Early Education Research, nationally 61% of 4-year-olds and 86% of 3-year-olds are not enrolled in any publicly-funded pre-kindergarten education. Providing high-quality educational services to those children might have a meaningful impact on their long-term trajectory.

We do not currently have good data on how many children are in fact a year behind at the end of second grade. NAEP data tells us that in 4<sup>th</sup> grade, 37% of children are reading at “Below Basic” levels, and 25% of children scored “Below Basic” in math. While the correlations between NAEP and state proficiency vary from state to state -- and it is important not to draw inappropriate conclusions from NAEP data -- the NAEP data are consistent with a hypothesis that there is a critical mass of students who are significantly behind in elementary school. This too should create a sense of urgency about addressing the early years.

NAEP data have shown meaningful differences among children of different races; the percentage of students who are “Below Basic” in both reading and math is higher for students who are Black and Hispanic than for students who are White and Asian. This tracks with historical state-level data on student proficiency. But data on district and school growth is much less correlated with race and income than data on proficiency; schools can help students advance regardless of the population they serve. Low proficiency rates among Black and Hispanic students should create an action imperative for equity-focused policymakers to address improvement in the years prior to third grade.

Having identified addressing the early years as a key value, it is then necessary to develop an indicator that can create an incentive to focus on those years – and that can fit appropriately into a state accountability system. This has proven to be a struggle for states, particularly within the limits of ESSA:

- There is very little appetite for imposing standardized tests on the K-2 years. Indeed, one of the major criticisms of standardized tests has been that they have caused a narrowing of focus in schools: if reading and math are the only things being measured, those are the things that will be prioritized. K-2 instruction is already potentially sub-standard in many places, with ample research showing that principals have a limited understanding of child development in those years – and that teachers may not be using best practices for instruction. Advocates for developmentally-appropriate instruction are rightly concerned that imposing a standardized testing regime on the K-2 years could exacerbate existing problems with instructional quality.
- ESSA has allowed the use of additional indicators, and some of those – such as chronic absenteeism – could be used to focus on the K-2 years. In some instances states could take the additional indicators they are already using and disaggregate data by grade span to emphasize those early grades; they could potentially even overweight those years in the accountability formula. A change along those lines would likely have an incremental positive impact, but would not necessarily drive a transformative change in focus on the earlier years.

In a next generation of accountability, federal law could allow for a new kind of indicator: an inspectorate, in which trained evaluators come into schools and measure the quality of the school’s processes. Because it focuses on process rather than outcomes, an inspectorate can gauge quality at all

age levels. Indeed, inspectorate-style accountability is increasingly common in early childhood programs; the federal Head Start program uses it, and some states have begun to do so as well.

Of course, inspectorates serve another value as well: providing actionable feedback to schools, in a way that test scores do not. Test scores provide useful information about how schools are doing, but no guidance on what it might take to get better. Given that in 47% of districts students are making less than .95 years worth of progress every year, that guidance is sorely needed. An inspectorate could fill that gap in an accountability system.

*Implementation Strategy*

A. The Relationship Among Indicators

While proficiency, growth, completion, and process measures all represent important elements of education accountability, it will be important for states to use them together effectively to drive improvement. In too many states multiple elements are combined into opaque formulas that end up masking the impact of any one element. To really drive improvement, states should use outcome and growth measures to sort schools as follows:

	<b>Strong Student Growth</b>	<b>Weak Student Growth</b>
<b>High Student Performance (Proficiency/Completion)</b>	Students are thriving, and schools are helping them do so. Lessons from these schools should be shared with other schools.	Students are doing well, but the situation should be monitored to make sure they do not fall off track. If growth trends are poor enough support for the school may be needed.
<b>Low Student Performance (Proficiency/Completion)</b>	Schools are doing their job, but the students need more help. Community-level interventions – included expanded early childhood services – may be the most promising strategy.	These situations are the most urgent. Both community-level and school-level interventions are likely to be needed.

Realistically, it will be difficult to develop an inspectorate that can serve all of the schools in a state; that might take years, and the expense would be significant. But federal accountability systems could allow states to try inspectorates in a targeted way, focusing on districts and schools where student growth is weakest. The inspectorates could then identify areas where the schools could improve their practices, and states could provide supports targeted to those needs. States will need to do a cost-benefit analysis on where the resources of an inspectorate can have the greatest impact, using assessment data to help determine where the benefits would be most significant.

As the table states, in some instances the needed changes will have to go beyond the school itself to include the early childhood offerings in a community. There are schools where proficiency is low, but students are making at least 1.0 years of progress every year; in those schools, the problem is that children came in too far behind to catch up. In those communities, the right strategy for improvement will be supporting improved kindergarten readiness. Accountability systems should be designed to look for those kinds of results. In schools with both low proficiency and low growth, increased investment in early childhood will likely need to be coupled with changes to the practices in K-12 schools.

In sum, test scores and graduation rates should be used to do an initial sorting of schools, and then an inspectorate can be used to get under the proverbial hood in places where more information is needed. Federal law should continue to mandate the use of proficiency, growth, and completion measures, with states encouraged to use them to sort schools as described above. States should then be authorized to use other measurement tools to diagnose existing processes – and needs – at the local level. It is important to have the inspectorate as part of the accountability formula rather than only the improvement process, so that the accountability system is properly identifying the districts and schools that would benefit most from resources to support improvement – even if those resources may be used primarily to improve school readiness, rather than the schools themselves.

Current law requires all indicators to be “valid, reliable, comparable, and statewide.” The federal government should consider giving states a waiver of this requirement for an inspectorate. It would be excessive to inspect every school every year; indeed, states might well want to use assessment data to identify districts and schools where an inspection would likely be most impactful. States could also require each school to be administered at least once in a period of multiple years to ensure that they can all benefit from the insights of an inspection, even if their assessment data is promising.

Indeed, a key goal of this approach is to be more thoughtful about school improvement. This approach will identify schools where unsatisfactory student outcomes are a product of inadequate educator practices, and will even identify specific practices that need to change to benefit student outcomes. States can then determine the best way to support the implementation of those practices. Historically accountability systems have identified schools for improvement but have been very weak when it comes to diagnosing what they need; because of that, there has often been a mismatch between diseases and cures. An inspectorate provides the kind of information needed to craft more thoughtful improvement strategies.

In that same vein, inspectorates will identify schools where unsatisfactory student outcomes are a product of factors outside the school’s control – specifically, where low kindergarten readiness is translating into disappointing high school outcomes. Having that information will allow states and communities to address the challenges at their root, rather than trying to play catch-up when students are already behind. This will make the improvement process even more likely to succeed.

## B. Building an Inspectorate

The development of an inspectorate not only provides a more holistic examination of how a school contributes to student achievement, it offers an opportunity for stakeholder engagement. States could develop more than one rubric for measuring school quality, recognizing that there may be different kinds of schools that should not necessarily all be evaluated in the same way. Local education leaders should be actively involved in the creation of a rubric or rubrics for the inspectorate to use; not only will that strengthen the quality of the rubrics, it will increase their credibility with districts and schools.

Federal support will likely also be needed for states that want to build out an inspectorate, which is a labor-intensive enterprise. The inspectors need not be employees of state government; this may be an area where public-private partnerships and multi-state collaboratives will have the flexibility to attract needed talent while meeting state needs. As with many startups, success will be most likely if the inspectorate starts small and develops good practices before expanding.

## *Analysis*

This section summarizes the potential impact of the proposal, and how it reflects the values of the Chamber's Future of Data Working Group.

- A. A description of the problem(s) with the current assessment and accountability paradigm that the Proposal is trying to address

As described above, the largest problem with the current assessment and accountability paradigm is that it drives educator focus toward the years from third through twelfth grade – and while those years are important, they are not enough time to achieve our national goals for high school outcomes. Indeed, they represent only half of the time children spend from birth through high school graduation. Moreover, child development experts would emphasize that earlier interventions can have the strongest impact on long-term child trajectories.

The data is clear: the best chance to dramatically improve student outcomes is by focusing on the earlier years. And the opportunity for improvement is real: our infrastructure in the years prior to third grade is much weaker than our infrastructure in the later years. The accountability movement has sparked meaningful change in the years from third grade through the end of high school, and now it is time to build on those changes by using the power of accountability systems to address the earlier years.

While that is the primary goal of this proposal, it does also address a key secondary goal: how to provide better information to educators to improve their performance at all ages. Indeed, this is the most significant shortcoming of the current accountability regime when it comes to grades 3-12. An inspectorate would address that issue, and provide key supports – particularly for the 47% of school districts where student growth is currently lowest.

- B. Objectives for policy change (i.e., at a high level, what requirements, guidance, and/or approaches to assessment and accountability will change?)

At the federal level, the ESSA's framework would need both a tweak and a major change. The tweak would be to add precision about how proficiency data and growth data should be used by states, so that their combined impact can be most powerful. The major change would be to allow states to use an inspectorate as part of their accountability system – potentially for only a subset of districts and schools. Ideally, the federal government would also provide some seed funding to help states launch inspectorates, individually or in combination.

It is important to acknowledge that this proposal is not designed to change the methods states currently use to conduct assessments. Assessment practices are extremely important; more precise measurement would benefit policymakers, educators, students, and others. This proposal focuses on the levers most likely to drive state and local behavior with regard to improving education prior to third grade, which is why it does not address changes to summative assessments. But this proposal could easily be harmonized or synthesized with reforms in at least two major areas:

- Improvements in how summative assessment is conducted in grades 3-12. Any leap forward in measuring proficiency and growth would have a positive impact on the framework proposed here.
- Improvements in formative assessments for all grades, particularly second grade and earlier. A renewed focus on the early years increases the importance of understanding child progress in those

years, and formative assessments can play a key role in that process. In fact, many districts with the resources to do so have rich assessment programs in the K-2 years; an increased focus on the processes of those years would benefit from stronger assessment practices.

Thus, while this proposal is primarily about accountability levers, it could easily be complemented by thoughtful changes to state assessment systems.

C. Desired outcome of the policy change (i.e., what does the Entrant hope will happen if the described policy objectives are achieved?)

The desired outcome of this policy change is that policymakers and local school officials will increase their focus on the years prior to third grade. We already have enough data to know that those years are critical. But state accountability systems send a clear message to school district leaders: focus on the later years, because that's what we're measuring. Mitigating that measure by including the early years in federal accountability is an important outcome of the proposed policy change.

School leaders tend to have backgrounds primarily in educating older children, and repeated studies have shown that even elementary school principals have very limited understanding of child development. So school leaders may already be inclined to focus on the later years, and accountability systems only reinforce that tendency. If the primary pressure schools are under is to improve test scores, early childhood investment is too much of a long-term strategy; by definition, a preschooler will not take an accountability test for another four years – and considering the rate of turnover among superintendents, they may reasonably believe they can't wait that long to show an impact.

Moreover, in high-mobility communities, as many as half of kindergartners may leave the district before third grade – meaning that the district might reasonably choose to focus its resources on children it knows will be included in its accountability metrics. Many of those children may continue to be educated in the same state, but not in the same district. That means the state should have a strong interest in creating incentives for districts to pay more attention to these young children.

Indeed, funding for early childhood services has historically been much more state-driven than local-driven. While the ratio of state to local funding varies from state to state, nationally state and local spending are relatively even. By contrast, early childhood funding comes primarily from state government: state pre-k programs generally support the entire cost of providing an education, perhaps with a small local match.

To date state governments have taken the primary responsibility for educating young children. This proposal is not designed to undercut states' leadership role in early childhood – it is instead designed to leverage that role to create greater incentives for school districts to step up. State accountability systems are one of the primary strategies for shaping district behavior, which is why it is important that they reflect the value of the pre-third-grade years.

D. Description of how the policy change will result in students who are better prepared for college, careers, and civic participation than in the current education system, with a focus on students who the system has historically underserved

While accountability systems have already resulted in students who are better prepared for college, careers, and civic participation, any system that focuses only on the years after third grade will keep the

ceiling of its impact artificially low. Historical data on student growth shows that existing achievement gaps for historically underserved students simply cannot be closed between third and twelfth grade. The only way to truly change the dynamic for underserved students is to focus on the earlier years, and shrink the third grade reading gap before it opens.

Indeed, the term “underserved” is never more powerful than when applied to the preschool years. In K-12, “underserved” refers to students attending schools that are not meeting their needs. In early childhood, “underserved” refers to students who may be receiving no educational services whatsoever.

E. An explanation of known tradeoffs, and potential negative consequences, as well as rationale for the approach selected

i. The Role of Schools and the Role of Not Schools

Schools are essential to student success. Critics of accountability systems sometimes point to all of the non-school factors that impact student outcomes. Those factors are real, but should not be used to avoid school accountability. Given the level of public investment in schools, it is entirely fair to measure the impact they are having – particularly the progress they are able to help students make.

But while it is reasonable to expect schools to provide children with a year’s worth of education every year, it is not reasonable to expect them to play catchup all the time. Given the many studies showing that third grade test scores are highly predictive of 12<sup>th</sup> grade test scores, it is important to create a greater focus on the years prior to third grade – and indeed, the years prior to kindergarten, which is the time when historically underserved students are actually losing the most ground.

It is appropriate for education accountability systems to measure the impact of schools on student performance. But it is also important for them to acknowledge that sometimes the key to improving student performance isn’t necessarily a school-based strategy. This proposal is meant to help isolate situations where the real key to improvement is investment in early childhood, and to set the table for states and districts to then make that investment.

ii. External and Internal Accountability

There is a tension at the heart of any external accountability system. External accountability is premised on a lack of trust: the federal government does not trust that states and school districts, left to their own devices, will provide historically underserved students with the education they deserve. That premise is a sound one – indeed, the very existence of “historically underserved students” justifies a lack of trust from the federal level. That is the justification for the federal government’s role.

But the best schools will be ones that operate on trust; educators, parents, and students will all trust in each other, and the school will succeed accordingly. There are schools operating successfully on trust all over the country, serving children of every demographic. Those schools will have what is sometimes called internal accountability: the key parties in the school feel accountable to each other, and behave accordingly.

That sets up the tension of accountability that will never go away: it is very hard to use systems based on a lack of trust to develop systems based on trust. That problem cannot be fully solved, but it can be mitigated. The current accountability system has bred distrust by focusing solely on assessment



measures, and during NCLB in particular emphasizing consequences that were seen as punitive. This approach would mitigate that distrust in important ways:

- Actionable Information: An inspectorate would provide useful information to districts and schools, and allow for a more interactive experience during the accountability process.
- Transparency: Developing rubrics for an inspectorate based on local educator input would help district and school leaders feel like they were being judged on criteria over which they had some control, rather than criteria imposed unilaterally from above.
- Solutions Orientation: Emphasizing solutions that do not imply the incompetence of educators – particularly expanded early childhood offerings – would show faith in the work of schools, and acknowledge factors outside their control. This would be particularly appropriate in schools showing strong growth even though their proficiency is low, a category of schools that historically has been subject to adverse consequences despite actually doing a good job with students.

These factors would help make the accountability system more of a trust-building exercise, even while preserving the important role of the federal government and states as a source of distrust for potentially inadequate local efforts.

### iii. Comprehensiveness and Simplicity

An accountability system with only one factor will skew behavior in potentially counter-productive ways. That is the heart of Campbell's Law, which observes that the more any metric is used for decision-making, the more it becomes subject to corruption pressures. As discussed above, if the only metric of school success was graduation rates, schools would have a strong incentive to water down their graduation requirements; if their only metric was proficiency, they would have a strong incentive to make low-achieving kids disappear from their rolls. Multiple metrics are needed to create the right balance to drive the right behaviors.

But having too many factors can create a jumble. At some point, measuring everything becomes measuring nothing; if too many factors are included in an accountability system, it does not really incentivize behavior, nor does it provide clear information about how schools are doing. Indeed, sometimes states quasi-intentionally create accountability measures that are difficult to understand, in order to avoid dealing with the consequences of low performance.

This proposal is meant to sit between those two extremes. It includes enough factors to provide a comprehensive picture of how districts and schools are doing, but not so many that it cannot be easily understood by a lay audience. That tradeoff is an important one in any accountability system, and this proposal is designed to strike a workable balance.

### iv. Principle and Reality

In principle, an inspectorate should provide useful information that a district or school can act on. In principle, early childhood investment should help improve kindergarten readiness, creating a ripple effect throughout the K-12 system. There is ample evidence to support both of those propositions, which are critical to the success of this accountability framework. But there is also enough evidence and experience to make clear that not every state, district, or school will experience the best-case scenario of this approach.

Thus, two of the most obvious potential negative consequences of this proposal are these: (1) States will build an inspectorate that costs a fair bit of money, but does not actually provide insights that anybody is acting on; and (2) States will identify the need for greater early childhood investment, and then either not make it – or make it in a way that does not actually address the educational needs of young children. Both of those potential pitfalls are readily avoidable; particularly with regard to early childhood investment, there are state and local programs that have demonstrated meaningful impacts, and could be used as models for future investment. But that is an opportunity, not a guarantee.

- F. An explanation of how the proposal adheres to the Chamber Foundation Future of Data Working Group's foundational principles.

*All students can learn, and all deserve to be held to high expectations and deserve to be served by a system that can meet their personal needs:* This proposal retains a focus on high expectations for high school graduates, and creates incentives for a focus on early education that is more likely to help children get on a trajectory toward success.

*Information regarding school and district inputs and student educational outcomes is critical for students, teachers, parents, and policymakers:* Under this proposal accountability systems would continue to produce meaningful data about student educational outcomes, and would in fact produce vastly more information about school and district inputs.

*A federal role is critical for establishing guardrails to ensure that all students can access a high-quality education:* This proposal would preserve the critical role of the federal government in establishing guardrails for a high-quality education for all students.

*Education funding shall strive for quality and equity, meaning students who the system has historically underserved shall have the most resources applied to their educational experience:* This proposal would create a strong focus on quality and equity. Arguably the most inequitable part of the education system is the early learning system, where many historically underserved children do not have access to any publicly-funded education at all. This proposal would seek to drive resources toward those years.

*Information and transparency into the K-12 system's performance shall expand access to high-quality learning opportunities. It shall allow resources to be targeted to improving low performing schools and also allow students and families to select other public options if their school is not meeting their student's needs or interests:* This proposal would help states be more precise in targeting resources to low performing schools – and to communities with high performing schools (as measured by growth) that need early childhood investments to achieve greater proficiency. Not only would it help to target those resources, it would provide meaningful guidance on how to use resources most effectively. It would also encourage investment in the early childhood system, which offers significant opportunity for parent choice – particularly given that many early childhood providers do not restrict enrollment to children who live in a particular school district.