

Seizing a Rare Opportunity: Design Considerations for Accountability under ESSA

Richard J. Wenning

January 24, 2016

INTRODUCTION

ESSA provides a vital and timely opportunity to recast educational accountability, repair and build trust, and generate the public will necessary to embrace a hopeful, modernized vision of public education and its purpose. This opportunity is particularly auspicious for school leaders, superintendents, and commissioners of education now beginning their tenure.

The overarching goal of the accountability system proposed in this paper is dramatic improvement in student outcomes and in closing performance and opportunity gaps. This requires that we align the design of state and local accountability systems with standards-based student progression and the prominence of student-centered, competency-based learning. Importantly, this opportunity allows us to prioritize development of a robust evidence base useful for formative and summative purposes; and separately, to tackle ratings, consequences, and other summative determinations. This proposal chooses to deliberately detach the first task from political cycles impacting the second, which tend to undermine fidelity of implementation and opportunities for innovation.

ESSA opens the door to a far more inclusive accountability system centered on benefiting students, families, and educators. Such a system must clearly and unequivocally reveal and build a common understanding of the inequities in opportunity and outcomes that constitute the gaps we must close to deliver on the American promise.

The shift to personalizing learning, progression based on demonstrated competency rather than seat time, student and family ownership of learning experiences, and alignment of in-school and out-of-school learning all require a balanced body of evidence reflecting the academic, social-emotional, and college and career planning progress and attainment of individual students.

How might that body of evidence become standards-based and credentialed in ways that employers, higher education institutions, and accountability systems recognize it as demonstrating student readiness for key transitions? How might that demonstration of readiness be expressed in a digital portfolio owned by each student and family, rather than on pieces of paper in file cabinets or in student information systems owned by school districts and states? The design objectives of the system proposed here address these questions.

A next-generation accountability system should supply information that helps marshal a consensus for change in our prevailing, antiquated educational delivery model and its allocation of resources. Dramatic improvement requires stakeholder trust. Thus, the design of the accountability system should inspire trust, offer needed autonomy to create new tools and approaches, and provide fair and transparent reporting of student outcomes using four layers of evidence, with comparable statewide measures as the first layer.

DESIGN OBJECTIVES

Design Objectives that Promote Trust and Learning while Providing Incentives for Innovation and Modernization

1. Student-centered accountability focused on key transitions and college and career readiness.

Design Objective #1: true student-centered accountability based on a balanced body of evidence to support competency-based learning and progression. The body of evidence also will yield a holistic and broadly understood view of school quality for public accountability purposes.

Key vehicle and essential design task: develop personalized individual student learning plans that eventually take the form of digital portfolios, incorporating four layers of learning evidence, spanning in-school and out-of-school settings, across key performance indicators (KPI).¹

The common purpose of public education: ensuring the adequacy of each individual student's progress toward college and career readiness, and ensuring opportunity for learning commensurate with that expectation. The individual student's digital portfolio becomes the comprehensive exit credential and key entry credential for colleges and employers. Each state will need to define college and career readiness (CCR) in partnership with its higher education system and business community. While individual state nuances will be necessary, the definition should include common anchor measures comprising the first layer of evidence described in Design Objective #2.

2. Balanced, useful, and engaging body of evidence developed through an inclusive design and implementation process to build ownership, insight, and will.

Design Objective #2: provide incentives for schools, districts, and states to develop the body of evidence needed to support full implementation of state content and performance standards at the student level. To be successful, the process for developing the body of evidence must build ownership, insight, and educator and public will to support change.

Key vehicle and essential design task: provide opportunities, incentives, and support to develop robust bodies of evidence that are useful for both formative improvement and external evaluation purposes, including annual ratings of school (and district) performance. Accomplishing both purposes effectively will require substantive stakeholder participation in a series of phased design processes.

These design processes will unfold over time, thus requiring greater reliance on statewide standardized evidence initially and then greater reliance on locally designed evidence as it emerges. Because there are not, nor should there be, a statewide or local standardized assessment for every desired standards-based competency, the proposed system develops four layers of evidence:

- a) Statewide standardized assessments
- b) Local standardized assessments

¹ The hierarchy of terms "indicators, measures, metrics, and targets" is used deliberately to differentiate the function and purpose of each component. See "A Framework for Academic Quality: A Report from the Consensus Panel on Charter School Academic Quality," National Alliance for Public Charter Schools, Colorado League of Charter Schools, National Association of Charter School Authorizers, CREDO at Stanford University, June 2008.

- c) Local assessment of discrete competencies demonstrated by student work, projects, and performances
- d) Educator determinations regarding the extent to which the three layers above demonstrate a student's progress and readiness for key transitions. This promotes educator professionalism, evidence triangulation, and collaborative examination of student work.

3. Transparent, engaging reporting that promotes public learning and will for change.

Design Objective #3: state and local development of high-quality public reporting systems featuring engaging visualizations of comparable evidence across each performance indicator and disaggregated to identify, diagnose, and reveal subgroup gaps. Data visualizations will be easily shared through digital media channels and use a common lexicon of plain language for students, educators, parents, and policy makers to promote shared understanding.

Key vehicle and essential design task: develop transparent, engaging systems that report student outcomes and learning opportunities, and because of their quality and support, are robust enough to survive the politics of consequences: cycles of debate over what "counts" for weightings, ratings, and stakes.

PROPOSED ACCOUNTABILITY SYSTEM

The proposed accountability system includes three key performance indicators and a fourth "indicator cluster." Each indicator and corresponding measure will be disaggregated by subgroup to reveal gaps. While each measure, within each layer of evidence, will be disaggregated at least at the local level, not all measures will be appropriate for statewide or local reporting or to establish school ratings.

Public reporting of outcomes (Design Objective 3) using statewide measures will continue while the longer-term design process described in Design Objectives 1 and 2 run concurrently. As schools, districts, CMOs, and states build their capacity to broaden the body of evidence, the centrality of statewide standardized evidence will diminish over time for annual determinations. To reward and encourage innovation, states should offer waivers from prevailing accountability systems to districts or school networks. And with support from the philanthropic sector and business community, states and districts should promote design processes (perhaps competitions) to fuel public/private R&D, and to develop and disseminate the attributes of emerging bodies of evidence.

Key Performance Indicators

1. **Academic attainment** (also known as achievement, proficiency, status) in all core subjects measured within each layer of evidence appropriate for the indicator. Annual statewide assessments of reading, mathematics, other core subjects, and college readiness comprise the first layer of evidence reported and employed for school ratings. The corresponding metric is the percentage of students reaching each performance level, with the highest level receiving the most value. What constitutes proficiency at each key transition must be aligned with developmentally appropriate college readiness content and performance standards.

Proficiency is a destination, a mile marker of attainment. It is one of several important indicators of "performance," but it is a lagging indicator. It does not in itself imply academic effectiveness. Note that the proposed system does not conflate "performance" and "proficiency." If an accountability system places greatest weight on proficiency, then it creates an incentive for schools to maximize their students' starting points through selectivity and to focus on kids on the cusp of proficiency—instead of focusing on all students and meeting their individual needs. This was a fundamental design flaw of NCLB's AYP measure and some states' ESEA waivers.

Therefore, to focus incentives on maximizing each individual student's progress toward CCR, the proposed system will weight academic attainment and its gaps far less than the leading indicator of academic growth at the elementary level. The weighting of academic attainment may increase at the secondary level, as time before graduation grows shorter. While a highly rated school may have subgroup gaps in attainment (simply because of different starting points), it cannot have subgroup gaps in growth.

- 2. Academic growth and its adequacy** (also known as normative and criterion-referenced growth, velocity, value-added, speed) in all core subjects measured within each layer of evidence appropriate for the indicator. Student Growth Percentiles² based on annual statewide assessments of reading, mathematics, other core subjects, and CCR selected by the state comprise the first layer of evidence reported and employed for school ratings. The corresponding metrics are Median Growth Percentiles, with 50th percentile growth reflecting the normative concept of a year's growth in a year's time; and Adequate Growth Percentiles, which provide a student-level growth target constituting "good enough" growth, and which yield the percentage of students on track to proficiency or on track to CCR.

Disaggregation and high weighting of growth and its gaps are essential because too often, poverty and growth are negatively correlated. The longitudinal (normative) growth of students from where they start is an essential indicator of performance and academic effectiveness. If an accountability system places greatest weight on growth, it creates an incentive to maximize the rate and amount of learning for all students and supports an ethos of effort and improvement.

Students who start behind need to grow faster to catch up than students who start proficient. There is no other way to close achievement gaps. Because the best sustained growth rates observed to date are insufficient to allow catching up by the vast majority of students who start behind for any number of reasons, we must allow more time for students who need it to catch up.

The adequacy of growth (growth to standard) is an important indicator of performance. It is highly correlated with student starting points, and thus not a good measure of effectiveness at the educator or school level. However, it is a very useful measure of the effectiveness of a system of schools at a state, district, CMO, or feeder-pattern level.

² The Student Growth Percentile model, originally known as the Colorado Growth Model, is the most widely used statewide growth model. It yields a normative and criterion-referenced growth percentile and is capable of measuring growth across different assessments. It was developed by the Colorado Department of Education in partnership with the National Center for Improving Educational Assessment and is available to the public on GitHub under a Creative Commons license.

3. **Indicator of progress toward English language proficiency.** To measure progress in achieving English language proficiency, the proposed system employs a statewide English language assessment capable of measuring longitudinal growth and attainment, supplemented by local evidence and educator determinations.

4. **Indicators of student success and school quality.** The proposed system features this component and prefers that these indicators be of low or no stakes beyond public reporting, given the early stage of the development of a number of applicable corresponding measures. Within this cluster of indicators lie the design and development of individual student digital portfolios that contain the evidence and credentials belonging to the student and provide views relevant for students, families, educators, colleges, and employers. The variety of measures below also will promote a public conversation about whether schools, districts, and states are ensuring opportunities for learning, both in and out of school, commensurate with the expectation of all students graduating CCR.
 - a. **Indicator of graduation readiness**
 - Graduation rates (using the common, four-year adjusted cohort graduation rate formula), with equal “credit” for 4, 5, 6, or 7-year rates. Rationale: create incentives to graduate only students who are ready; to welcome back dropouts; and to expand learning time.
 - Average composite college readiness assessment score and percentage of students reaching the college-ready cut score.
 - College enrollment, persistence, and completion.
 - Progress toward and attainment of college and career planning competencies defined at the school and district/CMO level with state support.
 - Progress toward and attainment of social-emotional competencies defined at the school and district/CMO level with state support.
 - Educator determinations of the extent to which an individual student is on track before each key transition, considering the complete body of evidence across all indicators.
 - The percentage of students with high-quality plans for their future and portfolios of evidence across all indicators demonstrating their readiness for college and skilled employment. Implicit is the need to define plan quality—a stakeholder design challenge.
 - b. **Indicator of student opportunity to learn**
 - Access to and completion of advanced coursework in each core subject, including music and the arts.
 - c. **Indicator of educator opportunity to learn and perform**
 - Statewide surveys of school climate, working conditions, and professional culture.
 - Measures of professional development opportunities.
 - Support for attaining graduate degrees or professional advancement.

Calculating Summative School Ratings for Each Indicator and Making Annual Determinations

Rating categories and their transparency. The proposed system would employ descriptive designations using the language of standards, not letter grades (as use of the latter degrades trust among stakeholders and their will to collaborate). The designations of Does Not Meet, Approaching, Meets, and Exceeds will be used to describe the level of performance on each indicator, measure, and metric. By

using a transparent index system with a rubric to assign points earned, summary determinations at the indicator level will be easily traced back to the outcomes reflected by their component measures and metrics.

Index scoring for each indicator. The index scoring system will be used to determine which schools are among the lowest-performing in the state and thus subject to “comprehensive support and improvement.” States should not be required to combine ratings at the performance indicator level into a single summative rating. Rather, the U.S. Department of Education should allow each state’s design and political processes to reach that conclusion. However, the system proposed here would not produce a single rating across indicators because a single rating combining so many measures would fail to promote public understanding, mask important strengths and weaknesses, waste political capital, and add unnecessary abstraction.

Weighting and subgroups. To meaningfully differentiate annual determinations of school quality, weighting will meet ESSA’s requirement that each of the four indicators count for a “substantial” weight, and that the first three “in the aggregate” be afforded “much greater weight” than the fourth. In addition:

- Each indicator will be **disaggregated by subgroup**, and the prominent weighting of growth gaps by subgroup will differentiate any school with low-performing subgroups.
- **Progress toward English proficiency** would count as much for a school with only a small number of English language learners as it would for a school with many. All means all.
- **Longitudinal student growth** would count as much at high-attainment schools as at low-attainment ones because the Student Growth Percentile Model calculates a full distribution of SGPs at each attainment level. However, a rubric accommodation can be established to give credit for meeting adequate growth with a lower normative growth threshold for high-attaining students.

Incentives to meet design objectives. To provide incentives for local development of the three additional layers of desired evidence, the proposed system would employ a “Request for Reconsideration” process like in Colorado, where schools and districts are invited to submit local evidence to challenge proposed state ratings of school quality. Such evidence will be evaluated by an expert panel and conclusions made public unless the superintendent or school leader chooses to withdraw the request following the expert panel’s feedback. Regardless of the outcome of the reconsideration process, consistent statewide evidence on the school’s annual results report (school performance framework, school report card) will still be disclosed, only augmented with the adjusted rating and its basis, if applicable.

The proposed system also will provide an annual determination for each indicator based on both one year of evidence and three consecutive years of evidence, and then assign the final ratings on whichever view is most favorable to the school. This creates an incentive to take a longer view, avoiding unproductive quick fixes like test prep, while also allowing a new leader’s school quality to be evaluated based on his or her shorter tenure.

Recommendations for the U.S. Department of Education

The U.S. Department of Education (ED) has key roles to play in the proposed accountability system (and analogously, so do State Education Agencies):

1. Reward innovative states and districts with flexibility and autonomy regarding the use of stakes and consequences to encourage and support stakeholder design cycles (Design Objective 2). This flexibility should not extend to the transparent and neutral reporting of statewide evidence (Design Objective 3).
2. Strengthen reporting and dissemination at a national level (Design Objective 3). For example, ED should report the percentage of students on track and attaining CCR by state and disaggregate this information in all relevant ways, both through engaging visualizations and ready public access to datasets suitable for secondary analysis.
3. Establish an ambitious national goal, such as leading the world in the percentage of students completing four-year college degrees.
4. Backward-map the performance improvement each state requires to achieve our national goal for all students and regularly diagnose state performance weaknesses, feeding that information back to states as part of support and improvement processes.
5. Provide differentiated support to states, similar to what states must do for districts and schools.
6. Influence state strategies by disseminating leading approaches and outcomes, and by providing direct and indirect support. Do less prescription and approval, and more disclosure and dissemination using modern digital channels.

Richard J. Wenning is executive director of BeFoundation, a family foundation devoted to dramatic improvement in the education of underserved children in Connecticut and the vitality of their communities. He also cofounded and directs SpreadMusicNow, a fund that supports music education for underserved youth, helping to shape their futures and build lifelong success. Richard has worked with federal, state, and local agencies and nonprofits for nearly 30 years. His role at BeFoundation brought him back to his home state of Connecticut after stints in Washington, D.C. and Colorado focusing on education policy. These included positions with the Colorado Department of Education as Associate Commissioner, Denver Public Schools as an executive on loan, and the Colorado League of Charter Schools as vice president. In Washington, D.C., Richard served as a senior advisor to the CEO of DC Public Schools and as professional staff for the Senate Committee on Appropriations.