

“Accountability Reimagined”

An American Youth Policy Forum Capitol Hill Forum

Monday, February 8, 2016

Background

The recently signed Every Student Succeeds Act (ESSA) provides states with greater flexibility to design accountability systems that use multiple measures beyond test scores. Educators and policymakers are increasingly realizing that a more nuanced approach to accountability is necessary if education systems are to prepare students with the knowledge and skills they need to be successful in college, careers, and life.

This forum highlighted approaches to thinking more broadly about accountability, moving away from a focus on assessments and considering recommendations from a recent Stanford Center for Opportunity Policy in Education (SCOPE) [report](#). This report encourages a new paradigm of accountability with a focus on “meaningful learning enabled by professionally skilled educators, supported by adequate and appropriate resources.”

The panelists for the session included:

- Dr. Linda Darling-Hammond, President and CEO, Learning Policy Institute
- Paul Leather, Deputy Commissioner, New Hampshire Department of Education

AYPF Senior Director Loretta Goodwin opened the session by noting that this an exciting moment in the world of public education. Following the recent enactment of ESSA, multiple entities have been holding forums and promoting discussions about its accountability provisions. The current session adds to that dialogue.

Presentations

Linda Darling-Hammond provided background information on ESSA and discussed the SCOPE report.

Darling-Hammond began her remarks with a focus on the future and the need to prepare children for a world significantly defined by rapid changes in the economy and domestic job market. Also important is the rapid growth in knowledge around the world—in the short period between 1990 and 2003 there was more knowledge created than in the rest of previous history. Children must take in this new knowledge, learn new technologies, and solve big problems such as climate change. These challenges require higher order learning and thinking that cannot be measured through simple multiple-choice tests. We must assess student ability to learn and find solutions. Future testing and evaluation systems will look less like they did under No Child Left Behind (NCLB) and more like *The Internship*, a movie starring Vince Vaughn

and Owen Wilson that depicts a multi-day interview process at Google. The plot depicted candidates being evaluated according to their skills, ability to solve problems, and teamwork.

NCLB encouraged far different approaches to assessment. The theory behind the legislation was that a focus on school achievement would lead educators and policymakers to improve education. Its strategies included annual testing, setting targets for improvement, identifying schools that failed to meet all the targets, implementing school consequences, and tying test scores to teacher evaluations.

The results were not overwhelming. Graduation rates went up somewhat. NAEP scores went up, but improvements were half as great as they had been prior to NCLB. Students' higher order skills actually declined.

What caused these results? The tests focused on lower order skills such as memorization as opposed to problem solving. There were no incentives to enrich curriculum offerings. Instead, emphasis was placed on basic course offerings with less attention being paid to classes such as social studies and computer science. The drivers of achievement were invisible. The focus on schools and teachers did not account for other important factors such as unequal school resources, growing poverty, and state and local policies. It was evident that mandated solutions are often unhelpful.

The SCOPE report highlights the value of a kitchen cabinet of advisors. It argues that accountability systems should:

- 1) Encourage high quality teaching and learning,
- 2) Provide tools for continuous improvement, and
- 3) Provide means for identifying and addressing problems.

Tests can offer information for an accountability system, but they do not by themselves create accountability. Further, to have real accountability, you must improve schools. Parents expect meaningful learning that keeps children engaged and prepares them for the real world. Offerings should be professional and schools should have adequate resources to realize their goals in an effective way.

In addition, accountability systems should . . .

- 1) Be **reciprocal** with each level of the system taking responsibility for the contributions it must make to serve each child well.
- 2) Be designed to produce **continuous system improvements** (at the school and district levels).
- 3) Include **assessments of varying kinds**. When your car stops on the freeway, you want more than a single indicator to help you figure out what's wrong. For young people, there should be a dashboard of indicators reflecting student success, engagement, and opportunities to learn.

- 4) Provide **transparent and accessible information** to the public.
- 5) Seek and reflect **student, parent, educator and community input**.
- 6) Require state/district **attention to struggling schools and flexibility for interventions** based on data.

One of the benefits of NCLB was data disaggregated by race and other sub-categories such as disability and English proficiency to measure achievement across student groups. These provisions remain under ESSA and can be incorporated into new accountability systems.

The U.S. is alone in the world in its frequency of testing and primary use of multiple-choice instruments. Other countries (and some American states) have been utilizing performance tasks and portfolios for some time. They have also been asking open-ended questions and allowing multiple days for thinking and revision. Schools with International Baccalaureate (IB) curriculums empower students to design scientific investigations and defend their work to a committee—these schools send young people (including those of varying races and incomes) to college at higher rates and with better outcomes. ESSA works to align the U.S. with these international best practices, allowing for multiple up-to-date measures of student academic achievement. States may apply for innovative assessment pilots to explore these best practices.

Darling-Hammond provided some information on how a new accountability system was being shaped in a sample state. California is examining school learning across multiple domains such as student achievement, student engagement, school climate, curriculum access, and parent engagement. Educators in the state now have the capacity to tell what percentage of their students are completing the college prep curriculum and thus they are starting to pay attention to how schools are preparing students for their futures. Importantly, California and other states are able to assign differing weights to each domain in their overall accountability systems.

Additionally, new forms of instructional support are emerging and expanding. Learning networks amongst schools and districts are facilitating exchanges of best practices. School Quality Reviews are sending evaluators to individual school sites to observe and help educators. High performing schools are being paired with low performing ones. Many states are placing high-performing teachers in leadership positions (e.g., making them mentors or master teachers). California started an Instructional Leadership Corps. Another option is to develop learning networks around math, English, or other subjects. Finally, engaging teachers in the scoring of performance assessments and School Quality Reviews is proving helpful to the way they think about their own teaching.

As the nation embarks on this new opportunity to redesign accountability systems, it is important to keep our eyes on the prize for the students—they must be ready to be leaders and meet the demands of a changing world.

Paul Leather discussed New Hampshire's Performance-Based Assessment and School Accountability Pilot.

New Hampshire's accountability system has been evolving over time. The 1.0 version was driven by NCLB and standardized testing. The state subsequently decided to apply for and received a NCLB accountability waiver to form the [Performance Assessment of Competency Education](#) (PACE) pilot. This was the 2.0 version. The current and third version of the accountability system is an innovation lab network, involving the input of a number of people over a period of several years. The emphasis is on being more personal and competency-based.

New Hampshire educators asked themselves the following question: What would the 51st state (reference in the SCOPE report) do if it had no prior accountability or NCLB history? They decided it was important to have locally selected assessments of student progress (performance-based and other), state validation assessments, and graduation portfolios.

There was also a need to support educators so they can better support their students. A [paper](#) written by Ben Jensen became a resource. It positions student learning as the core of all efforts. Educators must work to improve instruction but their professional judgment becomes a greater part of the process. Leather noted that Canada adheres to this philosophy and achieves impressive results.

Within New Hampshire's new system, students are called upon to confront real world problems (e.g., doing math and thinking about its applications). Leather provided an example of a test question that puts students in the role of community planners, asking them to use their math skills to develop a proposal designed to expand a town's water supply. The geometry-focused assessment was aligned with grade 9 math standards while also evaluating how students are able to collaborate and communicate. This approach is more reflective of the challenges students will face when they leave school and reflects a more intense focus on maximizing student learning, engagement, and outcomes. Leather stressed the need to create space for such innovative approaches while studying implementation and results.

Commenting on early accountability systems, Leather posited they were holding educators back. They were externally oriented, leaving educators feeling disenfranchised which caused communities to lose faith in their schools. New Hampshire wanted to harness educator ownership over the accountability system.

PACE only involves a limited number of school districts. Each had to demonstrate it was ready to participate in the pilot and a number of districts are still going through that process. It is challenging to get everyone in the state on board at the same time. Thus, it is helpful that ESSA allows states to focus on a limited number of districts at the beginning of their change processes.

Leather provided more detailed information about New Hampshire's process, detailing extended discussion that began in March, 2013 with Arne Duncan, the former Secretary of the

U.S. Department of Education. NH received final approval of its waiver request on March 3, 2015. The state did the necessary training in the interim period, before they even finished building the system. State-level organizers ensured that local teams and national supporters were involved and invested throughout the process. For example, they were included in meetings with the U.S. Department of Education. After hearing about the inclusive planning process, the USDOE said it would be a lot easier for the agency to support local planning if more states did the same.

PACE's state model competencies are based on national and state standards. Students are tested once in elementary, middle, and high school to track how they are growing developmentally. The state also urges PACE school districts to develop their own standards and assessments that are coherent and understandable to all. Districts use multiple performance assessments that are tied to competencies. They have annual reviews that compare data across schools (just like non-PACE districts). Having a common (to all PACE districts) performance assessment allows for various forms of comparability across districts and coordinated systems. Evaluation is not based on a single measure but a full body measures, most of them developed locally.

Leather also noted certain challenges with PACE. Disaggregated data shows that some student groups are struggling—for example, those with disabilities. And there is a possibility that local teachers are scoring children more highly than general assessments. However, teachers are invested in the assessment system and believe in the data they are collecting and what is being tested. Some describe PACE as the best professional development they have had over the course of their careers. Educators are once again at the center of developing systems.

Audience Q&A

Question 1 was an inquiry about state progress in developing new accountability systems.

States are moving away from the old systems established under NCLB. Most are at the beginning of the process while others had open-ended questions and performance tests prior to NCLB. Those that use performance-based assessments are injecting new life into their existing instruments. Those that had abandoned them under NCLB are stepping up to quickly reinstitute their previous assessment systems.

Question 2 reflected an interest in multi-domain evaluations. How do states/districts account for non-academic domains? What are the input and output measures? How do they weigh academic measures against non-academic ones?

Darling-Hammond commented that it is important to focus on the development of standards first, followed by accountability concerns afterwards. States must figure out what they are trying to accomplish, then they can figure out the appropriate measures to assess progress. A SCOPE paper (currently under peer review) itemizes the types of measures that can be employed. In addition to outcome indicators, there are potential “opportunity to learn”

indicators such as completion of college readiness courses or access to high quality curriculums. NCLB did not incorporate these types of measures since it focused on memorization and recall. If you believe socio-emotional approaches are important, she concluded, you should focus on them but not use that information as a flag for punishment.

Question 3 shifted the conversation to parents. How can parents be effectively involved in evaluation systems? At what point in the school year do assessment results come back? And how are they used to inform parents and learning?

New Hampshire and California have mechanisms that facilitate local community engagement. California passed a law requiring local community collaboration. The outcomes the state wants to monitor provide communities with information about resources being provided to students and schools. Parents may also be interested in knowing that the state 1) assumes the role of stepping in to assist struggling districts and 2) tracks concerning factors such as expulsion rates.

New Hampshire's PACE has introduced reciprocal expectations. It has an interest in seeing improvement—they believe this happens when people buy into what we are doing (including the parents). Performance assessments are embedded into the curriculum and classroom learning so students/teachers/parents are getting the results immediately.

Question 4 was about one of the New Hampshire slides. It seems to indicate that ELL students are doing as well as other students. Why is this the case?

Truthfully, Leather responded, the state is not very diverse. There are only a small number of ELL students and many come from Asian countries with high performance standards. Our concerns are mainly about poor kids and those with disabilities.

Question 5 was "What are you most excited about in ESSA?"

Darling-Hammond praised changes that allow for multiple evaluation measures—schools and students can be viewed holistically. The law requires states to think about teaching students higher-order skills. There is an opportunity to provide students and teachers with information that will inform learning. Assessments have the potential to be engaging and challenging. This is a departure from traditional tests that do not motivate students or help teachers learn how to teach more effectively. Other countries are much more advanced in this space so the U.S. will be able to start catching up.

Leather is really excited about section 1204 (the innovation assessment accountability demonstration authority). Innovation is now valued by Congress. There will be gains for certain students who were lost and ignored in the previous system. It is important that the nation is moving away from educator evaluation as the sole method for improving outcomes. We can now focus on building systems in which educators are valued and respected.

Question 6 was focused on ways to engage higher education teacher preparation programs.

New Hampshire's commissioner has a background in higher education teacher preparation. She has an innovation network consisting of multiple institutions that have been included in the PACE development process. The National Center for Assessment has doctoral students who have completed a lot of the psychometrics for New Hampshire as a part of their dissertation research. However, teachers are still not coming into the profession fully prepared for this new kind of teaching. We must do more.

California has been developing new standards for teachers. Student teachers are excited about the new reforms.

Question 7 pertained to American Indian students. How do we focus on language acquisition and the role of culture in learning? How can tribal leaders be involved in the assessment development process when they do not have a background in that area?

The panel was unsure of the specific answers to those questions. However, in general, many communities (and educators) have a lot to learn as they start to participate in these processes. Hopefully, communities will consider this as an opportunity to appreciate their differences and cultural richness. They should celebrate individuals who can move between languages and cultures, recognizing their strengths and building upon the value of those strengths.