Written Statement of Dr. Linda Darling-Hammond
President and CEO, Learning Policy Institute

Before the
Committee on Health, Education, Labor & Pensions
United States Senate

Full Committee Hearing:
ESSA Implementation: Perspectives from Education Stakeholders on Proposed Regulations

July 14, 2016

I. Introduction

Chairman Alexander, Ranking Member Murray, and Members of the Committee, thank you for your invitation to participate in this hearing.

My name is Linda Darling-Hammond. I am the Charles E. Ducommun Professor of Education Emeritus at Stanford University and serve as the President and CEO of the Learning Policy Institute (LPI).

The Institute conducts and communicates independent, high-quality research to improve education policy and practice. Working with policymakers, researchers, educators, community groups, and others, we seek to advance evidence-based policies that support empowering and equitable learning for each and every child.

I am honored to be here today.

As a parent, an educator, and a researcher, I want to begin by congratulating the Congress on the many ways in which the Every Student Succeeds Act (ESSA) builds on our knowledge of what works in education and how schools can be improved. To a much greater extent than its predecessor, ESSA affords states the opportunity to design accountability systems that both support continuous improvement across all schools while accurately identifying and assisting schools that are struggling to meet the needs of all students. It recognizes that educational improvement must increase students' ability to succeed in the 21st century, fostering such skills as critical thinking, complex problem-solving, effective communication and collaboration, and the ability to learn independently in a rapidly changing world.

The main point of my testimony is that the regulations for the law must allow for accountability that leads to this equity and improvement, by providing transparency and clarity for action both for schools that are failing overall or struggling in certain regards and for all schools to continually improve. At the same time, it must allow for the innovation that will carry this country and its people into the 21st century innovations in learning, teaching, and schooling that are necessary for our national success.

Changes are clearly needed in our educational systems. According to the most recent Program for International Student Assessment (PISA), a test of applied learning and higher order thinking skills released by the Organization for Economic Co-operation and Development (OECD), the United States ranked, among 34 countries, 27th in mathematics, 17th in reading, and 20th in science.[1] Between 2000 and 2012, when No Child Left Behind was in place, U.S. scores and rankings on PISA declined in all areas tested. The greatest challenges are in the schools serving our lowest-income students. And poor children are a growing share of the U.S. population, now comprising more than half all public school students.

It is clear that we need to do some things differently than we have attempted in the past. While we have some wonderfully successful schools, our system as a whole is not ensuring that all students can graduate from high school well prepared for their futures. To achieve this goal, we will need to redesign accountability and improvement systems to support these efforts.

As the U.S. Department of Education (the Department) works towards finalizing regulations to support the successful implementation of ESSA, these regulations must allow states to develop accountability systems that increase equity in educational opportunities and outcomes and that drive continuous improvement for all students and schools. These systems should offer transparency, without sacrificing the specific information needed to determine supports and interventions. And the regulations should support innovation, so that states can implement changes that will support and measure the kind of 21st century teaching and learning that are necessary for our national success.

Congressional intent to support these types of systems is evident throughout ESSA. I join many educators who applaud the wise approach Congress took to create a law that allows states to move away from the restrictions of NCLB that hampered continuous improvement and innovation. While well intentioned, NCLB resulted in rigid accountability systems that were often counterproductive to increasing equitable and meaningful educational opportunities for all students.

I also applaud the U.S. Department of Education for including in its proposed regulations a number of provisions that will help support the goal of creating accountability systems that drive improvement for all schools and all students. For example, the proposed regulations emphasize the need to ensure that interventions are not only based on each school's unique situation but they must also be evidence-based and locally-determined. The proposed regulations also support the use of indicators for both identification and diagnostic purposes. The use of diagnostic indicators can provide additional data to inform the use of funding for professional development, direct student services, and school improvement.

There are some areas, however, where the proposed regulations could be counterproductive to state efforts to support continuous improvement in all schools for all students. Other regulations could reduce the opportunities for states to develop much more effective systems for addressing inequalities and improving schools. The following testimony highlights these areas and provide

[1] Organization for Economic Co-operation and Development, "Program for International Student Assessment (PISA) Results from PISA 2012: United States" (Washington, DC: Organization for Economic Co-operation and Development, 2013).

recommendations on how the final regulations can support states in developing, implementing, and improving upon accountability systems that drive continuous improvement to ensure that all students develop the skills necessary to succeed in the 21st century.

II. Key issues

The final regulations issued by the Department, as well as any subsequent Guidance and Technical Assistance should:

1. Allow states to develop useful dashboards of information that provide transparency and guidance for productive action. ***The regulations should not require a single summative score***, which could limit a state's ability provide the data needed for schools and states to act wisely and well on behalf of the students and families, while hindering the ability of parents and community members to advocate wisely and well on behalf of their children.

2. Allow states to use additional indicators of school quality, beyond the four that are federally required, in meaningful ways that recognize and incentivize schools for their progress on these measures. The proposed regulations would essentially render the "5th indicator(s)" meaningless in the process of identifying schools, thus undermining efforts to eliminate disparities and increase student opportunities to learn. ***The regulations should not restrict state options for weighting and using these additional measures*** in meaningful ways to add to the information that is used to examine school success. This should include the meaningful use of extended-year graduation rates in state accountability systems, which incentivize schools to keep in, rather than pushing out, students who cannot graduate in 4 years and to re-attract those who have left.

3. Allow states to use continuous measures of achievement (such as scale scores, and movement across performance categories), in order to better measure progress and equity gaps. This approach encourages schools to pay attention to students at all points along the achievement continuum, and provides states with better information about progress and outcomes for all students. ***The regulations should not require reporting of student performance by the percentage of students who have met a single cut-point*** which has, under NCLB, focused attention disproportionately on assisting students near that cut point (the so-called "bubble kids") to the detriment of others.

4. Ensure sufficient time to implement thoughtful and effective accountability systems which incorporate stakeholder feedback and have the capacity to drive effective strategies for improvement in schools. ***The regulations should give states until the 2017-18 school year to use new systems for evaluating school progress and identifying schools for intensive assistance.***

Finally, it seems advisable for the Department to reconsider is its approach on how states respond to low participation rates on statewide assessments. The proposed regulations outline very specific consequences to be applied when there is a participation rate of less than 95% for any group in any school. The Department proposes a menu of options for states dealing with schools that fall below the 95% participation rate threshold, including:

- lower summative performance ratings
- lowest performance level on academic achievement indicator
- identified for targeted support and improvement
- state determined action that is equally rigorous and approved by ED[2]

These consequences would confuse actual student performance with the numbers of students taking tests, and reduce the clarity and transparency of ratings, decisions, and actions. A number of officials and educators have indicated that these approaches could backfire and cause greater challenges for them as they seek to build a culture of engagement in new assessments and systems. Encouraging states to determine and clearly articulate how they will factor the requirement for 95% participation in assessments without federally prescribed sanctions will likely better help address the previous misuse of and current responses to high-stakes testing.

For the purposes of my oral testimony, I will focus on the first item – how the Department, in its final regulations, can support states in developing accountability systems that are transparent, while also providing the information needed to drive improvement across all schools for all students.

III. §200.18 Meaningful differentiation of school performance: Preserving a Robust Dashboard to Guide Improvement

ESSA requires that states identify at least 5% of their Title I schools for comprehensive assistance based on their new accountability systems, which will include multiple measures, such as literacy and math achievement, English proficiency gains, graduation rates, and other indicators. The law does not prescribe a particular method for this identification, aside from noting that the 4 academic measures specified must have "much greater weight" than other measures the states add. However, the proposed regulations would require states to produce a "single summative score" on which to rank all of the schools in order to choose the "bottom 5%."

Many states – including California, Kentucky, Vermont, and Virginia, among others -- are well along a path toward developing new accountability systems focused on better information for school intervention and improvement that they believe will be undermined by this requirement, because it will mask important information and make it more difficult to target the right supports to the right schools in the right ways. Several of these states have used a single measure, such as an index or a grading scheme, in the past and have found that it impeded useful improvement.

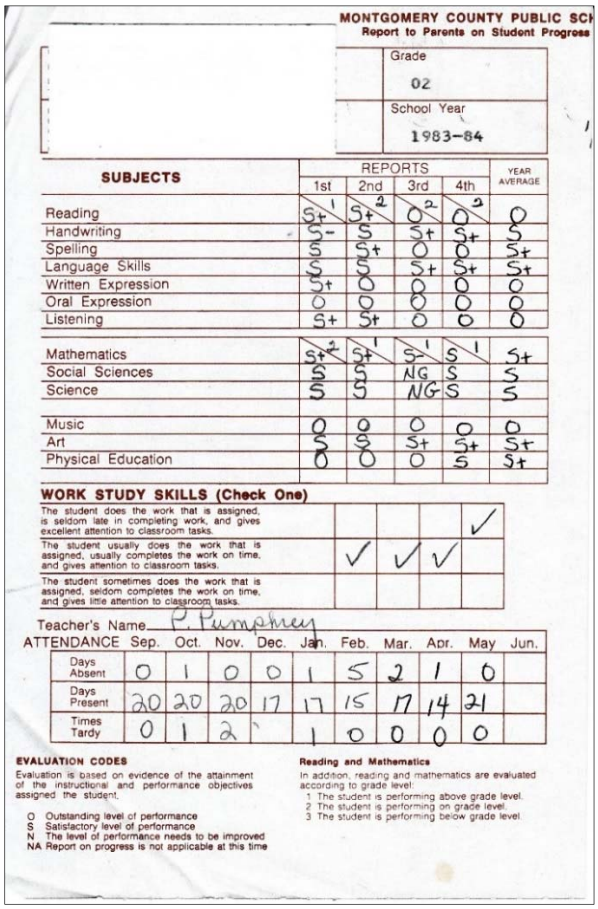
Their experience was that large amounts of resources and attention were directed to the single summative score at the expense of many other factors that impact teaching and learning. Schools could rest on their laurels if they ranked above an arbitrary cut point, rather than paying attention to continuously improving performance on every indicator. Important factors and data were forgotten because they were buried underneath the score. And important needs for groups of students and schools as a whole went unaddressed.

[2] Page 34548: Section §200.15

Parents and educators who work directly with children understand this from their personal experience. A single summative score is not needed and can get in the way of understanding where and how improvement efforts should be focused.

Years ago, when my three children were young, I eagerly awaited the report cards that told me how my children were doing in each of their school subjects, such as reading, writing, math, science, social studies, art, music, and physical education. The most useful of these report cards also provided information on such things as homework and study habits and citizenship. This information was clear and easy to understand, while also revealing specific areas where I could praise or help my child – and where the teacher and school needed to provide additional support.

Figure 1 – Student Report Card



MONTGOMERY COUNTY PUBLIC SCH
Report to Parents on Student Progress

Grade 02
School Year 1983–84

SUBJECTS	REPORTS 1st	2nd	3rd	4th	YEAR AVERAGE
Reading	S+ 1	S+ 2	O 2	O 2	O
Handwriting	S-	S	S+	S+	S
Spelling	S	S+	O	O	S+
Language Skills	S	S	S+	S+	S+
Written Expression	S+	O	O	O	O
Oral Expression	O	O	O	O	O
Listening	S+	S+	O	O	O
Mathematics	S+ 2	S+ 1	S 1	S 1	S+
Social Sciences	S	S	NG	S	S
Science	S	S	NG	S	S
Music	O	O	O	O	O
Art	S	S	S+	S+	S+
Physical Education	O	O	O	S	S+

WORK STUDY SKILLS (Check One)

The student does the work that is assigned, is seldom late in completing work, and gives excellent attention to classroom tasks.				✓
The student usually does the work that is assigned, usually completes the work on time, and gives attention to classroom tasks.	✓	✓	✓	
The student sometimes does the work that is assigned, seldom completes the work on time, and gives little attention to classroom tasks.				

Teacher's Name P. Pumphrey

ATTENDANCE	Sep.	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	May	Jun.
Days Absent	0	1	0	0	1	5	2	1	0	
Days Present	20	20	20	17	17	15	17	14	21	
Times Tardy	0	1	2		1	0	0	0	0	

EVALUATION CODES

Evaluation is based on evidence of the attainment of the instructional and performance objectives assigned the student.

O Outstanding level of performance
S Satisfactory level of performance
N The level of performance needs to be improved
NA Report on progress is not applicable at this time

Reading and Mathematics

In addition, reading and mathematics are evaluated according to grade level:
1 The student is performing above grade level.
2 The student is performing on grade level.
3 The student is performing below grade level.

In all of those years of parenting, it never once occurred to me to ask any of these schools for a "single summative score" to describe my child. I didn't need it to understand how my child was doing, and in fact it would have gotten in the way. I wanted and needed to know exactly where they were doing well and where they were in need of help, so that I could support them. The school needed that information as well. In fact, in my own personal experience, two of my children are dyslexic and while they performed well overall, the need for additional support in reading would have been masked if a single rating were the measure the school focused on.

Ranking all the first graders against each other, giving each an overall rating, and then identifying only the bottom 5 % for extra help, would have missed the mark.

The use of a single summative score would provide neither myself nor my child's teacher the information necessary to identify areas of improvement and act on them. Similarly, schools and districts need reporting systems that allow them to identify individual students and groups of students who may need intensive help in reading or math in order to design, target, and implement interventions like Reading Recovery or math lab. And they need to know which students are chronically absent in order provide organized outreach to the home for students who are not getting to school. This requires specific indicators that are individually reported, not a single summative score. And it requires a set of interventions that are targeted to the specific needs that are identified.

Some states are thinking about the same approach to identification and improvement of schools: On each of the indicators they use, they could identify schools that are low-performing and not improving (or that have large, persistent equity gaps), and provide focused intensive assistance to those schools to really help them improve in that area. For example, the state could identify and work with a group of schools that are not making sufficient progress in supporting English-language proficiency gains by organizing research about what works, examples of local schools that have strongly improved and can be visited and studied, curriculum materials and program models that can be adopted, professional development for educators, and coaches who work directly in the schools.

Just as targeted interventions can be organized for students who are struggling in a particular area, so such interventions can be organized to support networks of schools that share a common need. The same thing could be done with schools that are struggling in mathematics performance, for example, or graduation rates or high suspension rates, overall or for specific groups of students. The state might identify the neediest schools in each indicator area for intensive intervention. The total number of schools assisted might be more than 5%, but each could receive help for the specific areas of need. Across the set of indicators, some schools will be low performing in several and could receive more comprehensive services.

Research has demonstrated the power of the targeted interventions for networks of schools that share similar needs. As we describe in our LPI report, written in partnership with the Stanford Center for Opportunity Policy in Education, *Pathways to New Accountability Through the Every Student Succeeds Act,*[3] a number of states are developing accountability systems that incorporate this type of approach to school identification and continuous improvement. These systems aim to identify schools that are low-performing and not improving within each of several indicators, and/or have large equity gaps. Once identified, these schools can be provided with focused, intensive assistance to improve in the area or areas that are identified, such as English language proficiency, chronic absenteeism, or math assessment for a particular student subgroup.

[3] Darling-Hammond, L., Bae, S., Cook-Harvey, CM., Lam, L., Mercer, C., Podolsky, A., and Stosich, E. (2016). *Pathways to new accountability through the Every Student Succeeds Act*. Learning Policy Institute: Washington, DC. Retrieved from: https://learningpolicyinstitute.org/our-work/publications-resources/pathways-new-accountability-every-student-succeeds-act/.

Like some other states, California is exploring ways to examine both performance and improvement simultaneously on its dashboard of indicators and to classify school performance on each indicator. The example below -- for a college and career readiness index – would be replicated with the others, with data on subgroup performance also added. Schools falling within the red zone on any indicator would be identified for assistance. With the full set of indicators shown in the Figure 2B, the state could also identify all schools that are in the red zone (low performing and not improving) on at least 3 indicators, for example, as part of the group of schools to receive comprehensive intervention and assistance.

Figure 2A – Measuring Progress and Performance

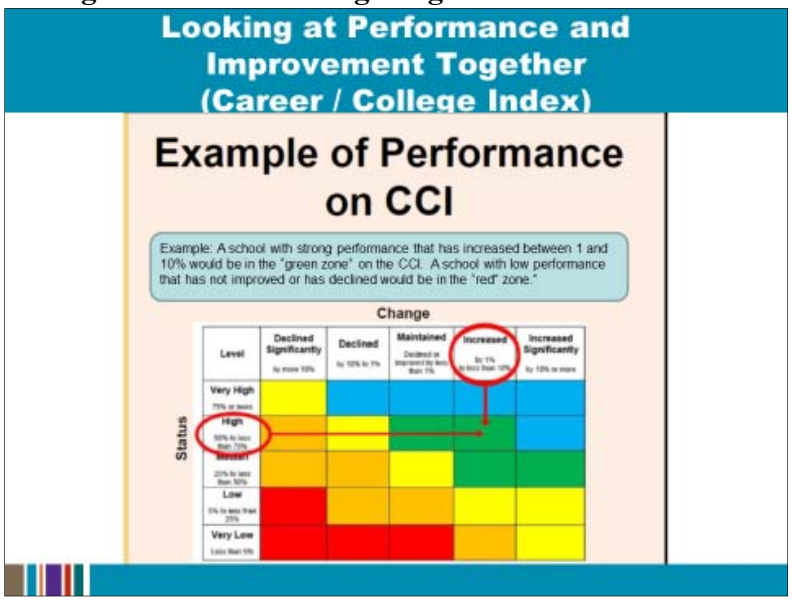
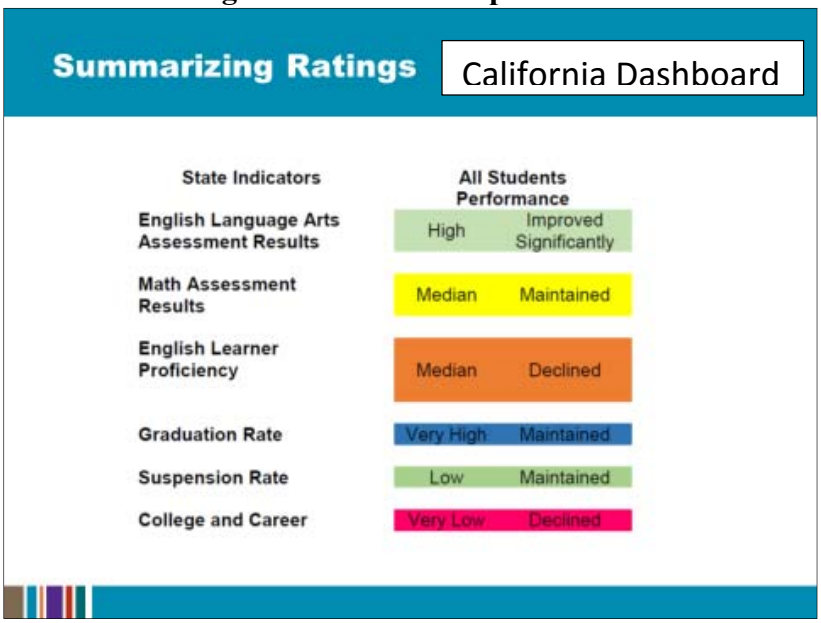


Figure 2B– School Report Card



Vermont is also in the process of determining ways to assess school performance and display the data for the purpose of identifying schools for targeted and comprehensive interventions. The state is in the process of piloting Education Quality Review protocols, or EQRs. EQRs comprise a system of inspection and improvement that is locally developed and implemented, which evaluates schools by measuring five dimensions of school quality:

- academic achievement in English language arts, mathematics, and science, plus graduation rates
- personalization, including personalized learning plans,
- safety and school climate,
- high-quality staffing, and
- financial efficiencies.

The Vermont EQRs will include two complementary processes for assessing these criteria: an Annual Snapshot Review, a multiple measures dashboard of quantitative data conducted by the State; and the Integrated Field Review, a system-level qualitative site review similar to the inspectorate model used in other countries. The Snapshot Reviews are designed to occur annually, whereas the more intensive Integrated Field Reviews occur at least every 3 years. Educators at all levels of the system, state and local, are invited to conduct the Integrated Field Reviews, including but not limited to members of the Vermont Agency of Education (AOE), superintendents, curriculum coordinators, principals, and teachers. During the Integrated Field Review, the review team will "engage in classroom observations, reviews of student work, panel discussions or interviews with parents, students and staff and collaborate to generate their assessments of school system performance."[4] If data from the EQR suggest that there is evidence of substantial inequity and insufficient improvement taking place, the Vermont AOE will intervene with support and sanctions designed to promote improvement.

In the course of consulting with stakeholders on developing a usable report card and school identification system, the state has also been evaluating several approaches (see figures 3A – 3C) which provide different kinds of information. Of note is the fact that, while a weighted index would identify a school like Frakes Secondary as the bottom 5% (figure 3A), it would miss the even lower graduation rates at Madson and Solina High Schools, the lower mathematics performance at Darwish, and the lower reading performance at Lindsay High School which are shown in the dashboard approach (figure 3B).

This critically important information could be taken into account with several kinds of decision rules for identifying the lowest-performing schools, including one that counts the number of struggling areas. (These counts could also be weighted to emphasize the 4 required academic indicators without losing valuable information from the dashboard.) Including improvement or growth information along with information about status (as in figure 3-C) would tell decision makers even more about what is happening in each school, including which of these schools is making progress and which is not.

None of these valuable kinds of information for deciding where and how to intervene would be available with a single summative score.

[4] http://education.vermont.gov/documents/edu-oped-education-quality-reviews.pdf.

Figure 3-A

Type of Aggregation- Index- Average with Weights

School	Reading/ELA	Math	Science	Grad. Rate	PLP	Climate	Staff Satisfaction	Return on Investment	Average
Jones High School	58	65	61	98	72	64	76	15	63.6
Smith Academy High	35	37	36	76	79	56	39	29	48.4
Frakes Secondary School	24	29	31	59	21	75	35	26	37.5
Madson High School	86	80	85	43	54	96	80	82	75.8
Darwish Secondary High School	32	25	35	72	70	57	58	56	50.6
Icenogle High School	86	84	79	84	61	25	72	78	71.1
Palmquist Secondary School	95	89	82	94	35	68	92	89	80.5
Solina High School	31	26	36	35	63	95	47	16	43.6
Spencer Community School	65	63	70	61	49	64	63	73	63.5
Lindsay High School	23	27	25	57	67	43	50	64	44.5

Source: CCSSO Conference, Ryan Reyna and Andrew Rice presenters 6/8/16

VERMONT
AGENCY OF EDUCATION

Figure 3-B

Type of Aggregation- Index- Counts of Struggling Areas

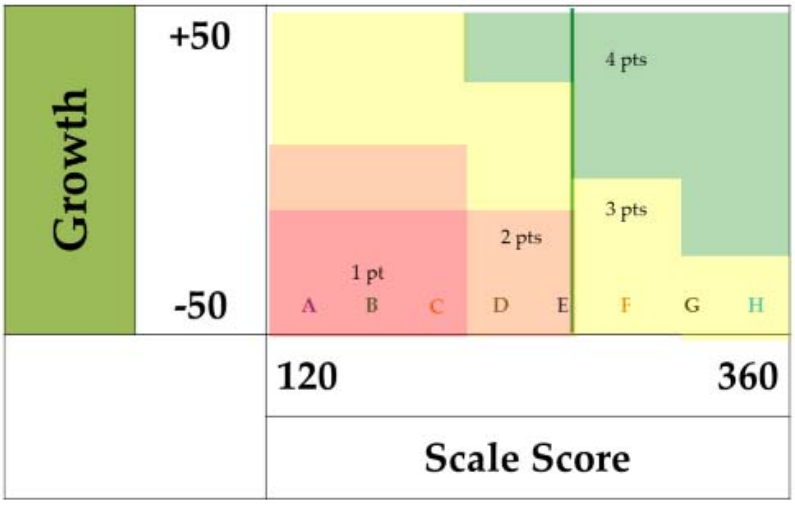
School	Reading/ELA	Math	Science	Grad. Rate	PLP	Climate	Staff Satisfaction	Return on Investment	Counts
Jones High School	58	65	61	98	72	64	76	15	1
Smith Academy High	35	37	36	76	79	56	39	29	0
Frakes Secondary School	24	29	31	59	21	75	35	26	2
Madson High School	86	80	85	43	54	96	80	82	0
Darwish Secondary High School	32	25	35	72	70	57	58	56	1
Icenogle High School	86	84	79	84	61	25	72	78	1
Palmquist Secondary School	95	89	82	94	35	68	92	89	0
Solina High School	31	26	36	35	63	95	47	16	1
Spencer Community School	65	63	70	61	49	64	63	73	0
Lindsay High School	23	27	25	57	67	43	50	64	2

Source: CCSSO Conference, Ryan Reyna and Andrew Rice presenters 6/8/16

VERMONT AGENCY OF EDUCATION

Figure 3-C

Type of Aggregation- Matrix



Source: CCSSO Conference, Ryan Reyna and Andrew Rice presenters 6/8/16

VERMONT AGENCY OF EDUCATION

Kentucky is another example of a state using a multiple measure approach which does not plan to rely on a single summative score to drive identification and improvement. Kentucky's Multiple Measures Dashboard "was designed to have a more balanced approach to determine school success by incorporating achievement, program reviews and effective teaching measures."[5] The Dashboard includes three components: (1) Next-generation learners, which measures performance on areas of achievement, gap, growth, college- and career-readiness, and graduation rates; (2) Next generation instructional programs and support, which includes program reviews for key instructional areas; and (3) Next-generation professionals, which includes data on educator qualifications and effectiveness. The state indicators are able to identify gaps in subgroup student performance and use the data to ensure that all students are developing the skills necessary in the 21st Century.

By contrast, when multiple indicators are aggregated together to yield a summative score, student subgroup performance can also be hidden from view. For example, in one state with an A-F system, the average proficiency rate for African American students in schools that received an A rating was only 58 percent.[6] In another state, 183 high schools received the highest rating within the State accountability system while having at least one subgroup with a graduation rate below 70%.[7]

Efforts by states that are working to develop new models for driving school improvement that privilege equity and innovation could be undermined by a requirement that they produce a single summative score. Their efforts to provide more nuanced, actionable data that is aligned with contemporary learning demands would be traded for simplicity that masks school needs and distracts attention from what should be done to improve performance.

ESSA does not require the use of a summative score and not every state would prefer to use a weighted index to combine indicators into a numerical score and a letter grade or similar rating scheme. There are a wide variety of methods that could meet ESSA's accountability requirements beyond the use of indices, such as a matrix approach that identifies where schools fall in terms of performance and growth with respect to each indicator – and includes schools for intervention on each of the separate measures – and/or decision rules that result in school classifications based on the number of areas in which schools fail to meet a standard.

The Department's regulations should seek to ensure that transparency is a major criterion for identifying schools, along with clear, rational decision rules based on actionable data. While some states may choose to have a system that produces a summative score, the Department should leave open the possibility of other systems of school identification, based on a robust data dashboard that provides information to stakeholders and informs improvement efforts.

IV. §200.18 – Weighting of Indicators

ESSA shifts from an old framework that primarily relied upon performance in math and reading to define a school's success or failure, to a new approach that measures school quality based on a

[5] http://education.ky.gov/AA/Acct/Pages/default.aspx.
[6] https://edtrust.org/resource/making-sure-all-children-matter-getting-school-accountability-signals-right/.
[7] Alliance for Excellent Education analysis of accountability data for Colorado.

combination, at the very least, of five separate measures. Section 1111(c)(4)(C)(i) of ESSA empowers states to design their own accountability systems to fit within a minimum set of Federal parameters: academic achievement in reading and math, the high school graduation rate, English proficiency gains for English learners, and one or more state selected measures of school quality and student success.

These measure(s) of school quality or student success offer the promise of a more comprehensive view for parents, students, educators, and stakeholders on how their school is performing on a variety of meaningful indicators. Each element of a state's emerging accountability system, if well-chosen, can create incentives and opportunities to move school practices forward in ways that better ensure all students are successful. With a well-designed dashboard of measures, educators and community members can track information about inputs, processes, and outcomes to inform a diagnosis of what is and what is not working in schools, along with the types and level of intervention needed.

States have the flexibility to determine the weights of the indicators used within each measure so long as academic achievement in English language arts and math, graduation rates, and EL proficiency are each considered substantial factors and in total, "afforded much greater weight" than the school quality/success indicator(s) within any state-designed accountability system.

However, the Department's proposed regulations would under Section 200.18 of the proposed rule would essentially render these additional indicators as meaningless in the accountability system. The Department describes how the first four indicators must be substantially weighted separately and much greater in weight together against the "fifth indicator" (which could be a set of multiple indicators) when identifying the lowest performing 5 percent of schools for comprehensive support and improvement. This identification also impacts which schools with consistently underperforming subgroups of students, specifically those performing as poorly as the lowest performing 5 percent of schools as one of the criteria, will be identified for targeted support and intervention. Specifically, the Department proposes that in order to meet the requirements for meaningful differentiation:

- A school's performance on the fifth indicator may not be used to change the identity of schools identified for Comprehensive Support and Improvement, unless it is making significant progress for the "all students" group on at least one of the indicators that is given substantial weight;
- A school's performance on the fifth indicator may not be used to change the identity of schools identified for Targeted Support and Improvement, unless each consistently underperforming subgroup in that school is making significant progress on at least one of the indicators given substantial weight; and
- A school performing in the lowest performance level on any of the substantially weighted indicators does not receive the same summative rating as a school performing in the highest level on all of the indicators.

Based on these rules, it is unclear how any indicator of school quality or success could be affirmatively used for its intended purpose unless a school shows major improvement on test scores, graduation rates, or EL progress. In other words, the school quality/success indicator

only serves as a downward ratchet for identification purposes: A school is unlikely to be recognized for positive performance on this indicator, significantly compromising its utility or effect on improving practice.

Yet these indicators can be critically important for leveraging equity and greater opportunity for students. For example, many community groups and civil rights advocates have fought hard to include suspension and expulsion data as a measure of school success, given the research which demonstrates both the strong relationship with graduation and the disproportionate rates by which students of color are often excluded from school due to suspensions and expulsions. Evidence shows that removing students from school for disciplinary purposes has a negative impact, sharply increasing the likelihood that they will drop out of school[8] and expanding the achievement gap, as students of color are typically suspended out of school at higher rates than their white peers.[9]

Research also indicates that tracking suspension and expulsion data by student groups can help highlight racially disparate practices and promote positive behavioral interventions in schools that will improve student engagement and academic success.[10]

Our experience in California, where the state includes this measure among the state priorities regularly tracked, is that changes in school policies have sharply reduced the rate of exclusions; school practices are beginning to support more productive approaches to behavioral interventions and social-emotional learning; and graduation rates have been climbing, for this and other reasons. Civil rights groups that are part of an Equity Coalition in my home state have advocated for consideration of these measures and other indicators of school climate as key levers for improving how schools serve all their students.

Similarly, the final regulations should encourage the meaningful use of extended-year graduation rates in state accountability systems, thereby incentivizing schools to keep in, rather than pushing out, students who cannot graduate in 4 years and to re-attract those who have left. The law explicitly allows for reporting of extended year graduation rates, along with 4-year graduation rates; however, the proposed regulations appear to restrict the ability of states to meaningfully count these extended-year rates in accountability determinations. Schools should be rewarded for keeping and ultimately graduating students who need extra support or time to catch up, such as students who may have immigrated to the U.S. as teenagers with little previous education, those returning to school after dropping out for work or childrearing, those who have been incarcerated, or those who simply need more time to reach high standards. Thus, the regulations

---

[8] American Psychological Association. (2008, December). Are zero tolerance policies effective in the schools? An evidentiary review and recommendations. *American Psychologist, 63(9), 852-862.* http://dx.doi.org/10.1037/0003-066X.63.9.852. See also Losen et. al. 2012; Lee, T., Cornell, D., Gregory, A., & Fan, X. (2011). High suspension schools and dropout rates for black and white students. *Education and Treatment of Children, 34(2), 167-192;* Fabelo, A. (2011). Breaking schools' rules a statewide study of how school discipline relates to students' success and juvenile justice involvement. New York, NY: Justice Center, Council of State Governments and Public Policy Research Institute. Retrieved from https://ppri.tamu.edu/breaking-schools-rules/.

[9] Skiba, R. J., Michael, R. S., Nardo, A. C., & Peterson, R. L. (2002, December). The color of discipline: Sources of racial and gender disproportionality in school punishment. *Urban Review*, 34(4), 317-341.

[10] Skiba, R., Chung, C., Trachok, M., Baker, T., Sheya, A., & Hughes, R. Parsing Disciplinary Disproportionality. *American Educational Research Journal, 51*(4), 640-670.

should allow states to use – and meaningfully count-- extended-year graduation rates in their accountability reporting and decision making.

Importantly, many states are working to create indicators of college- and career-readiness that can leverage much higher quality opportunities that are provided much more equitably to students. There is strong research demonstrating that taking college preparatory coursework in high school is correlated with several indicators of college readiness, from college enrollment[11] to grades[12] to persistence and completion.[13] Similar research shows that students who are enrolled in career academies enroll in community college at higher rates[14], are more prepared for college coursework[15], and experience higher wages and greater employment stability.[16]

As examples, Hawaii, Connecticut and New Jersey use the total percentage of students who enroll in any institution of higher education within sixteen months of earning a regular high school diploma as one way to indicate college- and career-readiness.[17] Georgia, Pennsylvania and Arkansas use evidence of rigorous course offerings, including the availability of Advanced Placement, International Baccalaureate, or college credit courses as part of their college- and career-readiness indicator.[18] Over eleven states, including Alabama, Florida, Kentucky and Illinois also use the percentage of students who receive industry certification to measure college- and career-readiness.[19]

Working hard to get more of these opportunities for a greater share of students could transform the futures of millions of young people. Diminishing the importance of indicators that are not in the set of federally prescribed measures indirectly limits the ability of states to meaningfully tackle many of the structural and societal challenges they face in locally relevant ways. In addition, the added language in the proposed rule essentially removes an aspect of state decision-making, which arguably oversteps the statutory boundaries surrounding state determination in the design of new state accountability systems.

The Department should allow states to make these kinds of indicators important and meaningful in their state accountability systems. By overly prescribing the weighting requirements and the

---

[11] Balfanz, R., & Legters, N. (2006). Closing "dropout factories": The graduation-rate crisis we know, and what can be done about it. *Education Week*, 25(42), 42-43.
[12] Adelman, C. (2006*). The toolbox revisited: Paths to degree completion from high school through college*. Washington, DC: U.S. Department of Education.
[13] Long, M.C., Conger, D., & Latarola, P. (2012). Effects of high school course-taking on secondary and postsecondary success. *American Educational Research Journal*, 49(2), 285-322: Willingham, W.W., & Morris, M. (1986). *Four year later: A longitudinal study of advanced placement students in college* (College Board Research Report No. 86-2, ETS RR No. 85-46). New York: The College Board.
[14] Center for Advance Research and Tecnology, (2011). *A model for success: CART's Linked Learning program increases college enrollment*. Clovis, CA: Center for Advanced Research and Technology.
[15] Dayton, D., Hester, C.H. & Stern, D. (2011). *Profile of the California Partnership Academies*, 2009-2010. Berkley, CA: Career Academy Support Network, University of California.
[16] Bishop, J.H., & Mane, F. (2004). The impacts of career-technical education on high school labor market success. *Economics of Education Review*, 23(4), 381-402.
[17] Alliance for Excellent Education. (2009). *Reinventing the federal role in education: Supporting the goal of college and career readiness for all students*. Washington, DC: Author. Retrieved from http://all4ed.org/reports-factsheets/reinventing-the-federal-role-in-education-supporting-the-goal-of-college-and-career-readiness-for-all-students/
[18] *Ibid.*
[19] *Ibid*.; Eleven states include: Alabama, Florida, Georgia, Illinois, Indiana, Kentucky, Maryland, New Mexico, Oklahoma, Louisiana, and Missouri.

uses of additional indicators, the proposed rule could create a perverse incentive for states actually to do less to solve pervasive problems that strongly affect student outcomes. Unless the Department adjusts the language, this provision as currently written would in effect discourage the use of these potentially powerful indicators for states and rollback community efforts underway to address the root cause of educational inequity.

The Department's final rules should support states in their efforts to implement accountability systems that advance equity by highlighting and measuring what matters most for student success and what provides the most useful levers for school improvement.

V. §200.33 – Calculations for Reporting on Student Achievement and Meeting Measurements of Interim Progress

Another area of concern is the way in which states are asked to demonstrate how students are progressing on academic measures. Although the law does not require a particular method of tracking students' proficiency levels, the proposed rules (see p. 34575) indicate that the determination of whether all students and each subgroup of students met or did not meet these State measurements of interim progress must be "based on the percentage of students meeting or exceeding the State's proficient level of achievement" and would be calculated using the method in proposed 200.15(b)(1), in which the denominator includes the greater of –

- 95 percent of all students and 95 percent of each subgroup of students who are enrolled in the schools, LEA, or State, respectively; or
- the number of all such students participating in these assessments

This rule replicates the "percent proficient" standard that was used under No Child Left Behind. However, research has found that a focus on the percentage of students who reach a particular cut point or proficiency standard incentivizes schools to focus only on a selected few students hovering around the proficiency cut score rather than paying attention to all students at all levels of achievement. Studies during the NCLB era characterized this well-documented practices as "educational triage," which resulted in focusing especially on students near proficiency and emphasizing test-specific rather than generalizable skills."[20] Furthermore, research found that the improvement gap was largest in the low-achieving schools, where focusing on students near the proficiency cut score came at the expense of attention to the lowest achieving students.[21] Measures that rely upon moving students across a threshold, for example from "Basic" to "Proficient" create the incentive to over-direct attention to students on the "bubble" at the expense of others.[22]

The use of the "percent proficient" measure also fails to make distinctions among students or schools who are farther away from or closer to the cut points, and those who have made significant progress or have largely stagnated in their progress. It is an uninformative measure

[20] Jennings, J., & Sohn, H. (2014). Measure for measure: How proficiency-based accountability systems affect inequality in academic achievement. *Sociology of Education*, 87(2), 125-141.
[21] Lauen, D., & Gaddis, S. (2015). Accountability pressure, academic standards, and educational triage. *Educational Evaluation and Policy Analysis*; Sparks, S. (2012). Study finds "bubble student' triage a gut reaction to rising standards." *Education Week*. Retrieved from: http://blogs.edweek.org/edweek/indside-school-research/2012/03/study_finds_bubble_student_tri.html.
[22] Jennings, J., & Sohn, H. (2014). Measure for measure: How proficiency-based accountability systems affect inequality in academic achievement. *Sociology of Education*, 87(2), 125-141.

for most purposes – and particularly useless for tracking gains and changes equity gaps in meaningful ways that can describe how students are actually doing and that can inform improvement efforts.

There are other, much more informative ways to report performance and growth, including progress along the entire scale used to reflect scores. For example, such reporting can reveal that students moved, on average, from a score of 234 to 250, while English learners moved from 208 to 240, a rate of improvement twice as great. All of these changes could occur without affecting the "percent proficient" measure at all, or in ways that do not show the actual gains made.

Many states are moving to use this kind of information in their systems in line with evidence which suggests that all test-based key indicators, including English Learners' progress towards English proficiency, should report progress using scale scores to demonstrate student growth and cohort improvement.

For example, Vermont has been collaborating with stakeholders to generate greater understanding around the importance of operating on a continuum of improvement that values growth rather than simply looking only at "above or below" cut scores. State officials note that a school that is 1 percent below an arbitrary target is not substantially different from a school that is 1 percent above the same target. However, a school that falls 1 percent below the target is likely substantially different than one that falls 30 percent below the target – yet both would be treated in the same say under a "percent proficient" reporting system. As opposed to cut scores, using scale scores can help reveal actual performance and showcase how far students progressed towards proficiency and gauge how much learning is taking place.

Although less severe, the Department's proposal in Section 200.18 that would require each state to have a minimum of three performance levels for each indicator could similarly distort the understanding of achievement and exacerbate continued misunderstandings about school quality. The category approach would provide much less information than scale scores, while insisting on status labelling rather than growth measures as the best way to understand school performance.

In the last decade, we have learned that status measures at particular cut points are not the most productive way to measure school contributions to student learning. Many states are moving to include a focus on both student-level achievement and growth over static measures for the well-understood reason that all students arrive at school with varying levels of preparedness – and schools should be recognized for having increased student learning. The construct as currently proposed does not actually describe change or measure the amount of academic progress each student makes over time. Instead, reverting to the old NCLB measures, the Department should make clear that states are allowed to report achievement in more productive ways by encouraging methods that provide increased accuracy and usefulness with information that shows status, progress, and improvement across the full range of proficiency levels.

VI. §200.19 – Timetable for Identification

As you are well aware, many states officials have expressed concern over the feasibility of implementing new systems that maximize the potential of ESSA by the 2017-18 school year.

Without the benefit of more deliberation and thoughtful planning, the currently proposed timeline precludes that opportunity. While calling for urgency for the sake of improving struggling schools is laudable, the unintended consequence may be that states end up resorting to only making minor tweaks to an existing system or worse, locking educators into old measures that maintain the status quo. A rushed timeline also undermines the public engagement process that is needed to ensure strong stakeholder input.

Using 2016-17 school year data to identify schools for intervention and support means relying on old information to inform a brand new system and restricting the entire accountability system to measures already in use, rather than taking advantage of the new opportunities for better information under ESSA.

States need time to revise their new accountability systems; this includes adding new indicators of English language proficiency and school quality or student success. Working in close collaboration with teachers, parents, civil rights groups, community-based organizations and other stakeholders, states also need to agree on how to combine indicators and establish criteria and procedures for school identification, all of which requires substantial time and effort. In addition, many states will need legislative or administrative approval in order to collect the data needed for school identification, including data for the indicators that might not yet exist. In essence, the Department's proposed timeline is unworkable.

As the Department take steps to make these regulations more workable for states, extending the timetable will allow for real stakeholder engagement and enhance the ability of states to implement high-quality accountability systems in 2017-18 while also using these systems to identify underperforming schools.

Thank you for the opportunity to provide my views on the Department of Education's proposed regulations on accountability and state plans. I would be happy to answer any questions that Members of the Committee may have.

Linda Darling-Hammond
President and CEO,
Learning Policy Institute
1530 Page Mill Road, #200
Palo Alto, CA 94304
(650) 332-9700