

---

# education policy analysis archives

A peer-reviewed, independent,  
open access, multilingual journal



Arizona State University

---

Volume 26 Number 13

January 29, 2018

ISSN 1068-2341

---

## How Do States Integrate Performance Assessment in Their Systems of Assessment?

*Elizabeth Leisy Stosich*  
Fordham University

*Jon Snyder*  
Stanford University



*Katie Wilczak*  
Summit Public Schools  
United States

**Citation:** Stosich, E. L., Snyder, J. & Wilczak, K. (2018). How do states integrate performance assessment in their systems of assessment? *Education Policy Analysis Archives*, 26(13).  
<http://dx.doi.org/10.14507/epaa.26.2906> This article is a part of the special issue, *Redesigning Assessment and Accountability for Meaningful Student Learning*, guest-edited by Soung Bae, Jon Snyder, and Elizabeth Leisy Stosich.

**Abstract:** This paper reviews state strategies for incorporating performance assessment in policy and practice. Specifically, the paper reviews the use of performance assessment in 12 states in the Innovation Lab Network, a group committed to developing systems of assessment that provide meaningful measures of college and career readiness. This review suggests that states relied on four central approaches for integrating performance assessment in state and local systems of assessment: 1) classroom purposes, 2) graduation requirement, 3) school accountability purposes, or 4) federal

Journal website: <http://epaa.asu.edu/ojs/>  
Facebook: /EPAAA  
Twitter: @epaa\_aape

Manuscript received: 1/26/2017  
Revisions received: 7/6/2017  
Accepted: 7/7/2017; Advance publication 8/1/2017

accountability. We review these approaches and the benefits and challenges associated with each strategy.

**Keywords:** performance assessment; education policy; accountability

### **Cómo los estados integran la evaluación del desempeño en sus sistemas de evaluación?**

**Resumen:** Este artículo revisa las estrategias estatales para incorporar la evaluación del desempeño en las pólizas y la práctica. Específicamente, el documento revisa el uso de la evaluación del desempeño en 12 estados en Innovation Lab Network, un grupo comprometido con el desarrollo de sistemas de evaluación que ofrecen medidas significativas de preparación universitaria y profesional. Esta reseña sugiere que los estados se basaron en cuatro enfoques centrales para integrar las evaluaciones del desempeño en los sistemas de evaluación estatales y locales: 1) propósitos de la clase, 2) requisito de graduación, 3) propósitos de responsabilidad escolar o 4) responsabilidad federal. Revisamos estos enfoques y los beneficios y desafíos asociados con cada estrategia.

**Palabras clave:** evaluación del desempeño; política educativa; responsabilidad

### **Como os estados integram a avaliação de desempenho em seus sistemas de avaliação?**

**Resumo:** Este artigo analisa as estratégias estaduais para incorporar a avaliação de desempenho em políticas e práticas. Especificamente, o documento analisa o uso da avaliação de desempenho em 12 estados da Innovation Lab Network, um grupo comprometido com o desenvolvimento de sistemas de avaliação que oferecem medidas significativas de preparação para universidades e profissionais. Esta revisão sugere que os estados se basearam em quatro abordagens centrais para integrar avaliações de desempenho em sistemas de avaliação estaduais e locais: 1) propósitos de classe, 2) requisito de graduação, 3) objetivos de responsabilidade escolar ou 4) responsabilidade federal. Revisamos essas abordagens e os benefícios e desafios associados a cada estratégia.

**Palavras-chave:** avaliação de desempenho; política educacional; responsabilidade

## **Introduction**

A fundamental purpose of educational accountability and assessment policies is to create conditions that increase the probability that each and every child will leave the formal education system with the opportunity to pursue a future of his or her own choosing. As President Obama stated at the signing of the Every Student Succeeds Act (ESSA), “With this bill, we reaffirm that fundamental American ideal that every child, regardless of race, income, background, the zip code where they live, deserves the chance to make out of their lives what they will” (The White House, Office of the Press Secretary, 2015). Further, the purpose is to assure that communities where this goal is achieved become the norm rather than the anomaly. The assessment of student learning, as discussed below, does not constitute the entirety of an accountability system that can meet this overarching purpose. Such assessment, however, is one essential component of an effective system of educational accountability policies.

This article is designed to advance the discussion prompted by Darling-Hammond, Wilhoit, and Pittenger (2014) in “Accountability for College and Career Readiness: Developing a New Paradigm” about how to strengthen state educational accountability systems. Darling-Hammond and colleagues proposed a framework for educational accountability that addresses (1) resources (fiscal resources, including how funding is used to structure the work and time of educators, as well as physical infrastructure, curriculum materials, etc.); (2) professional capacity (the abilities of the

adults who work with children directly, and those who support these adults, to meet goals for students); and (3) meaningful learning (access to the knowledge, skills, and dispositions that are necessary for young adults to pursue a future of their own choosing—including curriculum, instruction, and assessments). In this article we analyze the use of performance-based assessments in states. Specifically, we examine how states, districts, and schools integrate performance assessments in their systems of assessment and accountability to support more meaningful learning opportunities for children.

The role that assessment can and should play in accountability is a greatly debated topic. Recently, new policies (e.g., Marion & Leather, 2015; Menefee-Libey & Kerchner, 2015) and public opinion (e.g., Gewertz, 2014; Hagopian, 2014; Pizmony-Levy & Saraisky, 2016) are shifting the balance in accountability policy from an overreliance on standardized multiple choice tests in English Language Arts (ELA) and mathematics to include a more fulsome sampling of the goals of public education (e.g., deeper learning, social-emotional learning, 21st-century skills and dispositions). This paper focuses on how states and districts are “taking advantage” of the potential for changes in state and federal policy to support the creation of conditions that meet the goals of educational accountability policy in the area of meaningful learning (Darling-Hammond et al., 2014). In particular, it focuses on how states and districts encourage the use of performance assessments that support curriculum and instruction as well as, and most importantly, desired outcomes for our children.

Researchers increasingly agree that test-based accountability policies under No Child Left Behind (NCLB) failed to support the meaningful learning opportunities necessary to prepare all students for success in college, career, and life (Heilig & Darling-Hammond, 2008; Jennings & Sohn, 2014; Meier & Wood, 2004; Plank & Condliffe, 2013). Although the law brought much needed attention to the performance of historically underserved students, NCLB’s annual testing requirements led to an overreliance on “bubble” tests that emphasized low-level skills and narrowed the curriculum through a focus on mathematics and reading (Darling-Hammond & Adamson, 2014). Recent changes in federal educational accountability policy, including the reauthorization of the Elementary and Secondary Education Act (ESEA) as ESSA, have created new opportunities for states to use performance assessments as part of their efforts to develop systems of assessment that support deeper learning by more closely integrating assessment with curriculum and instruction.

Specifically, ESSA allows states to use multiple statewide interim assessments in ELA, mathematics, and science during the academic year that would contribute to a final score or scores, rather than a single summative assessment. These interim assessments could include performance assessments. Like NCLB, ESSA allows states to choose when and how to assess any content areas in addition to ELA, mathematics, and science. Additionally, ESSA explicitly permits the use of “portfolios, projects, or extended-performance tasks” as assessments in ELA, mathematics, and science (Every Student Succeeds Act of 2015, 2015). Thus, ESSA creates greater opportunities for states to use performance assessments in their systems of assessment than NCLB. By requiring students to apply their knowledge and skills to construct an original response, performance assessments can evaluate and support students in developing the critical abilities—such as critical thinking, inquiry, communication, and collaboration—that are essential for student success but poorly measured by many traditional assessments (Conley & Darling-Hammond, 2013).

In this article we briefly review the research on performance assessments, describe the present study’s focus on how 12 states use performance assessments, identify four strategies these states used to integrate performance assessments into their systems of assessment, and discuss implications for assessment policy and practice. We reviewed state assessment policies and capacity development efforts and provide examples to illustrate each of the four strategies we identified for

integrating performance assessments: (1) classroom purposes, (2) graduation requirement, (3) school accountability, and/or (4) federal accountability.

## Review of Research on Performance Assessments

Scholars argue that performance assessments can evaluate and develop students' higher level thinking skills by requiring them to construct an original response (Conley & Darling-Hammond, 2013; Parke, Lane, & Stone, 2006). Unlike traditional multiple-choice assessment items, performance assessments “ask students to do and to produce; to demonstrate learning through work authentic to the discipline and/or the real world” (Stanford Center for Assessment, Learning, and Equity, 2014). In this way, performance assessments can test the application of deep content knowledge and higher order thinking skills required for success in college and career while strengthening complex conceptual understanding through the process of completing the assessment (Chung & Baker, 2003; Wei, Pecheone, & Wilczak, 2014). Notably, the design, duration, and complexity of performance assessments can vary greatly from simply completing a sentence with a short phrase to writing an on-demand essay or conducting a multiple-day science investigation and writing up an analysis of the findings (Stecher, 2014). Thus, some performance assessments create more meaningful opportunities for learning than others.

Performance assessments not only measure student achievement but also are opportunities for student and teacher learning (Conley, 2015; Darling-Hammond & Falk, 2013); students learn through the process of completing the assessment, and teachers learn about their students' abilities and knowledge and are subsequently better able to address their students' learning strengths, interests, and needs. The process of designing, administering, and scoring complex performance assessments can provide a teacher information about students' abilities and support teachers in learning to integrate higher level thinking skills in their instruction (Darling-Hammond & Falk, 2013; Goldberg & Roswell, 2000). Research suggests that the use of performance assessments can produce positive instructional changes in classrooms, such as increased instructional time devoted to problem-solving and writing (Cohen & Hill, 2008; Koretz, Barron, Mitchell, & Stecher, 1996).

States that have introduced performance assessments into their systems of assessment have done so based on the theory of action that more complex and authentic forms of assessment, coupled with professional development to help teachers and systems to use them well, could lead to improvements in curriculum, instruction, and student learning (e.g., Cohen & Hill, 1998; Firestone, Mayrowetz, & Fairman, 1998; Marion & Leather, 2015). A key goal in the use of performance assessments is to shift instruction from a focus on discrete, lower level content acquisition, towards the development of deeper conceptual understanding and transferable skills. In the 1990s and early 2000s, several states included in this study engaged in large-scale use of performance assessments, including California, Kentucky, Oregon, and Vermont (Darling-Hammond, Rustique-Forrester, & Pecheone, 2005; Wei et al., 2014). In Kentucky, for example, as part of the Kentucky Instructional Results Information System, students produced a year-long portfolio in ELA and mathematics that was scored locally by teachers. Most teachers reported that the writing portfolios had a positive influence on instruction (Stecher, Barron, Kaganoff, & Goodwin, 1998) and research suggests that, with training, teachers increased their capacity to reliably score student portfolios (Tung & Stazesky, 2010).

Despite early evidence of the benefits of performance assessments, technical requirements under NCLB—particularly related to producing student-level data—and high costs associated with some performance assessment systems contributed to a decline in large-scale use of performance assessments (Parke et al., 2006; Pecheone, Kahl, Hama, & Jaquith, 2010; Wei et al., 2014). To

address cost and quality concerns, states have combined resources through testing consortia (e.g., Partnership for Assessment of Readiness for College and Careers [PARCC], Smarter Balanced Assessment Consortium [SBAC]) that support states in developing high-quality performance tasks that meet the federal government's rigorous technical requirements and reduce testing development and administration costs. Scholars argue that the cost of developing and scoring performance tasks is similar to the cost of more traditional item development when shared among states or districts (Chingos, 2012; Darling-Hammond, 2014; Topol, Olson, Roeber, Darling-Hammond, & Adamson, 2014). Furthermore, the increased flexibility allowed by ESSA may help to reduce the technical constraints of NCLB, which limited the large-scale use of performance tasks for accountability. Under ESSA, for example, states could use multiple, curriculum-embedded performance tasks that include elements of student choice and contribute to a summative score rather than a single end-of-year test (Darling-Hammond, 2017). Importantly, traces of previous performance assessment initiatives remain—including educators, schools, and systems that continued to use performance assessments—and can serve as a foundation for more recent state efforts to strengthen assessment and, ultimately, student learning (Wei et al., 2014).

## Study Design

This paper builds on the work of the Stanford Center for Opportunity Policy in Education (SCOPE) and the Stanford Center for Assessment, Learning, and Equity (SCALE) with the Innovation Lab Network (ILN), a working group of the Council of Chief State School Officers. Members of the ILN are committed to developing “comprehensive systems of formative and summative assessment, including performance-based measures of deeper learning, that provide meaningful measures of college and career readiness throughout a student’s education” (Council of Chief State School Officers, n.d.). At the time of the study, 12 states were members of the ILN: California, Colorado, Iowa, Kentucky, Maine, New Hampshire, Ohio, Oregon, Vermont, Virginia, West Virginia, and Wisconsin (see Figure 1). With varying levels of support from their states, districts, and/or schools as well as differing experiential histories with performance assessments, educators in each of these ILN states were using performance assessments to measure and further develop students’ knowledge, skills, and dispositions. Drawing on the efforts of these pioneering states and the resources that have supported their work, this paper clarifies the key decision points policymakers and practitioners from ILN states addressed in using performance assessments, provides examples of how the states used performance assessments, and provides guidance regarding the educational materials, development, and use processes that may assist the integration of performance assessments in state systems of assessment.

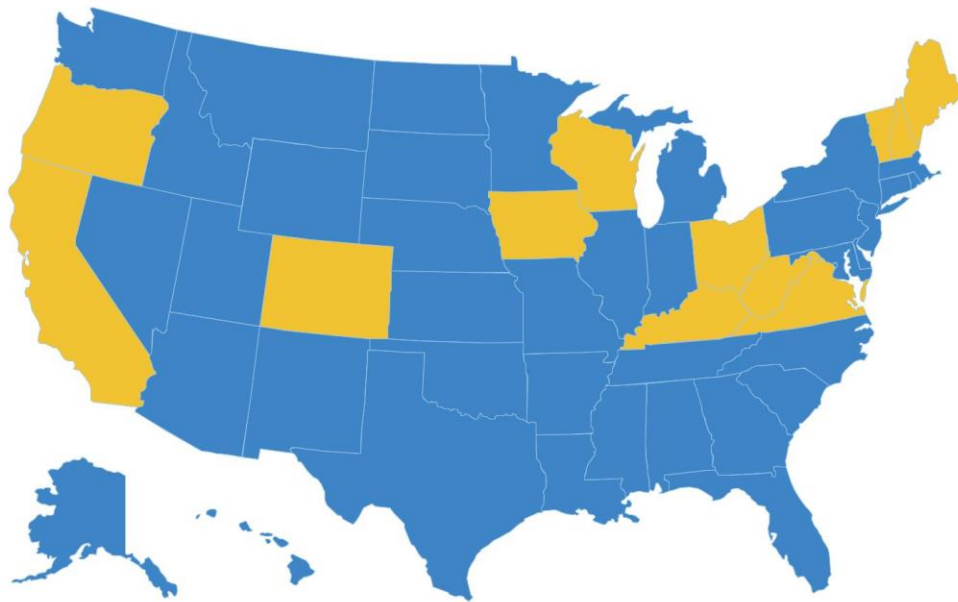


Figure 1. Member States in the Innovation Lab Network.

We conducted a descriptive policy analysis to understand differences in how states integrated performance assessments in their systems of assessment and the possible benefits and challenges of each approach. Specifically, we aimed to answer the following question: What actions do key stakeholders in ILN states take to support and/or require the use of performance assessments in PK–12 schools? Importantly, we analyze differences in the assessment approaches taken in each state, not the quality of these approaches or their outcomes for students. Furthermore, this paper focuses primarily on the potential educative benefits associated with using performance assessments; thus, an analysis of the psychometric quality and reliability of assessments used in these states is beyond the scope of this paper. Although some schools and school systems have implemented performance assessments across classrooms, all states included in this study had yet to completely implement their strategies for integrating performance assessments. Thus, more time would be needed to fully understand the effects of these policies on instruction and student learning. Nevertheless, ILN states varied greatly in the extent to which they had integrated performance assessments. As described in greater detail in the sections below, some states had been working to integrate performance assessments for more than a decade, whereas others were in their first year of implementation at the time of the study.

To collect information, we reviewed current policies and capacity development efforts related to performance assessments in the 12 member states of the ILN. Data collection focused on semistructured interviews with state education agency (SEA) personnel (17 total, including representatives from 9 of 12 states included in the study); information shared by SEA personnel during working group meetings (three online, four in-person); observations of professional learning opportunities for K–12 educators focused on performance assessments (three online, two in-person meetings); state assessment policies; and information from state websites regarding state legislation,

working papers, board meeting minutes, and other relevant documents. Interviews and analysis of state documents focused on the current status of performance assessments in each state and the policies, educator capacity, and technical considerations that supported or constrained the use of performance assessments in the state (Wei et al., 2014). Similarly, SEA personnel were asked to identify state goals, enabling conditions, and constraints for implementing performance assessments during online and in-person working group meetings.

We analyzed data collected from SEA personnel and websites through interviews, observations, and document review using thematic analysis (Boyatzis, 1998) to identify patterns in how states integrated performance assessments in their systems of assessment. We developed themes inductively, beginning with the descriptions from SEA personnel during interviews and observations and comparing these descriptions with state documents. During initial analyses, we met regularly as a research team to discuss points of disagreement, refine our definitions of each thematic code, and determine whether state evidence met these specifications. This process led to changes in our thematic codes to more accurately account for the varied approaches across states. For example, a preliminary code called “formative assessment purposes” was renamed “classroom purposes” since further analysis suggested that the use of performance assessments by teachers for formative and summative classroom purposes were closely linked but distinct from their use for school accountability purposes.

Our analysis yielded four distinct approaches for integrating performance assessments. Notably, some states used multiple strategies. We then revisited the data from each of the 12 ILN states to identify the primary strategy for integrating performance assessments employed in the state or school systems within the state. Finally, we conducted a member check (Merriam, 1995) by sharing initial findings with select state representatives to allow for input and further verify the accuracy and trustworthiness of our interpretations.

## **Strategies for Integrating Performance Assessments in State Systems of Assessment**

In this study, we review current policies and strategies for integrating performance assessments and provide some description of how state policies and capacity building efforts have evolved over time. The work described in this paper is ongoing. The world of educational policy is always changing, and the cast of characters are continually arriving and leaving. Work related to the ILN, for instance, began in the NCLB era, moved into the Race to the Top (and waivers to NCLB) era, and now continues into the ESSA era. In this turbulent world, the catalysts that spurred state interest in performance assessments varied over time. In some states, the adoption of college- and career-ready standards, including the Common Core State Standards and Next Generation Science Standards (NGSS), created demand for assessments that better evaluated the higher level and complex thinking required by these new standards. In other states, the adoption of competency- or proficiency-based learning models spurred interest in using performance assessments to evaluate students’ abilities to apply their knowledge in authentic ways and demonstrate mastery of both academic and so-called “21st-century” skills, such as communication, collaboration, problem-solving, and self-direction. Other states have created opportunities for using performance assessments, with a focus on locally developed and curriculum-embedded assessments, as an alternative to high-stakes standardized tests that are further removed from classroom instruction and curriculum.

Despite the different catalysts for the use of performance assessments, our review of state policies and practices reveals four common yet distinct strategies for incorporating performance assessments into systems of assessment:

- Supporting teachers, schools, and districts in their use of performance tasks for classroom purposes.
- Using performance tasks as one component of a graduation requirement, either by creating a graduation portfolio including performance tasks or establishing proficiency requirements that can be demonstrated through the use of performance assessments.
- Using performance assessments for school accountability purposes or replacing statewide tests with performance tasks when possible, including allowing the use of performance assessments in nontested grades/subjects or to fill the testing vacuum created by the introduction of the NGSS.
- Seeking a waiver from the federal government to alter federal testing requirements and allow performance tasks to be used for federal accountability.

We have identified these as the main approaches ILN states have pursued in their efforts to promote or require the use of performance assessments, regardless of the catalyst that spurred the action. As shown in Table 1, states included in this study employed multiple strategies for integrating performance assessments, and some strategies were more prevalent than others. These different strategies allow states to require and/or encourage the use of performance assessments in ways that consider local context and capacity as well as state-specific goals and priorities. While most of the studied activities took place prior to ESSA being passed in December 2015, under ESSA, states will be able to pursue these same strategies, with a few minor modifications. Significantly, rather than pursuing a waiver to develop and implement performance assessments for accountability purposes, states can now apply to the federal government to participate in a state-designed assessment innovation pilot that would allow some of the same flexibility provided by waivers. As described above, ESSA also includes increased flexibility for *all* states to design innovative assessment systems, including systems that incorporate performance assessments as part of interim assessments or as part of a portfolio of student work.

---

Table 1

*Number of States Employing Each Strategy for Integrating Performance Assessments*

---

<u>Strategy</u>	<u>States</u>
Classroom purposes	11
High school graduation	3
School accountability	7
Federal accountability	1

---

In the sections that follow we describe how ILN states used the four core strategies listed above for integrating performance assessments into their state systems of assessment. We provide examples from states and school systems to illustrate how each of these approaches advances the use of performance assessments.



## **Strategy 1: Performance Assessments for Classroom Purposes**

One strategy employed by almost all (11 of 12) of the states we studied was to support the use of performance assessments for classroom purposes. This strategy involves investment in growing educator capacity to design, develop, implement, and score performance tasks, as well as to integrate those tasks within the classroom curriculum. Underlying this approach is the theory of action that the use of performance assessments is a beneficial instructional strategy that leads to deeper student learning (Council of Chief State School Officers, 2016). Like NCLB, ESSA requires annual testing in ELA and mathematics in grades 3–8 and once in high school; however, performance assessments used for classroom purposes are, by definition, not part of formal state or federal policy. This means that states or districts can support the use of performance assessments without formally taking legislative action or seeking approval from the federal government.

States included in the study supported the use of performance tasks for classroom purposes by providing professional learning opportunities for teachers, allowing for local autonomy to set assessment and instructional priorities, pursuing investment in innovation grants, and establishing partnerships with technical assistance providers. This investment in developing the capacity of educators to support the use of performance assessments is an essential component of the foundation necessary to use performance tasks for reporting or accountability purposes, when and if that formal policy is pursued by the state. In the sections below, we describe two states, Ohio and Colorado, and one school network, Summit Public Schools, that have provided significant support for the use of performance assessments for classroom purposes.

Ohio is currently pursuing the use of performance assessments for classroom purposes through a coalition of six ILN districts working in collaboration with the Ohio Department of Education and regional education service centers to provide professional learning opportunities for teachers and school leaders around the design and development of performance tasks (Ohio Department of Education, 2012). The group of districts originally came together around the goal of acquiring a waiver from the federal government to use performance tasks for federal reporting purposes. In response to the recent passage of ESSA, the coalition refocused their attention on developing the school and district-based capacity required to implement performance assessments widely on a local level. As a grassroots collaborative now directly supported by the Ohio Department of Education, the group has visited schools successfully implementing performance assessments and project-based learning, developed a 2-year professional development plan with the support of technical assistance providers, and sought and won funding from several outside foundations to support its work. The goal of the collaborative is to lay the foundation for an official system of assessment that would include competency-based performance tasks used for state and federal reporting purposes. The purpose of the immediate investment in capacity-development strategies is to increase the probability that, regardless of state or federal policy, teachers will have the skills necessary to use performance assessments well and support student learning with appropriate instructional strategies.

Similarly, the Colorado Department of Education has formed Content Collaboratives—a set of content-specific teams comprising nearly 200 teachers from districts across the state—to review existing assessment items and develop common, discipline-specific performance tasks (Colorado Department of Education, n.d.). The Collaboratives have created and piloted 57 performance tasks in world languages, social studies, music, health, and mathematics, among other fields, and their work is directly supported by the Colorado Department of Education. The tasks created through the initiative are housed in an online resource bank that is accessible to all teachers in the state (and nationally), and the bank is intended to encourage teachers' use of performance assessments.

The Colorado Department of Education also supports the use of performance assessments by teachers around the state through its work with the Literacy Design Collaborative (LDC) and the Common Assignment Study, a project in which teachers build common units of study that include LDC modules (Bill & Melinda Gates Foundation, n.d.). LDC modules are subject-specific reading and writing performance tasks accompanied by an instructional plan that is taught over a 2- to 4-week period. Teachers throughout the state participate in LDC training to learn how to design modules, score student work, and understand the instructional implications of student performance on the task. The Common Assignment Study, a project facilitated in collaboration with the Colorado Department of Education and a team of research organizations, is investigating the feasibility of using LDC performance tasks embedded within common units of study to evaluate Student Learning Objectives (SLOs). SLOs are specific learning goals, along with an identified measure, that track student learning and progress towards the learning goal. SLOs are often identified by teachers, or school or district teams, and allow local educators to measure student growth over a period of time, often in content areas or courses not assessed by summative assessments. The Common Assignment Study approach provides an alternative to SLO evaluation based solely on traditional assessment item formats (i.e., multiple choice, fill in the blank) and encourages more authentic forms of assessment and instruction. While Colorado may ultimately include performance tasks as one measure of academic achievement used for federal accountability and high school graduation, it is now taking the important step of investing in local districts' abilities to design and use performance tasks as well as preparing students and teachers for the instructional shifts accompanying this transition.

School networks and districts are also forging ahead with systems that integrate performance assessments as a key component of their instructional and assessment strategies. Summit Public Schools, a charter management organization in California and Washington, integrates project-based learning and performance assessments to deepen students' development of content knowledge and "cognitive skills" such as problem-solving, communication, and creative thinking (Renner et al., 2015). Each course is driven by a set of projects, developed collaboratively by teachers, which include embedded performance assessments and are evaluated by a common cross-disciplinary rubric. Classroom instructional time is allocated to teacher and peer collaboration around projects, supported by Summit's Personalized Learning Plan, an online platform of "playlists" that allow students to self-direct content knowledge acquisition. Students are expected to reach "proficiency," or level 6, in each dimension of the rubric by graduation in order to demonstrate college readiness (Summit Public Schools, n.d.).

Summit's approach focuses on the learning opportunities afforded by project-based learning and performance assessments and the benefits of a system that links curriculum, instruction, and assessment. So far, it has been successful; a case study of Summit by the Progressive Policy Institute reports that in 2015, some 99% of Summit graduates were admitted to 4-year colleges (Osborne, 2016, p. 15). Summit is an example of a charter management organization independently pursuing the use of performance assessments for classroom purposes. Its students still must take state exams, and the learning opportunities provided by performance assessments and project-based learning may support their success on these external measures.

Although supporting the use of performance assessments for classroom purposes can provide a strong foundation for continued work in performance assessments, it is not without challenges. In-person professional development activities for teachers and school or district leaders can be costly and difficult to scale across numerous schools and districts. Sustaining initiatives over time is often difficult, leadership can be inconsistent, and foundation and state funding fluctuates. Evidence from ILN states suggests that these challenges may be minimized by establishing communities of practice, promoting strong communication among stakeholders, establishing shared

beliefs among leaders and educators about the benefits associated with performance assessments, and strategically using technology. Developing capacity among stakeholders to implement and understand the implications of student performance on performance tasks may encourage meaningful learning in classrooms and create more supportive conditions for adopting policies that support or require the use of performance assessments for consequential purposes.

## **Strategy 2: Using Performance Assessments for High School Graduation**

Maine, Vermont, and Oregon have adopted graduation requirements that allow for or require the use of performance assessments to earn a high school diploma, with the goal of evaluating students' abilities to apply their knowledge in ways that more closely reflect the demands of college, career, and life. Although both NCLB and ESSA require annual statewide testing in ELA and mathematics in grades 3–8 and once in high school, no specific federal requirements exist for determining high school graduation standards. Many states rely on narrow measures of student performance on traditional assessments and course completion to determine high school graduation. In contrast, state policymakers in Maine, Vermont, and Oregon, as well as districts and networks of schools in some states, have developed graduation requirements that assess students' readiness for college, career, and citizenship in more authentic ways. In these states, graduation requirements provide opportunities for measuring student proficiency more comprehensively through the use of performance assessments designed to measure students' mastery of both traditionally defined academic content and 21st-century skills, such as communication, collaboration, problem-solving, and self-direction (Conley, 2015; Conley & Darling-Hammond, 2013). We highlight how these three states and the Linked Learning pathways schools and districts in California have integrated performance assessments in high school.

The use of performance assessments in Maine was spurred by the adoption of proficiency-based diploma standards in 2011. Beginning in January 2018, high school students will be awarded a diploma based on demonstrations of proficiency (Maine Department of Education, n.d.). The diploma standards require that students “be allowed to demonstrate proficiency by presenting multiple types of evidence, including but not limited to teacher-designed or student-designed assessments, portfolios, performance, exhibitions, projects and community service” (Proficiency-Based Diploma Standards and Transcripts, 2015). Notably, performance assessments are only one of multiple assessment options that can be used to demonstrate student proficiency. The state is making concerted efforts to develop educator capacity and resources to support the use of performance assessments as a component of the proficiency-based learning system.

As part of the transition to proficiency-based diplomas, the Maine Department of Education adopted Guiding Principles, which describe five areas in which students must demonstrate competency at graduation: clear and effective communication, self-direction and life-long learning, creative and practical problem-solving, responsible and involved citizenship, and integrative and informed thinking. The Maine Department of Education formed a collaborative working group of school systems, education leaders, and organizations and enlisted the technical assistance of outside assessment experts to develop a set of initial performance tasks aligned with the Guiding Principles (Maine Department of Education, 2014). These are model tasks and rubrics that districts can use or look to as they design their tasks. The piloting of these model tasks will also result in samples of student work that demonstrate proficiency. Model tasks with student work samples can provide information to students and teachers about what attaining proficiency on state standards requires. In addition, these model tasks can support educators as they work to develop their own assessments of student learning. Teachers from 22 districts have engaged in the development, piloting, and revision

of these performance tasks. Significantly, teacher involvement in assessment design can increase the capacity of educators to use performance assessments to evaluate student proficiency (Darling-Hammond & Falk, 2013).

Similar to Maine's efforts, Vermont's efforts to integrate performance assessments as part of its proficiency-based learning system include a focus at the high school level. Specifically, the state is transitioning to proficiency-based requirements for both awarding diplomas at graduation and measuring progress in secondary school. In April 2014, the Vermont Board of Education approved the Education Quality Standards (EQS), which describe expectations for student learning in literacy, mathematical content and practices, scientific inquiry and content knowledge, global citizenship, physical and health education, artistic expression, and transferable skills (Vermont Board of Education, 2014). The EQS includes the Common Core State Standards for ELA and mathematics and the NGSS for science. In addition, the EQS details the state's shift from units and seat time to proficiency as the sole means for determining student progress and high school graduation by 2020. According to the EQS, "students must be allowed to demonstrate proficiency by presenting multiple types of evidence, including but not limited to teacher- or student-designed assessments, portfolios, performances, exhibitions and projects" (Vermont Board of Education, 2014, p. 7).

Vermont is currently growing the capacity of educators to respond to the pending proficiency-based graduation requirements. To assist educators in learning to use performance assessments to support and evaluate student proficiency, the Vermont Agency of Education brought together educators from across the state to develop sample tasks and scoring criteria for assessing transferable skills over the course of 2014–2015. The tasks and processes were piloted during the 2015–2016 school year (Vermont Agency of Education, n.d.). As described in the EQS, "Transferable Skills" refer to a broad set of knowledge, skills, work habits, and character traits that are essential for college and career readiness, including skills related to communication, collaboration, creativity, innovation, inquiry, problem-solving, and the use of technology. Complex performance assessments are well designed to assess student mastery of these crosscutting skills since they require students to apply knowledge and skills in authentic ways, including through collaborative projects, science investigations, or presentations. Importantly, Vermont's efforts to develop educator capacity to use performance assessments as part of proficiency-based graduation requirements grow from the state's history with performance assessments as well as current complementary efforts to develop K–12 educators' capacity to use performance assessments in ELA and science (Dolezal, Grube, & Watterson, 2017; Vermont Agency of Education, 2017).

Oregon has used performance assessments as part of its approach to measuring students' readiness for high school graduation for more than a decade (Darling-Hammond et al., 2005). In 2007, the Oregon State Board of Education adopted new high school graduation requirements aimed to better measure readiness for college and career, including requiring students to demonstrate proficiency in Essential Skills and complete an increased number of required credits (Oregon Department of Education, n.d.). Similar to how Vermont defines "Transferable Skills," Oregon defines "Essential Skills" as the crosscutting skills that can be applied across content and settings, including the ability to read and comprehend various kinds of text, write clearly, apply mathematics in a variety of settings, think critically and analytically, and demonstrate teamwork skills. Students have multiple options for demonstrating proficiency in Essential Skills, including through the use of "Work Samples," a representative sample of individual student work to demonstrate proficiency in one or more Essential Skills. According to state legislation, these Work Samples could include authentic performance assessments, including student work from "research papers, statistical experiments, speaking presentations, theatrical performances, [and] work experience" (Assessment of Essential Skills, 2011). Work Samples can be developed locally or drawn

from a statewide bank of sample tasks. The Work Samples are scored by teachers using a statewide rubric, and an online system is available to assist with the scoring process.

As noted, students can demonstrate proficiency in the Oregon Essential Skills in multiple ways, including through the use of scores from statewide assessments or other standardized assessments (e.g., SAT or Advanced Placement exams). Notably, all students are required to take Oregon's statewide assessments in ELA, mathematics, and science at specific grade levels for accountability purposes. In ELA and mathematics, the statewide assessment is currently administered by SBAC. School districts and public charter schools determine whether all students will also have opportunities to demonstrate proficiency using Work Samples. Consequently, in some districts, only those students who fail to demonstrate proficiency on statewide assessments have the opportunity to complete performance assessments, or Work Samples, to meet graduation requirements. The state department of education is supporting the use of performance assessments as part of graduation requirements by developing an online bank of Work Samples for each required Essential Skill, which can be used directly by educators, thus reducing the labor associated with using Work Samples.

In California, the use of performance assessments to meet graduation requirements has combined the efforts of a network of schools and districts with support from state policy. In 2011, California passed Assembly Bill 790, which authorized the Linked Learning Pilot Program. Districts had the opportunity to apply for competitive funds to develop Linked Learning pathways, career-themed high school pathways aligned with performance-based graduation requirements and developed in partnership with community organizations and businesses to better prepare students for success in life after high school.

The aim within Linked Learning pathways is to offer students an integrated approach to instruction, assessment, and curriculum that privileges engaging students in projects and tasks in which they can learn by doing and demonstrate their understanding through performance assessments that require complex thinking, problem-solving, judgment, and creativity. Tasks and projects in Linked Learning pathways are often multidisciplinary and problem based, and connections to the real world aim to be authentic and transparent (Jaquith, Martin, & Johnston, 2014, p. 4).

Specifically, students engage in projects with authentic performances and products to demonstrate competency in key areas. These opportunities for performance assessments are integrated into students' Linked Learning pathway experiences at each grade level. As part of this process, benchmarks are established at each grade level to measure progress towards graduation. Each participating district designed its own graduate profile. Importantly, the Linked Learning Pilot Program grew from and provided further support for career-themed pathways that had been in place for a decade or more in some schools and extended this opportunity to many more California students (Jaquith et al., 2014). In fact, the pilot served approximately 600,000 high school students in its first year and will be available to more than one-third of the state's high school students when it is fully implemented (Linked Learning, n.d.). By incorporating work-based learning experiences, research suggests that the Linked Learning pilot may improve student attendance and reduce dropout rates (Colley & Jamison, 1998; National Academy Foundation, 2012).

Using performance assessments as part of graduation requirements can help to assess students' readiness to apply content knowledge and essential skills to solve complex challenges, the kind of challenges they will face in college, career, and life. Furthermore, the use of performance assessments en route to graduation can help provide guidance to teachers and educational leaders about the kinds of learning experiences students need to prepare them to be successful after high school. Thus, integrating performance assessments into state graduation requirements is challenging because it requires not only changing the way students are assessed at the end of high school but

also changing instruction and curricula throughout high school to prepare students for success at graduation. Consequently, in some states, only those schools or districts with the capacity for carrying out a comprehensive performance-based graduation system are able to use performance assessments to meet graduation requirements and, in other states, using performance assessments at graduation is an option rather than a requirement. Notably, states have supported implementation and technical quality of performance tasks used for graduation purposes by providing sample tasks, student work, and scoring criteria to strengthen the validity and comparability of these tasks.

### **Strategy 3: Performance Assessments for School Accountability**

Some states have incorporated performance assessments into their systems of assessment by replacing traditional summative tests with performance tasks or including performance assessments as one measure in a larger system of assessment. As mentioned previously, NCLB and ESSA require annual testing of each student in ELA and mathematics in grades 3–8 and once in high school. They also require testing in science at three points throughout a student’s K–12 schooling. These assessments must meet a detailed list of requirements provided by the federal government, including high standards of technical reliability, validity, and comparability, but states have latitude in designing or selecting assessments within the federal parameters. The federally required tests are not usually the only assessments within a state; states also assess students in history/social studies, science, the arts, and other areas. Colorado, New Hampshire, Oregon, Vermont, Virginia, West Virginia, and Wisconsin include performance assessments alongside other, more traditional, standardized measures of student achievement in a system of assessment.

States have stayed within federal requirements in three main ways yet managed to include performance assessments in their state systems of assessment:

- Including performance tasks as one portion of statewide assessments (e.g., PARCC and SBAC exams)
- Leveraging the adoption of new standards (e.g., NGSS) to develop assessments that use performance-based items
- Allowing student achievement in subjects/grades not required by the federal government (e.g., history/social studies) to be measured through performance assessments

#### **New Consortia Assessments Include Performance Assessments**

Nearly half of the states in the United States have adopted either the PARCC<sup>1</sup> or SBAC<sup>2</sup> assessments, new collaboratively developed exams that are aligned with the Common Core State Standards. Of the ILN states reviewed in this study, Colorado has adopted PARCC and New Hampshire, Oregon, Vermont, West Virginia, and Wisconsin have adopted SBAC. Both these summative exams, which include assessments in ELA and mathematics, meet federal reporting criteria and include performance tasks. The assessments have been carefully designed by assessment experts and educators to be reliable and valid measurements of college and career readiness (Doorey & Polikoff, 2016; Schultz, Michaels, Dvorak, & Wiley, 2016; Smarter Balanced Assessment Consortium, 2016).

---

<sup>1</sup> Eight states and the District of Columbia used the PARCC annual assessment during the 2014–2015 school year (Partnership for Assessment of Readiness for College and Careers, n.d.).

<sup>2</sup> Fifteen states and the U.S. Virgin Islands are current members of the SBAC (Smarter Balanced Assessment Consortium, 2016).

The PARCC and SBAC performance tasks are administered by classroom teachers during a specified testing period and include a diverse set of stimuli, open-response items, and a writing portion that asks students to synthesize sources, conduct research, analyze data, and draw conclusions. These on-demand performance tasks, in combination with the more traditional multiple-choice portions of the PARCC and SBAC assessments, test students' content knowledge and application of skills while meeting the technical measurement standards required by the federal government. States can adopt one of these two exams—or develop their own assessment using a similar model—and use performance assessments for reporting purposes without seeking special permission from the federal government.

Significantly, the inclusion of performance tasks on these high-stakes exams has led states and districts to invest in instructional strategies that prepare students for success on performance tasks—including strategies that promote critical thinking, the application and synthesis of knowledge and skills, and text-based research and analysis. For example, the Building Educator Assessment Literacy Project provides a 2-day training during which teachers analyze student work from SBAC ELA and mathematics performance tasks, learn about the scoring process and expectations for student performance, and discuss the instructional implications of the demands of the assessment (Stanford Center for Assessment, Learning, and Equity, n.d.). Several states and districts we studied have participated in the training, including California, Oregon, and New Hampshire. California districts have used the training as a way of increasing teacher assessment literacy specifically around the expectations of the Common Core State Standards and the SBAC performance task as well as introducing the instructional changes that teachers might make to increase student success.

For many states and districts, the performance task portion of the PARCC or SBAC assessment is a very new form of assessment that has prompted significant changes throughout the system to prepare students for success. This preparation involves the instructional and curricular shifts that lead to development of transferable skills, deep content knowledge, and critical thinking. This is a significant shift that states can make within the current confines of federal policy. States that currently use state-specific, nonconsortia assessments can also use this model to include performance tasks in a summative assessment for school accountability.

### **NGSS and New Opportunities in Science**

With the introduction and adoption of the NGSS, students and teachers are refocusing science teaching and learning on deeply understanding core scientific concepts, applying scientific practices, and making connections across underlying concepts. The standards may cover fewer topics but go more deeply into specific content and focus on developing disciplinary core ideas in the context of key scientific practices and crosscutting concepts (National Research Council, 2011). Like the transition from prior state content standards to the Common Core, this transition to the NGSS represents a significant shift from previous state science standards and the ways in which they have been taught and assessed. To appropriately assess students on these new standards, states are challenged with transforming their science assessments. Not only do some stakeholders want to redesign science assessments, it is now imperative to design and adopt new assessments as existing tests are not aligned to the NGSS. This provides a timely opportunity for states to include forms of authentic, performance-based assessment as measures of science learning. Notably, some states were already including performance-based measures as part of their approach to assessment in science. The New England Common Assessment Program, used in Maine, New Hampshire, Rhode Island, and Vermont includes a combination of multiple-choice and short-answer questions as well as “an inquiry task that requires students to analyze and interpret findings from an actual science experiment” (Vermont Agency of Education, 2017).

In California, the state department of education is developing a new system of science assessment aligned to the NGSS with the support of science content experts, researchers, test providers, assessment experts, and educators. The proposed system comprises one exam in elementary, middle, and high school and includes a combination of item types including those that “deeply measure the full universe of CA NGSS Pes [performance expectations]” (California Department of Education Executive Office, 2016, p. 6). The new assessment system was approved by the California State Board of Education, and the state plans to move forward with a tentative timeline to launch a spring 2017 pilot, 2018 field test, and 2019 statewide operational system (Collier, 2016). Unlike the current science assessments in California, this new system includes more authentic forms of assessment that will measure students’ content knowledge and skills, including the ability to synthesize information, apply scientific practices, and transfer concepts and skills across content areas.

Other states, such as a group of states organized into a professional learning community through the Council of Chief State School Officers (Council of Chief State School Officers, n.d.) are exploring this opportunity in science. Just as the Common Core State Standards presented an opportunity to rethink assessment in ELA and mathematics, the adoption of the NGSS is spurring a shift in science assessment towards the inclusion of performance tasks that measure not only knowledge and skills but also the ability to apply and transfer key scientific concepts. The book *Developing Assessments for the Next Generation Science Standards*, published by the National Research Council (2014), discusses the technical considerations of including performance tasks in science assessments used for accountability and recommends several options for doing so, including an assessment comprising mixed-item formats, performance tasks including student collaboration, and performance events incorporating hands-on activities.

Efforts to integrate performance assessments in science assessments in California and other states illustrate how the adoption of new standards can create an opportunity to develop assessments that better measure students’ abilities to apply their content knowledge and skills in authentic ways.

### **Performance Tasks for State Reporting Purposes**

Many states assess their students in content areas or grade levels that are not required by federal law, such as history/social studies, science in multiple grade levels, and arts or humanities. While assessments used for federal reporting (i.e., ELA, mathematics, science) have strict technical standards and must meet a long list of requirements, state assessments not used for federal accountability often have more flexibility. In addition, local districts may have greater control over developing these assessments, allowing them to closely align assessments with their local curricula and learning objectives. For example, legislation passed in Virginia is encouraging schools and school systems to use performance assessments to evaluate student learning in “nontested” subjects and grades.

In 2014 the Virginia General Assembly passed House Bill 930, replacing five state-directed exams with local assessments (Virginia Department of Education, 2014). The Standards of Learning exams that were replaced with local alternative assessments include Grade 3 History, Grade 3 Science, Grade 5 Writing, United States History to 1865, and United States History: 1865 to the Present. The Virginia Board of Education subsequently developed guidelines to assist local districts with the transition, specifically recommending the use of performance assessments to meet local alternative assessment guidelines (Virginia Department of Education, 2014). While it is not required that districts choose to use performance assessments, the state is encouraging and supporting their use.



According to Virginia Department of Education personnel, this legislation is intended to support an incremental shift in assessment towards an increased use of performance assessments (and other “authentic assessments”) in place of more traditional multiple-choice exams. The state has sponsored several assessment workshops and issued grant applications for regional cooperative training sessions to enhance teachers’ skills in creating, implementing, and scoring local assessments (Virginia Department of Education, 2014). Additionally, the state has partnered with several technical assistance providers to build the capacity of teachers and districts to develop performance tasks and shift to instructional strategies that align with the learning assessed by performance tasks. While the current assessments affected by this policy are not assessments reported for federal accountability, the state may ultimately use locally crafted, performance-based assessments in a reformed state accountability system.

This strategy of including performance assessments where currently permitted in an overall system of assessment allows states and districts to assess students in more meaningful ways within the confines of existing policy. Some states have focused on encouraging the use of locally developed performance tasks, which may be closely integrated with curriculum and instruction, and allow schools or districts to assess locally determined student outcomes, in addition to state standards. At the same time, locally designed assessments are likely to vary in the extent to which they reliably measure students’ performances on state standards for learning.

While it can be difficult to develop assessments that include performance items while meeting technical requirements set forth by the federal government, the efforts of PARCC and SBAC demonstrate that it can be done. The cost of scoring and developing performance tasks for use as summative assessment items is comparable to the cost of developing more traditional items when shared across a consortium of states or districts (Darling-Hammond, 2014; Topol et al., 2014). Furthermore, the effects of including even one performance task as part of a summative assessment can be significant and include beneficial instructional shifts that lead to more meaningful, content-rich learning opportunities for students, a greater understanding among teachers of the expectations of proficient student performance, and decreased “test prep” activities (Ferrara, 2009; Guthrie, Almasi, Schafer, & Afflerbach, 1994; Koretz, Stecher, Klein, McCaffrey, & Diebert, 1993). For those states that choose to require state assessments in subjects other than those required by the federal government, the use of performance assessments is an opportunity to influence instruction and assessment in meaningful ways. Significantly, all of the options discussed as strategies for using performance assessments for school accountability purposes are possible for states to pursue under ESSA without seeking specific approval or exemptions from the federal government.

#### **Strategy 4: Seeking a Waiver to Allow Performance Tasks to Be Used for Federal Accountability**

Some states have sought greater flexibility in responding to federal requirements for assessment through ESEA waivers (under ESSA this strategy would include applying to be part of the Innovative Assessment and Accountability Demonstration Authority). The technical requirements and frequency of testing required by NCLB made it challenging for states to integrate performance assessments more comprehensively into their state systems of assessment. In fact, many states that had sophisticated performance assessment systems in place before NCLB abandoned these systems given the cost and technical requirements of assessing more students in more grades and reporting student-level performance annually (Pechone et al., 2010; Wei et al., 2014).

Given the delay in reauthorizing ESEA, the U.S. Department of Education allowed states to apply for ESEA flexibility, commonly referred to as an ESEA waiver, beginning in 2011 (U.S. Department of Education, 2011). ESEA waivers created opportunities for states to transition students, teachers, and schools to systems of accountability and assessment that were aligned to newly adopted college- and career-ready standards for all students. Notably, only New Hampshire was approved for a pilot to design and implement a performance-based system of assessment as part of its ESEA waiver. Colorado and Ohio were preparing to apply to implement a system of assessment that integrated performance assessments when the recent reauthorization of ESEA as ESSA (2015) passed. ESSA has the potential to further advance these efforts because it includes support for an innovative assessment pilot. Specifically, up to seven states at a given time can apply to develop and implement an innovative assessment pilot, which could include the piloting of performance-based assessments for accountability purposes. Thus, New Hampshire's approach to piloting a system that integrates performance assessments as part of its state accountability system can inform future state assessment pilots.

The New Hampshire Department of Education's ESEA waiver request was approved in August 2015 and allowed for greater flexibility in local assessment practices, with a focus on integrating local performance assessments into the state system of assessment. Importantly, New Hampshire's assessment pilot built on nearly a decade of capacity development efforts related to competency-based learning and performance assessments. New Hampshire has become a national leader in transforming its educational system to support deeper learning through competency-based education and performance assessments. In 2005, the state was the first to pass statewide regulations to create a competency-based education system and has maintained a competency-based credit system since 2009 (Freeland, 2014). As schools and districts have worked to design systems to assess mastery, incorporating performance assessments in curricula has become increasingly widespread. Scott Marion and New Hampshire Deputy Commissioner of Education Paul Leather (2015) argue that complex performance assessments are well matched with a competency-based approach to education when they are closely connected to curriculum and instruction and "can provide opportunities for teachers to develop sophisticated understandings about the nature of student learning" (p. 6).

The New Hampshire Department of Education actively supported the use of performance assessments as part of a competency-based system by first developing educator capacity to develop, use, and score performance assessments, and then seeking a waiver to reduce federal testing requirements and create a system of assessment that integrated local and common performance tasks. In 2012, the New Hampshire Department of Education established a network of 20 high schools to develop educator capacity to develop and use performance assessments to measure and support student competency through assessment literacy training, performance assessment development and peer review, regional scoring sessions, and the development of a Web-based bank of local and common performance tasks (Freeland, 2014). Annual federal testing requirements in grades 3–8 and once in high school, however, presented a challenge since these tests were disconnected from classroom instruction and measured student performance at one point in time, regardless of whether students had reached mastery.

In 2014, the New Hampshire Department of Education submitted a waiver to the federal government to reduce annual, summative statewide testing requirements under NCLB to once in elementary, middle, and high school as well as to establish the Performance Assessment of Competency Education (PACE) pilot program. Marion and Leather (2015) explain,

The PACE system is designed to foster deeper learning on the parts of students than is capable under current systems. This requires timely assessments linked closely with curriculum and instruction. The PACE system is based on a rich system of local and common (across multiple

districts) performance-based assessments that are necessary for supporting deeper learning as well as allowing students to demonstrate their competency through multiple performance assessment measures in a variety of contexts. Thus, the accountability option was established to enable schools and districts to demonstrate student achievement and learning growth through means other than or in addition to standardized tests, with an emphasis on performance assessment. (pp. 6–7)

In essence, PACE aims to connect curriculum, instruction, and assessment more closely by establishing a system of assessment that integrates locally developed and curriculum-embedded performance assessments, common (across districts) summative performance assessments, and SBAC assessments rather than relying solely on standardized end-of-year assessments to measure student progress for accountability purposes. During the school year 2016–2017, nine districts in New Hampshire were fully engaged in the PACE pilot (New Hampshire Department of Education, n.d.); the ultimate goal is for the PACE system to be implemented in districts more broadly. Importantly, the New Hampshire Department of Education worked closely with multiple organizations and assessment experts to develop educator capacity and design its system of assessment.

Developing a system of assessment that incorporates locally developed and statewide performance assessments has the potential to encourage more meaningful learning opportunities for all students by creating cohesion between curriculum, instruction, and assessment. As seen in New Hampshire, implementing such a system requires significant investments in not only assessment infrastructure but also educators' capacity to develop, use, and score these assessments. By involving educators in performance task development and scoring, this approach has the potential to support ongoing educator learning about students and the learning opportunities necessary for supporting student success. Given the need for high levels of educator capacity to implement such a system, states seeking to develop a performance-based assessment system may be more successful if they begin by piloting the system in districts and networks of schools with greater capacity to carry out performance assessments and allow additional time for growing capacity in districts or schools with lower levels of capacity to do so. Under ESSA, states can consider applying for the innovative assessment pilot to redesign their systems of assessment to support more meaningful learning through performance assessments or other means. Interest among ILN states in applying to participate in the innovative assessment pilot is high and states have considered multiple approaches, including but not limited to systems that rely extensively on performance assessments, through course assessments, and locally developed assessments.

## **Discussion and Implications**

Analyzing the policies and capacity building efforts of the 12 ILN states included in this descriptive review of state policy revealed that states have used different strategies to integrate performance assessments and were motivated to do so by different catalysts. Despite these different motivations, our review of policy documents and interviews with SEA personnel suggest that actors at the school, district, and/or state level in each state had taken steps to encourage or require the use of performance assessments as a strategy for supporting more meaningful learning for students. Each of the four strategies used to integrate performance assessments in state systems of assessment reflect a professional approach that relies extensively on growing the capacity of educators. Some states began their transition to performance assessments by encouraging the use of performance tasks for classroom use, while others established more formal policies allowing for performance assessments to be used consequentially (oftentimes as part of the PARCC or SBAC assessments) and then focused attention on the instructional shifts necessary for student success. All, however,

have found that investing in the capacity of educators to use performance tasks is essential. The overall goal of these policies is to increase the use of performance assessments in practice; therefore, teachers must have the skills and will to do so effectively. By developing the capacity of educators to use performance assessments as part of an accountability system as well as in their classrooms, states and districts may be better able to strengthen the system of curriculum, instruction, and assessment as a whole, with the intent of supporting meaningful student learning and college and career readiness for all.

Importantly, assessment systems are influenced by actors at all levels of the educational ecosystem, creating a multilayered system of assessment practices that can provide more coherent or fragmented support for performance assessments. Educators and policymakers at the classroom, school, district or network, state, and federal level all make decisions about assessment, and these decisions may support or constrain the beneficial use of performance assessments. Performance assessments in each state represented one part of a larger system of assessments, and the four strategies used to encourage or require the use of performance assessments varied in terms of the extent to which they integrated performance assessments, the degree to which they developed professional capacity, and the level of technical quality required. When states encourage the use of performance assessments for classroom purposes but not for graduation, school, or federal accountability purposes, performance tasks may be integrated in curriculum and instruction at the school level yet be disconnected from assessments administered by the district and/or state. On the other hand, including performance tasks in assessments used for accountability purposes without supporting their use in classrooms may fail to foster teacher support for this approach to assessment. Using performance assessments for classroom purposes in the states we studied seemed to represent a foundational step for developing educators' capacity to design, implement, score, and use performance assessments to support student learning. Thus, this strategy has the potential to both improve students' opportunities for growth and development and prepare educators to use performance assessments for additional consequential purposes.

By contrast, policies that encourage or require the use of performance assessments at graduation, for school accountability, or for federal reporting integrate performance assessments in the school, district, or state accountability system. This approach may create positive pressure for teachers to make instructional changes, such as increased focus on higher level thinking (Cohen & Hill, 2001; Koretz et al., 1996). However, these efforts to integrate performance assessments in the accountability system may remain disconnected from the work of teachers and students in classrooms if educators are not involved as partners in developing and implementing performance assessments as part of the system of assessment and accountability.

The use of performance assessments for accountability purposes can also raise technical concerns related to whether performance assessments, including portfolios in which teachers may design their own task or select tasks from a resource bank, are comparable measures of whether students have met expectations for learning and whether teachers are able to reliably score student work against shared expectations for performance (e.g., state standards). These technical questions are significant because valid and comparable performance assessments are needed to ensure that all students, including students with varied backgrounds and abilities and with different teachers and school contexts, have an equitable opportunity to demonstrate their knowledge and skills; students are held to the same high standards; and students, teachers, schools, and systems receive accurate information about students' learning outcomes. Designing performance tasks with clear goals for performance, using scoring guides that clarify expectations for performance and minimize the need for raters to make inferences about student work, providing sufficient, high-quality training for raters, and monitoring assessment and scoring practices over time are all essential for maintaining high-quality performance-based systems of assessment (Stecher, 2014). Educators can play an

essential role in supporting high-quality design, scoring, and use of performance assessments. Previous efforts to implement large-scale performance assessment systems suggest significant investments in assessment development and teacher training can support high-quality development, implementation, and scoring of performance assessments (Pecheone et al., 2010; Wei et al., 2014).

### **Implications for Policy and Practice**

States aiming to deepen student learning and promote instructional practices focused on the higher order thinking skills required for success in college and career might pursue the use of performance assessments through the application of one or more of the four strategies discussed in this state policy review. When considering which strategies might best achieve their desired goals, states can consider the context within their states, specifically, the historical use of performance-based assessments; the capacity of educators to design, implement, and score assessments (i.e., to use performance assessments well in support of student growth and development); the political climate; the current state assessment system; and existing state policies related to graduation. In instances where educator capacity is less developed, states might initially encourage rather than require the use of performance assessments while investing in developing professional expertise and instructional practices over time. Likewise, in states that lack the political support or capacity to use performance assessments for federal accountability, incorporating performance assessments in “nontested” subjects could be a productive strategy for developing the necessary political will for further integration of performance assessments in the larger state system of assessment. In states that are in the process of reviewing or developing graduation requirements, incorporating performance assessments into a graduation portfolio or demonstration of proficiency could be an appropriate entry point.

The state and local considerations described above will influence the strategy ultimately pursued by states and will influence the level of adoption, and perhaps effects, of the use of performance assessments within states. As noted, some states invested in performance assessment initiatives in the 1990s and early 2000s and, significantly, retain some of the practices and educator capacity developed at that time. Those key resources can be leveraged as states seek to strengthen students’ opportunities for learning through the use of performance assessments.

## References

- Assessment of Essential Skills, Or. Admin. R. 581-022-0615 (2011).
- Bill & Melinda Gates Foundation. (n.d.). Common Assignment Study. Retrieved from <http://k12education.gatesfoundation.org/2015-common-assignment-study-coky/>
- Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code development*. Thousand Oaks, CA: Sage Publications.
- Briggs, D., Diaz-Bilello, E., Peck, F., Alzen, J., Chattergoon, R., & Johnson, R. (2015, April). *Using a learning progression framework to assess and evaluate student growth*. Boulder, CO: Center for Assessment Design Research and Evaluation. Retrieved from [http://www.colorado.edu/education/sites/default/files/attached-files/CADRE.CFA-StudentGrowthReport-Final\\_0.pdf](http://www.colorado.edu/education/sites/default/files/attached-files/CADRE.CFA-StudentGrowthReport-Final_0.pdf)
- California Department of Education Executive Office. (2016). *California State Board of Education March 2016 agenda*. Sacramento, CA: Author.
- Chingos, M. M. (2012). *Strength in numbers: State spending on K–12 assessment systems*. Washington, DC: Brown Center on Education Policy at The Brookings Institution.
- Chung, G. K. W. K., & Baker, E. L. (2003). *The impact of a simulation and problem-based learning design project on student learning and teamwork skills* (CSE Technical Report 599). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Cohen, D. K., & Hill, H. C. (1998). *Instructional policy and classroom performance: The mathematics reform in California*. Philadelphia, PA: Consortium for Policy Research in Education.
- Cohen, D. K., & Hill, H. C. (2008). *Learning policy: When state education reform works*. New Haven, CT: Yale University Press.
- Colley, D. A., & Jamison, D. (1998). Post school results for youth with disabilities: Key indicators and policy implications. *Career Development for Exceptional Individuals*, 21, 145–160.
- Collier, M. (2016, March 9). State board gives green light to next phase of new science tests. *EdSource*. Retrieved from <http://edsources.org/2016/state-board-gives-green-light-to-next-phase-of-new-science-tests/561636>
- Colorado Department of Education. (n.d.). Colorado Content Collaboratives. Retrieved from <https://www.cde.state.co.us/contentcollaboratives>
- Conley, D. T. (2015). A new era for educational assessment. *Education Policy Analysis Archives*, 23(8). <http://dx.doi.org/10.14507/epaa.v23.1983>. This article is part of EPAA/AAPE's Special Series on *A New Paradigm for Educational Accountability: Accountability for Meaningful Learning*. Guest Series Edited by Dr. Linda Darling-Hammond.
- Conley, D. T., & Darling-Hammond, L. (2013). *Creating systems of assessment for deeper learning*. Stanford, CA: Stanford Center for Opportunity Policy in Education.
- Council of Chief State School Officers. (n.d.). Innovation Lab Network. Retrieved from [http://www.ccsso.org/What\\_We\\_Do/Innovation\\_Lab\\_Network.html](http://www.ccsso.org/What_We_Do/Innovation_Lab_Network.html)
- Council of Chief State School Officers. (n.d.). Science assessment item collaborative item specifications guidelines. Retrieved from [http://www.ccsso.org/Resources/Publications/Science\\_Assessment\\_Item\\_Collaborative\\_Item\\_Specifications\\_Guidelines.html](http://www.ccsso.org/Resources/Publications/Science_Assessment_Item_Collaborative_Item_Specifications_Guidelines.html)
- Darling-Hammond, L. (2014, March). Assessment for deeper learning: What's next and what will they cost? *The State Education Standard*, 14, 32–36.
- Darling-Hammond, L. (2017). *Developing and measuring higher order skills: Models for state performance assessment systems*. Retrieved from Learning Policy Institute website:

- [https://www.learningpolicyinstitute.org/sites/default/files/product-files/Models\\_State\\_Performance\\_Assessment\\_Systems\\_REPORT.pdf](https://www.learningpolicyinstitute.org/sites/default/files/product-files/Models_State_Performance_Assessment_Systems_REPORT.pdf)
- Darling-Hammond, L. & Adamson, F. (2010). *Beyond basic skills: The role of performance assessment in achieving 21st century standards of learning*. Stanford, CA: Stanford Center for Opportunity Policy in Education.
- Darling-Hammond, L. & Adamson, F. (Eds.). (2014). *Beyond the bubble test: How performance assessments support 21<sup>st</sup> century learning*. San Francisco, CA: Jossey-Bass.
- Darling-Hammond, L., & Falk, B. (2013). *Teacher learning through assessment: How student-performance assessments can support teacher learning*. Washington, DC: Center for American Progress. Retrieved from <https://www.americanprogress.org/wp-content/uploads/2013/09/TeacherLearning.pdf>
- Darling-Hammond, L., Rustique-Forrester, E., & Pecheone, R. L. (2005). *Multiple measures approaches to high school graduation*. Stanford, CA: Stanford University School Redesign Network. Retrieved from <https://edpolicy.stanford.edu/sites/default/files/publications/multiple-measures-approaches-high-school-graduation.pdf>
- Darling-Hammond, L., Wilhoit, G., & Pittenger, L. (2014). Accountability for college and career readiness: Developing a new paradigm. *Education Policy Analysis Archives*, 22(86), 1-38. <http://dx.doi.org/10.14507/epaa.v22n86.2014>
- Dolezal, L., Grube, M., & Watterson, T. B. (2017). *Strengthening and streamlining local comprehensive assessment systems: Guidelines and support for leadership teams*. Retrieved from Vermont Agency of Education website: <http://education.vermont.gov/sites/aoe/files/documents/edu-student-learning-assessments-state-assessments-strengthening-and-streamlining-local-comprehensive-assessment-systems.pdf>
- Doorey, N., & Polikoff, M. (2016). *Evaluating the content and quality of next generation assessments*. Washington, DC: Thomas B. Fordham Institute.
- Every Student Succeeds Act of 2015, Sections 1111(b)(2)(B)(vi) and 1111(b)(2)(J)) (2015).
- Ferrara, S. (2009, December). *The Maryland School Performance Assessment Program (MSPAP) 1991–2002: Political considerations*. Paper presented at the National Research Council Workshop, Washington, DC.
- Firestone, W. A., Mayrowetz, D., & Fairman, J. (1998). Performance-based assessment and instructional change: The effects of testing in Maine and Maryland. *Educational Evaluation and Policy Analysis*, 20(2), 95-113.
- Foote, M. (2005). *New York performance standards consortium: College performance study*. New York, NY: New York Performance Standards Consortium.
- Freeland, J. (2014). *From policy to practice: How competency-based education is evolving in New Hampshire*. Lexington, MA: Clayton Christensen Institute for Disruptive Innovation.
- Gewertz, C. (2014). Opting out of testing: A rising tide for states and districts? *Education Week*. Retrieved from [http://blogs.edweek.org/edweek/curriculum/2014/03/opting\\_out\\_of\\_testing.html?qs=parents+common+core+testing](http://blogs.edweek.org/edweek/curriculum/2014/03/opting_out_of_testing.html?qs=parents+common+core+testing)
- Goldberg, G. L., & Roswell, B. S. (2000). From perception to practice: The impact of teachers' scoring experience on performance-based instruction and classroom assessment. *Educational Assessment*, 6(4), 257-290.
- Grossman, T., & Shipton, S. (2012). *State strategies for awarding credit to support student learning*. Washington, DC: National Governors Association.
- Guthrie, J. T., Almasi, J. F., Schafer, W. D., & Afflerbach, P. P. (1994, April). *Policies for integrated reading instruction related to a state-wide improvement program*. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA.

- Hagopian, J. (Ed.). (2014). *More than a score: The new uprising against high-stakes testing*. Chicago, IL: Haymarket Books.
- Heilig, J. V., & Darling-Hammond, L. (2008). Accountability Texas-style: The progress and learning of urban minority students in a high-stakes testing context. *Educational Evaluation and Policy Analysis*, 30(2), 75–110.
- Hill, R. (2000, June). *A success story from Kentucky*. Paper presented at the Annual National Conference on Large-Scale Assessment, Council of Chief State School Officers, Snowbird, UT.
- Jaquith, A., Martin, D., & Johnston, J. (2014). Developing a performance assessment system from the ground up: Lessons learned from three Linked Learning pathways. Stanford, CA: Stanford Center for Opportunity Policy in Education.
- 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.
- Koretz, D. M., Barron, S., Mitchell, K. J., & Stecher, B. M. (1996). *Perceived effects of the Kentucky instruction results information district*. MR-792-PCT/FF. Santa Monica, CA: RAND.
- Koretz, D., Stecher, B., Klein, S., McCaffrey, D., & Diebert, E. (1993, December). *Can portfolios assess student performance and influence instruction? The 1991–1992 Vermont experience* (CSE Technical Report 371). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CREST).
- Linked Learning. (n.d.). Independent Evaluation Shows that Linked Learning is Working for California Students. Retrieved from <http://www.linkedlearning.org/en/independent-evaluation-shows-that-linked-learning-is-working-for-california-students/>
- Literacy Design Collaborative. (n.d.). [Literacy Design Collaborative homepage]. Retrieved from <https://ldc.org>
- Maine Department of Education. (n.d.). Implementation of proficiency-based diploma extension request for SAUs awarding diplomas. Retrieved from <http://www.maine.gov/doe/proficiency/standards/Implementprofbaseddipextreqsauawarddip.html>
- Maine Department of Education. (2014). Guiding Principles performance assessment tasks to support proficiency-based diplomas. Retrieved from <http://mainedoenews.net/2014/02/19/guiding-principles-performance-assessment-tasks-to-support-proficiency-based-diplomas>
- Marion, S., & Leather, P. (2015). Assessment and Accountability to Support Meaningful Learning. *Education Policy Analysis Archives*, 23(9), n9.
- Meier, D., & Wood, G. (Eds.). (2004). *Many children left behind: How the No Child Left Behind Act is damaging our children and our schools*. Boston, MA: Beacon Press.
- Menefee-Libey, D. J., & Kerchner, C. T. (2015). California's first year with local control finance and accountability. *Education Policy Analysis Archives*, 23(22), n22.
- Merriam, S. (1995). What can you tell from an N of 1? Issues of validity and reliability in qualitative research. *PAACE Journal of Lifelong Learning*, 4, 51–60.
- National Academy Foundation. (2012). *A guide to work-based learning: A continuum of activities and experiences*. New York, NY: Author.
- National Research Council. (2011). *A framework for K–12 science education: Practices, crosscutting concepts, and core ideas*. (p. 10). Washington, DC: The National Academies Press. Retrieved from [http://www.nap.edu/catalog.php?record\\_id=13165](http://www.nap.edu/catalog.php?record_id=13165)
- National Research Council. (2014). *Developing assessments for the next generation science standards*. Washington, DC: The National Academies Press.
- New Hampshire Department of Education. (n.d.). Performance Assessment of Competency Education (PACE). Retrieved from <http://education.nh.gov/assessment-systems/pace.htm>



- New Hampshire Department of Education. (2014). [Elementary and Secondary Education Act waiver].
- New York Performance Standards Consortium. (n.d.). *Educating for the 21<sup>st</sup> century: Data report on the New York Performance Standards Consortium*. New York, NY: Author. Retrieved from [http://performanceassessment.org/articles/DataReport\\_NY\\_PSC.pdf](http://performanceassessment.org/articles/DataReport_NY_PSC.pdf)
- Ohio Department of Education. (2012). *Expanded learning: An innovative school improvement strategy*. Columbus, OH: Author. Retrieved from <https://21stcenturylearningcenters.files.wordpress.com/2012/08/expanded-learning-web.pdf>
- Oregon Department of Education. (n.d.). Oregon diploma. Retrieved from <http://www.ode.state.or.us/search/results/?id=368>
- Osborne, D. (2016, January). Schools of the future: California's Summit Public Schools. Retrieved from Progressive Policy Institute website: [http://www.progressivepolicy.org/wp-content/uploads/2016/01/2016.01-Osborne\\_Schools-of-the-Future\\_Californias-Summit-Public-Schools.pdf](http://www.progressivepolicy.org/wp-content/uploads/2016/01/2016.01-Osborne_Schools-of-the-Future_Californias-Summit-Public-Schools.pdf)
- Parke, C. S., Lane, S., & Stone, C. A. (2006). Impact of a state performance assessment program in reading and writing. *Educational Research and Evaluation*, 12(3), 239-269.
- Partnership for Assessment of Readiness for College and Careers. (n.d.). States. Retrieved February 19, 2016, from <http://www.parcconline.org/about/states>
- Pecheone, R., Kahl, S., Hama, J., & Jaquith, A. (2010). *Through a looking glass: Lessons learned and future directions for performance assessment*. Stanford, CA: Stanford Center for Assessment, Learning, and Equity.
- Pizmony-Levy, O., & Saraisky, N. G. (2016). *Who opts out and why? Results from a national survey on opting out of standardized tests*. New York: Columbia University Academic Commons.
- Plank, S. B., & Condliffe, B. F. (2013). Pressures of the season: An examination of classroom quality and high-stakes accountability. *American Educational Research Journal*, 50(5), 1152–1182.
- Proficiency-Based Diploma Standards and Transcripts, Me. Rev. Stat. Ann. tit. 20-A, § 4722-A (2015).
- Reform Support Network. (n.d.). *Targeting growth: Using student learning objectives as a measure of educator effectiveness*. Retrieved from <https://www2.ed.gov/about/inits/ed/implementation-support-unit/tech-assist/targeting-growth.pdf>
- Renner, N., Schultz, S. E., Dieckmann, J., Martin, D., Reyes, G., Tweedy, A., ... Holthuis, N. (2015). *Curriculum audit and recommendations for Summit Public Schools*. Stanford, CA: Stanford Center for Assessment, Learning, and Equity.
- Schultz, S. R., Michaels, H. R., Dvorak, R. N., & Wiley, C. R. H. (2016). *Evaluating the content and quality of next generation high school assessments*. Alexandria, VA: Human Resources Research Organization.
- Smarter Balanced Assessment Consortium. (n.d.). Member states. Retrieved February 19, 2016, from <http://www.smarterbalanced.org/about/member-states>
- Smarter Balanced Assessment Consortium. (2016). *Smarter Balanced Assessment Consortium: 2014–2015 technical report*. Retrieved from [http://www.smarterbalanced.org/wp-content/uploads/2015/08/2014-15\\_Technical\\_Report.pdf](http://www.smarterbalanced.org/wp-content/uploads/2015/08/2014-15_Technical_Report.pdf)
- Stanford Center for Assessment, Learning, and Equity. (n.d.). Building educator assessment literacy. Retrieved from <https://scale.stanford.edu/content/building-educator-assessment-literacy>
- Stanford Center for Assessment, Learning, and Equity. (2014). Principles of performance assessment [PowerPoint slides].
- Stanford Center for Assessment, Learning, and Equity. (2015). *Educative assessment and meaningful support: 2014 edTPA administrative report*. Stanford, CA: Author.

- Stecher, B. M. (2014). Looking back: Performance assessment in an era of standards-based educational accountability. In L. Darling-Hammond & F. Adamson (Eds.), *Beyond the bubble test: How performance assessments support 21st century learning*. San Francisco, CA: Jossey-Bass.
- Stecher, B. M., Barron, S., Kaganoff, T., & Goodwin, J. (1998). *The effects of standards-based assessment on classroom practices: Results of the 1996–97 RAND Survey of Kentucky teachers of mathematics and writing* (CSE Technical Report 482). Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST); RAND Education.
- Steele, J. L., Lewis, M. W., Santibanez, L., Faxon-Mills, S., Rudnick, M., Stecher, B. M., & Hamilton, L. S. (2014). *Competency-based education in three pilot programs: Examining implementation and outcomes*. Santa Monica, CA: RAND Corporation. Retrieved from [http://www.rand.org/pubs/research\\_reports/RR732.html](http://www.rand.org/pubs/research_reports/RR732.html). Also available in print form.
- Summit Public Schools. (n.d.). Our approach: Our results. Retrieved from <http://summitps.org/approach/results>
- Topol, B., Olson, J., Roeber, E., Darling-Hammond, L., & Adamson, F. (2014). Investments in assessments of deeper learning: The costs and benefits of tests that help students learn. In L. Darling-Hammond & F. Adamson (Eds.), *Beyond the bubble test: How performance assessments support 21st century learning*. San Francisco, CA: Jossey-Bass.
- Tung, R., & Stazesky, P. (2010). *Including performance assessments in accountability systems: A review of scale-up efforts*. Boston, MA: Center for Collaborative Education.
- U.S. Department of Education. (2011). Letters from the education secretary or deputy secretary. Retrieved from <http://www2.ed.gov/policy/gen/guid/secletter/110923.html>
- Vermont Agency of Education. (n.d.). Proficiency based learning. Retrieved from <http://education.vermont.gov/proficiency-based-learning/performance-assessments>
- Vermont Agency of Education. (2017). New England common assessment program. Retrieved from <http://education.vermont.gov/data-and-reporting/educational-performance/science-assessment-results>
- Vermont Board of Education. (2014). Education quality standards: State board rule 2000. Retrieved from [http://education.vermont.gov/documents/EDU-FinalEQS\\_AsAdopted.pdf](http://education.vermont.gov/documents/EDU-FinalEQS_AsAdopted.pdf)
- Virginia Department of Education. (2014). Alternative Assessment Assistance Incentive Grant program application. Richmond, VA: Author. Retrieved from [https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&ved=0ahUKEwjh79WtoLLAhVP6WMKHVzHCaUQFggkMAE&url=http%3A%2F%2Fwww.doe.virginia.gov%2Fadministrators%2Fsuperintendents\\_memos%2F2014%2F292-14a.docx&usq=AFQjCNHI9zeBICK93IrYY5ELIcQmXSaT4w&sig2=BUSjqFq3z9K9MKEUQCSnYg](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=2&ved=0ahUKEwjh79WtoLLAhVP6WMKHVzHCaUQFggkMAE&url=http%3A%2F%2Fwww.doe.virginia.gov%2Fadministrators%2Fsuperintendents_memos%2F2014%2F292-14a.docx&usq=AFQjCNHI9zeBICK93IrYY5ELIcQmXSaT4w&sig2=BUSjqFq3z9K9MKEUQCSnYg)
- Virginia Board of Education. (2014, September). *Guidelines for local alternative assessments for 2014–2015*. Richmond, VA: Author. Retrieved from [http://townhall.virginia.gov/L/GetFile.cfm?File=C:\TownHall\docroot\GuidanceDocs\201\GDoc\\_DOE\\_5611\\_v1.pdf](http://townhall.virginia.gov/L/GetFile.cfm?File=C:\TownHall\docroot\GuidanceDocs\201\GDoc_DOE_5611_v1.pdf)
- Wei, R. C., Pecheone, R. L., & Wilczak, K. L. (2014). *Performance assessment 2.0: Lessons from large-scale policy & practice*. Stanford, CA: Stanford Center for Assessment, Learning, and Equity.
- White House, Office of the Press Secretary. (2015). *Remarks by the President at Every Student Succeeds Act signing ceremony* [Press release]. Retrieved from <https://www.whitehouse.gov/the-press-office/2015/12/10/remarks-president-every-student-succeeds-act-signing-ceremony>

## About the Authors

### **Elizabeth Leisy Stosich**

Fordham University

[estosich@fordham.edu](mailto:estosich@fordham.edu)

Elizabeth Leisy Stosich is an Assistant Professor in Educational Leadership, Administration, and Policy at Fordham University. Previously, she was a Research and Policy Fellow at the Stanford Center for Opportunity Policy in Education. Her research interests include education policy, assessment and accountability, school and district leadership, school improvement, and teachers' professional learning.

### **Jon Snyder**

Stanford University

[jdsnyder@stanford.edu](mailto:jdsnyder@stanford.edu)

Jon Snyder is the Executive Director of the Stanford Center for Opportunity Policy in Education (SCOPE). His research interests include teacher learning, conditions that support teacher learning, and the relationships between teacher and student learning.

### **Katie Wilczak**

Summit Public Schools

[katie.wilczak@gmail.com](mailto:katie.wilczak@gmail.com)

Katie Wilczak is the Director of Research Projects at Summit Public Schools. Previously, she was the Director of Learning Design & Engagement at the Stanford Center for Assessment, Learning, & Equity (SCALE). Her research interests include assessment policy, project based learning, teacher education, student engagement, and system improvement.

## About the Guest Editors

### **Elizabeth Leisy Stosich**

Fordham University

[estosich@fordham.edu](mailto:estosich@fordham.edu)

Elizabeth Leisy Stosich is an Assistant Professor in Educational Leadership, Administration, and Policy at Fordham University. Previously, she was a Research and Policy Fellow at the Stanford Center for Opportunity Policy in Education. Her research interests include education policy, assessment and accountability, school and district leadership, school improvement, and teachers' professional learning.

### **Soung Bae**

Stanford University

[soungb@stanford.edu](mailto:soungb@stanford.edu)

Soung Bae is a Senior Learning Specialist and UDL Innovation Studio Manager at the Schwab Learning Center at Stanford University. Formerly, she was a Senior Research and Policy Analyst at the Stanford Center for Opportunity Policy in Education. Her research interests focus on school accountability, student engagement, and designing learning environments that appreciate and support learner variability.

**Jon Snyder**

Stanford University

[jdsnyder@stanford.edu](mailto:jdsnyder@stanford.edu)

Jon Snyder is the Executive Director of the Stanford Center for Opportunity Policy in Education (SCOPE). His research interests include teacher learning, conditions that support teacher learning, and the relationships between teacher and student learning.

**SPECIAL ISSUE**  
**Redesigning Assessment and Accountability**

## education policy analysis archives

Volume 26 Number 13

January 29, 2018

ISSN 1068-2341



Readers are free to copy, display, and distribute this article, as long as the work is attributed to the author(s) and **Education Policy Analysis Archives**, it is distributed for non-commercial purposes only, and no alteration or transformation is made in the work. More details of this Creative Commons license are available at <http://creativecommons.org/licenses/by-nc-sa/3.0/>. All other uses must be approved by the author(s) or **EPAA**. **EPAA** is published by the Mary Lou Fulton Institute and Graduate School of Education at Arizona State University. Articles are indexed in CIRC (Clasificación Integrada de Revistas Científicas, Spain), DIALNET (Spain), [Directory of Open Access Journals](#), EBSCO Education Research Complete, ERIC, Education Full Text (H.W. Wilson), QUALIS A1 (Brazil), SCImago Journal Rank; SCOPUS, SOCOLAR (China).

Please send errata notes to Audrey Amrein-Beardsley at [Audrey.beardsley@asu.edu](mailto:Audrey.beardsley@asu.edu)

Join EPAA's Facebook community at <https://www.facebook.com/EPAAAPE> and Twitter feed @epaa\_aape.

education policy analysis archives  
editorial board

Lead Editor: **Audrey Amrein-Beardsley** (Arizona State University)

Consulting Editor: **Gustavo E. Fischman** (Arizona State University)

Associate Editors: **David Carlson, Lauren Harris, Eugene Judson, Mirka Koro-Ljungberg, Scott Marley, Iveta Silova, Maria Teresa Tatto** (Arizona State University)

**Cristina Alfaro** San Diego State University

**Gary Anderson** New York University

**Michael W. Apple** University of Wisconsin, Madison

**Jeff Bale** OISE, University of Toronto, Canada

**Aaron Bevanot** SUNY Albany

**David C. Berliner** Arizona State University

**Henry Braun** Boston College

**Casey Cobb** University of Connecticut

**Arnold Danzig** San Jose State University

**Linda Darling-Hammond** Stanford University

**Elizabeth H. DeBray** University of Georgia

**Chad d'Entremont** Rennie Center for Education Research & Policy

**John Diamond** University of Wisconsin, Madison

**Matthew Di Carlo** Albert Shanker Institute

**Sherman Dorn** Arizona State University

**Michael J. Dumas** University of California, Berkeley

**Kathy Escamilla** University of Colorado, Boulder

**Melissa Lynn Freeman** Adams State College

**Rachael Gabriel** University of Connecticut

**Amy Garrett Dikkers** University of North Carolina, Wilmington

**Gene V Glass** Arizona State University

**Ronald Glass** University of California, Santa Cruz

**Jacob P. K. Gross** University of Louisville

**Eric M. Haas** WestEd

**Julian Vasquez Heilig** California State University, Sacramento

**Kimberly Kappler Hewitt** University of North Carolina Greensboro

**Aimee Howley** Ohio University

**Steve Klees** University of Maryland

**Jaekyung Lee** SUNY Buffalo

**Jessica Nina Lester** Indiana University

**Amanda E. Lewis** University of Illinois, Chicago

**Chad R. Lochmiller** Indiana University

**Christopher Lubienski** Indiana University

**Sarah Lubienski** Indiana University

**William J. Mathis** University of Colorado, Boulder

**Michele S. Moses** University of Colorado, Boulder

**Julianne Moss** Deakin University, Australia

**Sharon Nichols** University of Texas, San Antonio

**Eric Parsons** University of Missouri-Columbia

**Amanda U. Potterton** University of Kentucky

**Susan L. Robertson** Bristol University, UK

**Gloria M. Rodriguez** University of California, Davis

**R. Anthony Rolle** University of Houston

**A. G. Rud** Washington State University

**Patricia Sánchez** University of University of Texas, San Antonio

**Janelle Scott** University of California, Berkeley

**Jack Schneider** College of the Holy Cross

**Noah Sobe** Loyola University

**Nelly P. Stromquist** University of Maryland

**Benjamin Superfine** University of Illinois, Chicago

**Adai Tefera** Virginia Commonwealth University

**Tina Trujillo** University of California, Berkeley

**Federico R. Waitoller** University of Illinois, Chicago

**Larisa Warhol** University of Connecticut

**John Weathers** University of Colorado, Colorado Springs

**Kevin Welner** University of Colorado, Boulder

**Terrence G. Wiley** Center for Applied Linguistics

**John Willinsky** Stanford University

**Jennifer R. Wolgemuth** University of South Florida

**Kyo Yamashiro** Claremont Graduate University

archivos analíticos de políticas educativas  
consejo editorial

Editor Consultor: **Gustavo E. Fischman** (Arizona State University)

Editores Asociados: **Armando Alcántara Santuario** (Universidad Nacional Autónoma de México), **Jason Beech** (Universidad de San Andrés), **Angelica Buendia** (Metropolitan Autonomous University), **Ezequiel Gomez Caride** (Pontificia Universidad Católica Argentina), **Antonio Luzon** (Universidad de Granada), **Angelica Buendia** (Metropolitan Autonomous University), **José Luis Ramírez** (Universidad de Sonora)

**Claudio Almonacid**

Universidad Metropolitana de  
Ciencias de la Educación, Chile

**Miguel Ángel Arias Ortega**

Universidad Autónoma de la Ciudad  
de México

**Xavier Besalú Costa**

Universitat de Girona, España

**Xavier Bonal Sarro** Universidad  
Autónoma de Barcelona, España

**Antonio Bolívar Boitia** Universidad  
de Granada, España

**José Joaquín Brunner** Universidad  
Diego Portales, Chile

**Damián Canales Sánchez** Instituto  
Nacional para la Evaluación de la  
Educación, México

**Gabriela de la Cruz Flores**

Universidad Nacional Autónoma de  
México

**Marco Antonio Delgado Fuentes**

Universidad Iberoamericana, México

**Inés Dussel**, DIE-CINVESTAV,  
México

**Juan Carlos González Faraco**

Universidad de Huelva, España

**María Clemente Linuesa**

Universidad de Salamanca, España

**Jaume Martínez Bonafé**

Universitat de València, España

**Alejandro Márquez Jiménez**

Instituto de Investigaciones sobre  
la Universidad y la Educación,  
UNAM, México

**María Guadalupe Olivier Tellez**,  
Universidad Pedagógica Nacional,  
México

**Miguel Pereyra** Universidad de  
Granada, España

**Mónica Pini** Universidad Nacional  
de San Martín, Argentina

**Omar Orlando Pulido Chaves**

Instituto para la Investigación  
Educativa y el Desarrollo  
Pedagógico (IDEP)

**Paula Razquin** Universidad de  
San Andrés, Argentina

**Miriam Rodríguez Vargas**

Universidad Autónoma de  
Tamaulipas, México

**José Gregorio Rodríguez**

Universidad Nacional de Colombia,  
Colombia

**Mario Rueda Beltrán** Instituto de

Investigaciones sobre la  
Universidad y la Educación,  
UNAM, México

**José Luis San Fabián Maroto**

Universidad de Oviedo,  
España

**Jurjo Torres Santomé**,

Universidad de la Coruña, España

**Yengny Marisol Silva Laya**

Universidad Iberoamericana,  
México

**Ernesto Treviño Ronzón**

Universidad Veracruzana, México

**Ernesto Treviño Villarreal**

Universidad Diego Portales  
Santiago, Chile

**Antoni Verger Planells**

Universidad Autónoma de  
Barcelona, España

arquivos analíticos de políticas educativas  
conselho editorial

Editor Consultor: **Gustavo E. Fischman** (Arizona State University)

Editores Associados: **Geovana Mendonça Lunardi Mendes** (Universidade do Estado de Santa Catarina),  
**Marcia Pletsch, Sandra Regina Sales** (Universidade Federal Rural do Rio de Janeiro)

**Almerindo Afonso**

Universidade do Minho  
Portugal

**Alexandre Fernandez Vaz**

Universidade Federal de Santa  
Catarina, Brasil

**José Augusto Pacheco**

Universidade do Minho, Portugal

**Rosanna Maria Barros Sá**

Universidade do Algarve  
Portugal

**Regina Célia Linhares Hostins**

Universidade do Vale do Itajaí,  
Brasil

**Jane Paiva**

Universidade do Estado do Rio de  
Janeiro, Brasil

**Maria Helena Bonilla**

Universidade Federal da Bahia  
Brasil

**Alfredo Macedo Gomes**

Universidade Federal de Pernambuco  
Brasil

**Paulo Alberto Santos Vieira**

Universidade do Estado de Mato  
Grosso, Brasil

**Rosa Maria Bueno Fischer**

Universidade Federal do Rio Grande  
do Sul, Brasil

**Jefferson Mainardes**

Universidade Estadual de Ponta  
Grossa, Brasil

**Fabianny de Cássia Tavares Silva**

Universidade Federal do Mato  
Grosso do Sul, Brasil

**Alice Casimiro Lopes**

Universidade do Estado do Rio de  
Janeiro, Brasil

**Jader Janer Moreira Lopes**

Universidade Federal Fluminense e  
Universidade Federal de Juiz de Fora,  
Brasil

**António Teodoro**

Universidade Lusófona  
Portugal

**Suzana Feldens Schwertner**

Centro Universitário Univates  
Brasil

**Debora Nunes**

Universidade Federal do Rio Grande  
do Norte, Brasil

**Lílian do Valle**

Universidade do Estado do Rio de  
Janeiro, Brasil

**Flávia Miller Naethe Motta**

Universidade Federal Rural do Rio de  
Janeiro, Brasil

**Alda Junqueira Marin**

Pontifícia Universidade Católica de  
São Paulo, Brasil

**Alfredo Veiga-Neto**

Universidade Federal do Rio Grande  
do Sul, Brasil

**Dalila Andrade Oliveira**

Universidade Federal de Minas  
Gerais, Brasil