Pathways to New Accountability Through the Every Student Succeeds Act

Linda Darling-Hammond, Soung Bae, Channa M. Cook-Harvey, Livia Lam, Charmaine Mercer, Anne Podolsky, and Elizabeth Leisy Stosich

April 2016

About the Learning Policy Institute

The Learning Policy Institute conducts and communicates independent, high-quality research to shape evidence-based policies that support equitable and empowering learning for every child. Nonprofit and nonpartisan, the Institute connects policymakers at the local, state, and federal level with the evidence, ideas, and actions needed to strengthen the pre-k to grade 12 education system and address the complex realities facing public schools and their communities. Working with policymakers, researchers, educators, community groups, and others who care about improving public schools, the Institute advances. More information is available at http://learningpolicyinstitute.org.

About the Stanford Center for Opportunity Policy in Education (SCOPE)

The Stanford Center for Opportunity Policy in Education (SCOPE) was founded in 2008 to address issues of educational opportunity, access, and equity in the United States and internationally. SCOPE engages faculty from across Stanford University and from other universities to work on a shared agenda of research, policy analysis, educational practice, and dissemination of ideas to improve quality and equality of education from early childhood through college. More information about SCOPE is available at https://edpolicy.stanford.edu.

LEARNING POLICY INSTITUTE

SCOPE
Stanford Center for
Opportunity Policy in Education

Pathways to New Accountability Through the Every Student Succeeds Act

Linda Darling-Hammond, Soung Bae, Channa M. Cook-Harvey,
Livia Lam, Charmaine Mercer, Anne Podolsky, and Elizabeth Leisy Stosich

Abstract

This paper examines the options available to states to redefine their accountability systems as they begin to implement the Every Student Succeeds Act (ESSA). The new law provides the possibility that states can create more balanced systems of support and accountability focused on educating young people so they can become productive, engaged citizens who are prepared for 21st century college and careers. We examine these possibilities, beginning with an overview of the law's requirements, including its allowances for indicators of school progress, methods of identifying schools for support and intervention, and requirements for the use of evidence-based interventions. We then look more closely at the range of indicators that might be considered in a multiple measures accountability system as evidence of learning, opportunities to learn, and student engagement. Next we discuss how these indicators might be combined to identify schools for intervention and support, and how they could be used within a continuous improvement system that also examines school practices through school visits and observations. We close with a discussion of research supporting evidence-based interventions that may be worth considering to support school improvement in a new accountability system.

Table of Contents

Figures and Tables

Figures

Tables

Pathways to New Accountability Through the Every Student Succeeds Act

Introduction

In December 2015, President Obama's signature reauthorized the 1965 Elementary and Secondary Education Act as the Every Student Succeeds Act (ESSA), which opened up new possibilities for how student and school success are defined and supported in American public education. One of the most notable shifts from ESSA's immediate predecessor, the No Child Left Behind Act (NCLB), is that states have greater responsibility for designing and building their state accountability systems and for determining supports and interventions for schools and districts.

For more than a decade, the federal government has attempted to drive student achievement through the use of targets and sanctions tied to a narrow definition of student success—i.e., student test scores in reading and math and, later, high school graduation rates. Although graduation rates improved during the NCLB era,[1] concern has grown that test-based accountability has resulted in a narrowing of the curriculum through an emphasis on math and reading at the expense of untested subjects like science, history, art, and music.[2] In addition, instruction has tended to focus on the format of the required multiple-choice tests. It has emphasized the recall and selection of right answers on tests of low-level skills at the expense of deeper analysis and problem-solving; research and inquiry; oral and written communications; and uses of technology and other tools to develop, evaluate, and use knowledge in real-world applications.

Despite gains on state tests that were the focus of these accountability efforts, progress slowed on the National Assessment of Educational Progress (NAEP), where the rate of gain was about half that of the pre-NCLB era. And on the Program for International Student Assessment (PISA)—a more open-ended test that evaluates how students apply their knowledge and demonstrate their reasoning—U.S. performance declined in math, reading, and science between 2000 and 2012, both absolutely and in relation to other countries.[3]

Policymakers learned that the reliance on student test scores as a measuring stick for gauging school effectiveness did not always translate into schools that were teaching students the relevant skills needed to apply knowledge to real-world situations. Instead, in many cases, improved scores signaled the greater use of test-taking strategies, rather than more durable learning; sometimes, gains were also achieved by eliminating low-scoring students from the testing pool.[4]

ESSA marks an important move toward a more holistic approach to accountability by encouraging multiple measures of school and student success. This shift creates new opportunities for local innovation by giving states the opportunity to create new approaches to accountability and improvement. The law also gives state leaders the challenging responsibility of designing systems that can address enduring inequalities in student learning opportunities and outcomes.

Principles for a New Accountability

While the approach of NCLB was problematic, its intent was to ensure that the success of traditionally underserved students mattered as much as that of other students. Now, under ESSA, states are largely responsible for creating a system that supports the success of all students. Figuring out how to use the new flexibility to achieve both greater equity and deeper learning is the challenge for the new ESSA.

ESSA eliminates NCLB's Annual Yearly Progress (AYP) system. This system set unrealistic targets for improving student performance based solely on test scores in two subjects, attached to sanctions for failing to reach those targets, thereby focusing schools' attention almost exclusively on those tests. States can now select indicators beyond those the federal government requires, including additional measures of student outcomes, school functioning, and students' opportunities to learn. They can also establish goals and determine school interventions based on their own contexts rather than adhering to a one-size-fits-all nationwide metric.

This transition could enable states to build systems of accountability that contain more robust measures of student learning and achievement aimed at preparing students for life after graduation; that provide educators with information to improve their practice; and most importantly, that support schools' capacities to reflect on and adjust their efforts to support students and educators.

This more expansive view of accountability systems is based on a perspective that accountability should be designed to help leverage improvement, not just to label or sanction schools. Given that there are multiple outcomes of schooling that we care about, and that outcomes can only be understood in relation to inputs and processes, a helpful accountability system will consider inputs, processes, and outcomes simultaneously, and enable its users to begin to understand the relationships among them, so that they can pursue useful changes. Furthermore, a productive accountability system should acknowledge that schools, districts, states, and the federal government bear different responsibilities for inputs, processes, and outcomes. Accountability strategies should be structured so that each level of the system is expected to wield the levers it controls to create equity and quality.

Achieving an equitable school system that leads to meaningful, relevant, and engaging learning opportunities for all children will require, for example, that states and districts provide funding and other resources that target the schools and districts serving high populations of students who live in poverty, have special needs, or are English language learners. It will also require that states, districts, and schools undertake the different tasks—such as curriculum design, access to materials, and educator development—that will enable these students to develop much richer learning supported by quality instruction.

A system that focuses on the whole child and the whole school requires a more comprehensive set of indicators that measure the range of skills and competencies students need to be successful upon graduating from high school. These include the mastery of core academic content; the ability to think critically, collaborate, and communicate; the development of academic mindsets; and the capacity for independent learning. The resources and conditions that support students' opportunities to learn must also be included. When systems include information about school resources and supports, the staff can readily identify inequities and respond appropriately with supports to aid improvement.

To meet these goals, some states are utilizing a framework that proposes a new paradigm for how to view educational accountability, illustrating how a hypothetical "51st state" might construct and implement policy strategies to ensure students are college-, career-, and life-ready.[5] A redesigned accountability system that accomplishes these goals for students would rest on three key pillars:

1. A focus on meaningful learning.

2. Professionally skilled and committed educators.

3. Adequate and appropriate resources that enable and support the first two pillars.

Such a system of accountability should be animated by processes for continuous evaluation and improvement that lead to problem-solving and corrective action at the local level, and that are supported by the state (see Figure 1). The system should be:

*   reciprocal and comprehensive, with each level of the system—school, district, state, and federal government—held accountable for the contributions it must make to produce an effective system;

*   focused on capacity building, including the knowledge, skills, and improvement processes needed to support high-quality education;

*   performance based in its means for gauging progress and success; and

*   informed by multiple measures that illuminate what is working and what needs to be improved or fixed.

Figure 1: **Key Elements of an Accountability System**
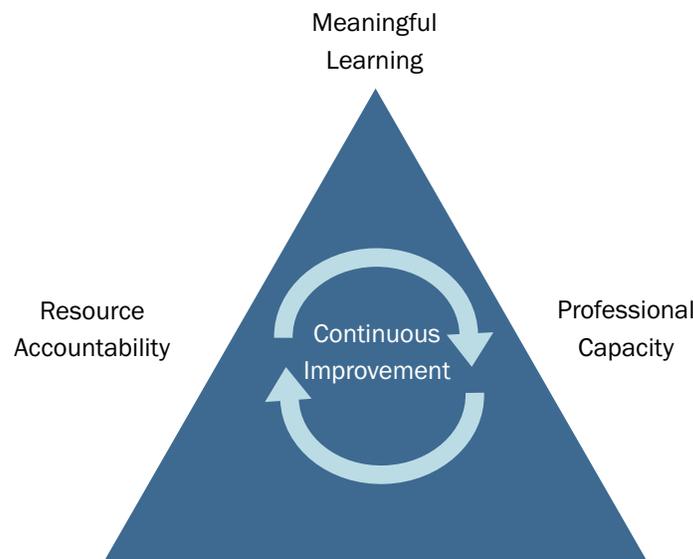
An ideal accountability system should highlight and measure both the things that matter most for student success and those that provide the most useful data and incentives for school improvement. Thus, states should thoughtfully consider how each element of their accountability system creates incentives and opportunities to move school practices forward in ways that better ensure that all students are successful in their learning and their lives beyond school. Part of this process is identifying what kinds of conditions and practices have been found to lead to better outcomes for students. Those that have greater influence on student success should have special consideration as potential indicators. Furthermore, measures of learning should seek to capture the aspects of student performance that have greatest traction for later success—including those that evaluate higher-order thinking and performance capacities.

Transparency is a critical component of any accountability system, as data drawn from well-chosen indicators provides public access to information while enabling more intelligent problem solving. Data dashboards using multiple measures can track information about inputs, processes, and outcomes to inform a diagnosis of what is and what is not working in schools and for which students. The data can be supplemented by school quality reviews that provide a qualitative look at school functioning. Such reviews can occur on a periodic basis for all schools or more frequently for schools in need of improvement to guide them in planning and developing new strategies.

Our aim in this paper is to illuminate potential approaches to developing more balanced systems of support and accountability focused on educating young people so they can become productive, engaged citizens armed with the knowledge, skills, and dispositions to participate fully in our society. We examine how these approaches can take advantage of the new flexibility under ESSA, beginning with an overview of the law's requirements and allowances for indicators, school identification, and evidence-based interventions. We then look more closely at the range of indicators that might be considered as evidence of learning, opportunities to learn, and student engagement. Next we discuss how these indicators could be used within a continuous improvement system and how they might be combined to identify schools for intervention and support. We close our paper with a discussion of research supporting evidence-based interventions that may be worth considering in a new accountability system.

What Does ESSA Require and Allow?

The concept of student learning under ESSA is much broader than it was under NCLB. States are required to implement assessments that measure "higher-order thinking skills and understanding." The law explicitly allows the use of "portfolios, projects, or extended performance tasks" as well as adaptive assessments.

In addition, states must use multiple measures of student and school performance in an accountability system they design. The act allows states wide latitude in selecting measures beyond the test scores and graduation rates that are required, limiting the authority of the Department of Education to influence decisions about indicators and how they are used. The law states that the Secretary of Education may not prescribe the indicators that states must use, the weight of measures or indicators, or the specific methodology used to differentiate or identify schools using these indicators. Thus, states have the opportunity to consider which indicators could best leverage improvements in teaching and learning.

Indicators

ESSA requires that a state's accountability system must be based on its adopted academic standards. Each state must establish long-term goals that include interim measurements of progress toward those goals for all students, as well as for certain subgroups of students—i.e., economically disadvantaged students, students from major racial ethnic groups, children with disabilities, and English language learners.

The law specifies that, at a minimum, the state's accountability system must have the following indicators:

- Academic achievement as measured by proficiency on annual assessments in English language arts and math (in each of grades 3–8, plus one grade in high school).

- Another "valid and reliable statewide academic indicator" for elementary and middle schools, which can be a measure of student growth.

- The four-year adjusted cohort graduation rate for high schools (states may add an extended-year adjusted cohort graduation rate if they choose).

- A measure of progress in English language proficiency for English language learners (in each of grades 3–8, plus one grade in high school).

- At least one measure of school quality or student success that is valid, reliable, and comparable across the state and allows for meaningful differentiation in school performance. These measures may include student engagement, student access to advanced coursework, postsecondary readiness, school climate and safety, or other measures.

The state must annually measure and report these data for all students and, separately, for each identified group of students. However, the law does not limit the number or kinds of indicators, nor does it require that the indicators be rolled up into a single index or letter grade. As we describe below, states can include multiple indicators of school quality and success to inform the identification of schools for intervention and support and for diagnostic and school improvement purposes.

Identifying Schools for Assistance

ESSA requires states to identify the "lowest performing" 5% of all public schools that receive Title I funding, all public high schools that fail to graduate one-third or more of their students, schools previously identified by the state for support and improvement that have not satisfied the state-determined criteria for exit within the period specified by the state (which shall not exceed 4 years), and any additional statewide categories that a state deems appropriate. For schools that fall into one of these categories, school districts must complete a **comprehensive support and improvement plan** that:

• is informed by the indicators and long-term goals from the state's accountability system;

• includes evidence-based interventions;

• is responsive to a school-level needs assessment; and

• identifies resource inequities that will be addressed.

In identifying schools for intensive assistance (which must be done at least once every three years), each of the first four indicators previously listed should be of "substantial weight." In the aggregate, the academic indicators must be of "much greater weight" than the other school quality indicator(s).

There are a number of ways that indicators can be given weight within the context of a multiple measures framework. For example, the state could pay greater attention to certain indicators within the set of multiple measures in the way it responds to school performance or identifies schools for additional assistance. We describe such strategies later in this paper.

The plan must then be approved by the school, district, and state education agency and periodically monitored and reviewed by the state education agency. In addition to those schools identified for comprehensive support and improvement, on an annual basis the state must identify schools where there are consistently underperforming subgroups of students. The district is then responsible for supporting the school in creating a school-level targeted support and improvement plan. Similar in structure to the comprehensive plan, the targeted support plans require evidence-based interventions and must be approved and monitored by the school district.[6]

Evidence-based Interventions

States must demonstrate that the selected interventions are evidence-based. ESSA defines as "evidence-based" an activity, strategy, or intervention that demonstrates a statistically significant effect on improving student (or other relevant) outcomes based on strong, moderate, or promising evidence from at least one well-designed and well-implemented experimental or quasi-experimental study, or a rationale based on high-quality research findings or positive evaluation which suggests the intervention is likely to improve outcomes. States have flexibility to allow schools and districts to determine which evidence-based interventions are most likely to work in which contexts and with which students.[7]

What Indicators Might States Consider?

ESSA creates opportunities for states to design accountability systems that provide a more comprehensive picture of student outcomes and opportunities to learn. Although the indicators required by ESSA reflect a minimum standard, states can take the initiative to design systems that capture more information about the factors that matter most for student success and that provide the most useful incentives for school improvement. Indicators of college and career readiness, student engagement, social-emotional supports, access to a rich curriculum, school climate and organizational functioning, and access to qualified teachers all provide information about the broader set of outcomes and opportunities that shape student success.

In this section, we provide examples of specific indicators states might consider to evaluate and support student and school progress, and we offer existing state and district examples to illustrate how these indicators are currently being used. Table 1 provides examples of three kinds of indicators that could be part of a balanced system of accountability and support: academic outcomes, opportunities to learn, and engagement. We then provide examples of how specific systems define and use these kinds of indicators.

The accountability system may include some indicators that are reported and used only to offer diagnostic information for improvement, and others that guide decision-making about schools for purposes of intervention. In some states, local districts may add indicators for their own purposes that are distinct from the state system. The indicators below could be used in any of these distinctive ways.

Indicators of Academic Outcomes

ESSA requires all states to collect and use at least two indicators of academic achievement and one indicator of English proficiency. As with NCLB, mathematics and reading/language arts assessments must be administered in grades 3 through 8 and at least once in grades 9 through 12. Science assessments must be administered at least once in grade 3–5, 6–9, and 10–12. States may choose when and how to assess any additional content areas, as long as assessment results can be broken out by subgroup and allow for meaningful differentiation in school performance.

- **Measures of academic achievement.** To measure academic achievement in mathematics, reading/language arts, and science, states may use a single summative assessment or "multiple statewide interim assessments during the course of the academic year that result in a single summative score that provides valid, reliable, and transparent information on student achievement or growth." These can include traditional sit-down tests or performance tasks. This strategy might allow schools to better integrate assessment into curriculum and to provide timely information to inform instruction. An additional benefit is that when teachers use and score performance assessments, they can also develop a deeper understanding of academic standards and their implications for classroom practice.[8]

  The law encourages states to consider including measures of performance to evaluate critical abilities—such as critical thinking, inquiry, communication, and collaboration—that are part of the new standards most states have adopted and essential for student success, but poorly measured by many traditional tests.[9] To support comparability, states can use

Table 1: **Potential Indicators for a Multiple Measures System**

Academic Outcomes	Opportunities to Learn	Engagement
Achievement on Assessments • Standardized test results, reported in terms of status and growth for individual students and/or student cohorts • Performance assessment results from common state tasks • Progress toward English language proficiency/EL reclassification rates • Students meeting college standard on AP/IB or other college-readiness tests or dual-credit college coursework **Graduation/School Progress** • 4-, 5-, and 6-year adjusted cohort graduation rates • Proportion of 8th graders who progress to 9th grade • Dropout rates **Career and College Readiness** • Students completing college preparatory coursework, approved career technical education (CTE) sequence, or both • Students meeting standard on graduation portfolios, industry-approved certificates, licenses, or badges recognized by post-secondary institutions and businesses	**Curriculum Access** • Access to a full curriculum, including science, history, and the arts, as well as reading and math • Availability of and participation in rigorous courses (e.g., college preparatory, Advanced Placement), programs, etc. • Availability of standards-based curriculum materials, technology resources **Access to Resources** • Ratios of students to counselors and specialists to students • Teacher qualifications • Safe, adequate facilities **School Climate** • Evidence from student and staff surveys about school offerings, instruction, supports, trust, and belonging **Teachers' Opportunities to Learn** • Access to and participation in professional development and support	**Student Participation** • Average daily attendance/chronic absenteeism rates • Suspension and expulsion rates • Student perceptions of belonging, safety, engagement, and school climate per student surveys **Social-Emotional Learning** • Student attitudes towards learning (e.g., academic mindset) • Indicators of social-emotional skills from student assessments • Indicators of social-emotional supports from student surveys **Parent/Community Engagement** • Indicators of participation and engagement from parent surveys **Teacher Engagement** • Indicators of participation and engagement from teacher surveys

common performance tasks or common scoring criteria with training that ensures consistency in applying standards.

Both the Smarter Balanced and PARCC (Partnership for Assessing Readiness for College and Careers) tests include short (1-3 hour) performance tasks that call on students to investigate questions, find and evaluate evidence, and use critical thinking in written analyses, as well as to engage in modeling and complex problem solving in mathematics tasks. Many countries and some states use more extended performance tasks, including projects lasting several days or weeks, to evaluate students' abilities to design and conduct investigations, frame and solve problems, and produce designs or products as well as written and oral responses. The same is true of systems like the International Baccalaureate (IB) and a subset of Advanced Placement (AP) examinations that include written essays (e.g., English, world languages, history), performance tasks (new science courses), and portfolios (art). The new AP senior seminar course includes a set of embedded performance assessments that teachers organize and score throughout the year.

> There is a substantial knowledge base about how to develop, administer, and score reliable and valid performance assessments from the United States and around the world.

There is a substantial knowledge base about how to develop, administer, and score reliable and valid performance assessments from the United States and around the world.[10] Although tests of these skills became rare in the United States during the NCLB era, many countries and provinces, ranging from the United Kingdom and Australia to Singapore and Hong Kong, use such tasks in their examination systems. Teachers trained to evaluate responses reliably score the products students produce—essays, mathematics tasks, research papers, scientific investigations, literary analyses, and artistic exhibitions.

During the 1990s, many states developed systems of assessment that included performance tasks and portfolios requiring students to demonstrate their learning through research, analysis, writing, quantitative displays, oral presentations, and uses of technology. Those who stayed with the assessments over time succeeded in developing well-designed tasks that were comparable across settings and in training teachers to score them reliably. For example, a study of Kentucky's writing portfolios, which included three writing samples in different genres, with specific guidelines and scoring rubrics for each task, found rates of agreement between teacher raters and auditors of 99% for exact or adjacent scores.[11] Studies in states including California, Kentucky, Maine, Maryland, Vermont, and Washington[12] found that these assignments improved the quality of instruction and improved achievement on both traditional standardized tests and more complex performance measures.[13]

An important tool for states wanting to develop richer tasks is the **Performance Assessment Resource Bank**,[14] which is a joint project of the Stanford Center for Assessment, Learning, and Equity (SCALE) and the Stanford Center for Opportunity Policy in Education (SCOPE), in collaboration with the Council of Chief State School Officers' (CCSSO) Innovation Lab Network. This online resource provides performance tasks within key subjects and across disciplines linked to new standards and learning progressions. Developed with educators across the states, these tasks have been piloted and vetted for quality and offered with rubrics and scoring protocols. The

site also includes portfolio frameworks, learning progressions, curriculum units in which tasks are embedded, and tools to help educators design and review tasks, and score them with consistency.

New Hampshire, Colorado, and Virginia are among the states planning to use the bank to support their performance assessment strategies. New Hampshire's Performance Assessment for Competency Education (PACE) pilot may be the most developed; it involves a growing number of districts in a combination of state and local performance-based assessments under a federal flexibility waiver from NCLB requirements. The PACE system relies on a competency-based approach to instruction, learning, and assessment. In addition to the use of the Smarter Balanced English Language Arts (ELA) and math assessments once per grade span, participating districts use state-developed common performance tasks in intervening years, supplemented by local performance tasks designed to support deeper learning.[15]

Another approach has been used in Washington, where the state provides districts and schools with common curriculum-embedded performance tasks in subjects other than ELA and math—including civics, the arts, physical education, science, social studies, and others—along with rubrics and specifications for scoring. Local districts administer the tasks and organize the scoring. These are not currently part of the accountability system but could be under ESSA.

- **Measures of high school graduation.** At the high school level, states are required to include four-year adjusted cohort graduation rates as their additional academic indicator. States may also choose to include an extended (e.g., 5- or 6-year) adjusted cohort graduation rate. This creates positive incentives for schools to admit, keep, and support students with challenges that prevent them from graduating in the standard four years. These include students who may have dropped out for a job or childrearing, been incarcerated, immigrated to the U.S. as teenagers with little previous education who need more time to catch up, or those who simply need more time to reach high standards. Including extended-year graduation rates provides an important protection against the perverse incentives that existed under NCLB for schools to exclude lower-performing students in order to boost accountability metrics.

  In middle school, states may also include an indicator for the proportion of 8th grade students who do not drop out before 9th grade—a significant statistic in many school districts—and those who are on track to graduate. This measure can serve as an early warning, triggering intervention and support for students at risk of failing to graduate or for schools with high concentrations of students who are not on track. California's CORE (California Office to Reform Education) districts, a consortium of nine districts that received an ESEA flexibility waiver from the federal government, include an "on track to graduate" indicator that measures the percentage of eighth graders who meet the following criteria, designed to predict whether students are likely to graduate from high school on time: grade point average (GPA) of 2.5 or better; attendance rate of 96% or better; no D's or F's in ELA or math; and no suspensions.[16]

- **Measures of college and career readiness.** Collecting data on students' participation and performance in college and career pathways can serve as both a measure of students' opportunities to learn and a powerful predictor of post-secondary success. Moreover, using multiple measures of college and career readiness—including participation in college preparatory coursework, dual enrollment, and work-based experiences—provides a fuller, more accurate picture than performance on state tests alone, which are less powerful predictors of success later in life.[17]

To measure college readiness, states could include:

- measures of student participation in/completion of college preparatory coursework, or the proportion of those participating in AP and IB programs;

- scores on college entrance exams, including SAT/ACT or AP/IB;

- success in dual enrollment courses (concurrent enrollment in high school and community college); and/or

- post-secondary enrollment, persistence, and graduation.

To measure career readiness, indicators could include:

- the proportion of students who complete a comprehensive sequence of courses and internships in career technical education (CTE);

- the proportion of students who complete work-based learning experiences that meet certain standards; and/or

- the proportion of students reaching a defined level of achievement as documented through graduation portfolios, industry-approved certificates, licenses, and badges recognized by post-secondary institutions and businesses.

South Carolina, for example, publicly reports both college-readiness indicators (e.g., participation and success in AP/IB programs, dual enrollment) and career-readiness indicators (e.g., the number of students enrolled in work-based learning experiences, career technology courses, attending career technology centers).[18] Two states, Virginia and Kentucky, include in their state report cards the number of students who earn industry-recognized credentials. Starting in 2015–16, Alabama will incorporate into its new accountability system career readiness indicators such as students earning an approved industry credential. The very comprehensive National Academy Foundation Student Certification Assessment System— which certifies rigorous coursework and satisfactory performance on end-of-course exams, culminating projects, a work-based learning internship, and high graduation requirements— also performs this function and could be similarly recognized.

California Department of Education reports indicators including the percent of students who have completed the state-approved (called A-G) college preparatory curriculum and those who have completed state-approved CTE sequences of courses, including internships. Some districts also recognize the proportion of students who have completed both, as students in the state's Linked Learning pathways academies do. An indicator that looks at students completing college and career pathways could serve as an important protection against re-creating a two-tiered system that limits students' post-secondary choices.

Combining college and career readiness indicators can be tricky as students have different goals for their steps immediately after high school, and different pathways to achieve those goals. Some states look at the proportion of students who have completed a college or career preparatory pathway, with the goal of assuring that 100% of students have completed at least one of these and that no students are unprepared for a productive next step. Some jurisdictions have created interesting indices that are combinations of a variety of measures. We describe these in the next section.

- **Measures of English language proficiency.** All states are required to include a measure of progress in achieving English language proficiency that compares student proficiency with the previous year. The best measures are individually administered assessments of listening and speaking as well as reading and writing within content areas, so that mastery of academic language can be assessed. Using students' gain scores along a continuum of proficiency, rather than a metric such as "percent proficient," can better assess the full range of English language learners and their needs. In turn, this can incentivize schools to provide support to English language learners at all stages of language acquisition.

  Although not required by the federal accountability system, some states have considered including an indicator for rates of English learner reclassification, as such as the one required by California's redesigned accountability system. Reclassification is the process whereby an English language learner is designated as a Fluent English Proficient student after meeting linguistic and academic criteria set by the state and district. In some settings, where English language learners are segregated for instruction, this redesignation can enable greater access to the core curriculum. In other contexts in which English language learners are mainstreamed for instruction irrespective of this designation, there often is not a change in instruction. However, once students are reclassified, students typically lose the additional support provided to English language learners and districts lose funding to provide special services for these learners.

  Just as NCLB's emphasis on "percent proficient" encouraged schools and educators to focus on "bubble kids"[19]—those students just below the proficiency threshold—tallying reclassifications could incentivize schools to focus on students who score just below the cut off of English proficiency and pay less attention to those who are newcomers. It could also create incentives to reclassify students even when they are not ready to be fully independent, thereby losing them access to needed services. Thus, such indicators should be considered with care to avoid unintended negative outcomes.

Indicators of Opportunities to Learn

ESSA requires that states include at least one other indicator of school quality or student success in addition to the two academic outcome and English proficiency indicators. By looking at students' opportunities to learn, state accountability systems can provide information about the resources and conditions that influence student learning outcomes. In addition, opportunity-to-learn outcomes can be used to hold the districts and the state accountable for providing the resources necessary for schools to meet ambitious goals for student learning. The legislation suggests that states could include indicator(s) of student engagement, educator engagement, student access to and completion of advanced coursework, postsecondary readiness, or school climate and safety. In addition, access to a full curriculum, a supportive learning environment, adequate resources, and experienced, in-field, qualified, and effective teachers are all essential for student success.

In Monroe County, Georgia, the district has adopted a comprehensive set of indicators of school conditions that influence students' opportunities to learn, including facilities quality, Internet access, new teacher retention, staff attendance, and professional learning.[20] Some of these are also used in state accountability systems, such as those in New York and California. This information can be used by educators at the school, district, and state level to ensure all students have access to equitable and adequate opportunities to learn and that educators are supported in working with students.

- **Measures of school conditions.** Surveying parents, teachers, and students can provide information about the school conditions that support student learning and emotional well-being. A rich research base exists that describes the school conditions that matter most for improving student learning, including strong principal leadership, a coherent instructional program, parent involvement, and high levels of safety and order.[21] In New York City, for example, the district surveys parents, teachers, and students on academic expectations, communication, engagement, and safety and respect. These data are reported as part of a School Quality Guide (see pp. 22-24), along with indicators of student achievement, student progress, achievement gaps, and information from a school quality review. Similar indicators are used in Alberta, Canada, and the CORE districts in California (described on pp. 15-20).

- **Measures of access to a rich curriculum.** Measures of participation in or completion of college- and career-ready pathways can provide information about students' opportunities to learn in high school and indicate whether students are on track for post-secondary success. These outcome measures also gauge opportunities to learn, and thus they do double duty in expanding educational equity. In addition, as in California, states can collect information about students' access to a full curriculum, including science, history-social studies, art, music, world language, and physical education. This could provide a corrective to the narrowing of curriculum that occurred in many schools under NCLB. It could also strengthen students' long-term performance, given the evidence that literacy and critical thinking are not generic skills but instead rely on the content knowledge students have had the opportunity to learn.[22]

- **Measures of access to resources.** A multiple measures approach can serve as part of a reciprocal accountability system that holds schools accountable for student learning while simultaneously holding the state and district accountable for providing sufficient resources to support student learning. Including indicators of key resources—adequate funding, staffing, and facilities; up-to-date, standards-based materials and technology; qualified and experienced teachers—can help to identify schools in need of support as well as inequities in resource distribution across schools.

  Access to such tools as professional learning for teachers can be an important variable associated with the expertise students ultimately encounter. Some systems examine these kinds of issues through tallies (e.g., the number of students per computer, student-teacher ratios), while others use survey data from students, teachers, or parents to report aspects of resource availability. For teachers, access to high-quality, appropriate professional development is an indicator of teaching conditions that will ultimately translate into learning conditions for students as well.

- **Measures of access to qualified teachers.** Scholars find that high-need students—low-income students, English language learners, those with low prior academic performance—are less likely to have access to highly qualified or effective teachers, whether measured by experience, training, certification for the field taught, or evaluation ratings, and are much more likely to be taught by novices and those who have not completed training.[23] These inequalities influence student achievement.[24] Schools with large numbers of inexperienced, uncertified, or out-of-field teachers place students at an acute disadvantage in their learning.[25] The same is true with respect to school principals, who are also inequitably distributed, although less data are typically available on this point. Including indicators of school-level access to qualified

educators (e.g., proportions of teachers who are fully certified, teaching in field, and have more than three years of experience) can encourage districts to pay attention to and take action to close gaps in access to qualified educators.

Indicators of Engagement

Indicators of engagement can provide information about the culture and climate of schools, including data on parent and community involvement, safe and supportive school conditions, student participation and engagement in schooling, and measures of social-emotional learning. States can choose to include indicators of engagement to meet or exceed the requirements of ESSA, which requires only one indicator of school progress or student success.

- **Measures of attendance and chronic absenteeism.** The number of days that students attend school is positively associated with student achievement, affecting their grades and achievement on reading and math assessments[26] as well as on-time graduation.[27] Including attendance as part of a multiple-measures data dashboard can motivate districts to quickly respond when student attendance stagnates or declines, overall or for specific groups.[28]

  In particular, chronic absenteeism, commonly measured as missing 10% or more of the school year, is associated with lower academic performance, lower likelihood for graduation, and increased achievement gaps in elementary, middle, and high school.[29] Furthermore, the negative effects of chronic absenteeism can increase over time, because students who are chronically absent in one year are often chronically absent in multiple years. Including this indicator in a multiple measures system, as the CORE districts do, can help to identify and provide targeted support to schools that need to develop strategies to better address the needs of chronically absent students.

- **Measures of student suspension and expulsion.** Because evidence shows that removing students from school increases their likelihood of dropping out, many districts have committed to reducing students' out-of-class time due to suspensions and expulsions, with particular attention given to the disproportionate rates of suspension experienced by students of color. The federal Office of Civil Rights uses these measures as key indicators for its civil rights enforcement activities. California, which has included this measure in its state accountability system since 2013, has seen a precipitous drop in student suspension and expulsions rates as schools have also been adopting restorative justice programs and other alternatives to suspension.[30]

- **Measures of social-emotional learning.** There has been growing interest in measuring student engagement in learning using recently developed measures of student attitudes towards learning, such as indicators of a growth mindset, grit, or persistence. California's CORE districts were some of the first in the country to experiment with using social-emotional learning indicators as part of their accountability systems.[31] After reviewing research on the social-emotional factors that promote long-term learning, the CORE districts developed a student survey and a teacher documentation system to measure four key factors considered important

for students' academic performance that were also measurable and actionable predictors of student success:[32]

1. Growth mindset.

2. Self-efficacy.

3. Self-management.

4. Social awareness.

There are two ways to think about such measures: First, using social-emotional indicators as part of a multiple measures system of accountability encourages a broader definition of student success, and second, it incentivizes schools to create opportunities for meaningful learning that will foster students' growth as self-directed learners.

- **Measures of parent engagement.** Parents and community members play an essential role in supporting student and school success. Extensive research in the Chicago Public Schools, for example, has shown that parent engagement can be a critical factor for improving student learning outcomes.[33] Consequently, the district uses parent surveys to measure engagement, thereby incentivizing schools to develop stronger ties with parents and the community.

How Might States Use Multiple Measures in Data Dashboards?

If the purpose of accountability is to monitor progress on meaningful goals and to support continuous improvement, then it is critical to have access to accurate and up-to-date information on how well a school is progressing.

Several states have developed data dashboards that use indicators (such as those described above) to report the data needed to foster continuous improvement and identify schools requiring support. Like automobile dashboards that provide information on a car's functioning—gas gauge, engine temperature, tire pressure, and fluid gauge—a data dashboard provides critical information about what is working and what needs attention. Dashboards create transparency, a key aspect of accountability, for educators and the local community. They allow data to be accessed in a timely way to guide action. Such data also help prioritize limited resources so that they address the most pressing needs, and to recognize practices that have been proven successful in advancing teaching and learning.

In this section, we provide examples of data dashboards used in the accountability systems of Alberta, Canada; California's network of CORE districts; and the New York City Department of Education. In a later section, we provide examples of different ways to evaluate and combine data for the purpose of identifying schools for assistance.

Alberta's Results Report

The Alberta Results Report is an online reporting tool that contains data for the province's seven sets of indicators, which are organized around three main goals in its multiple measures accountability system (see Figure 2). The goals and measures are as follows:

> If the purpose of accountability is to monitor progress on meaningful goals and to support continuous improvement, then it is critical to have access to accurate and up-to-date information on how well a school is progressing.

1. **High Quality Learning Opportunities:**

    a. Safe and Caring Schools. Based on surveys, this is measured by the percentage of surveyed teachers, parents, and students who agree that students are safe at school, learning to be caring individuals, learning respect for others, and treated fairly in school.

    b. Student Learning Opportunities. Based on surveys, this is measured by the percentage of teachers, parents, and students who are satisfied with the opportunity for students to receive a broad program of studies, and the percentage of teachers, parents, and students who are satisfied with the overall quality of basic education. In addition, administrative data are used to measure the percentage of students aged 14–18 registered in the K–12 system who drop out the following year and the percentage of students in the grade 10 cohort who have completed high school by the end of their third (grade 12) year.

2. **Excellence in Learner Outcomes:**

    a. Student Learning Achievement (Grades K-8). This is measured by the percentage of students who achieve the "acceptable" standard on the Provincial Achievement Test and the percentage of students who achieve the "excellence" standard on the Provincial Achievement Test.

    b. Student Learning Achievement (Grades 9-12). This is measured by the percentage of students who achieve the "acceptable" standard on a diploma exam; the percentage of students who achieve the "excellence" standard on a diploma exam; the percentage of students in the grade 10 cohort who have taken four or more diploma exams by the end of their third year in high school; and the percentage of grade 12 students who have met the eligibility criteria for a Rutherford Scholarship based on course marks in grades 10, 11, and/or 12.

    c. Preparation for Lifelong Learning, World of Work, Citizenship. This is measured by the percentage of students in the grade 10 cohort who have entered a post-secondary-level program at an Alberta post-secondary institution or registered in an Alberta apprenticeship program within six years of entering grade 10; the percentage of teachers and parents who agree that students are taught attitudes and behaviors that will make them successful at work; the percentage of teachers, parents, and students who are satisfied that students model the characteristics of active citizenship; and the percentage of teacher and parent satisfaction that students are demonstrating the knowledge, skills, and attitudes necessary for lifelong learning.

3. **Highly Responsive and Responsible Jurisdiction:**

    a. Parental Involvement. This is measured by the percentage of teachers and parents satisfied with parental involvement in decisions about their child's education.

    b. School Improvement. This is measured by the percentage of teachers, parents, and students who indicate that their schools and the schools in their jurisdiction have improved or stayed the same over the last three years.

The Results Report details how schools in each province are performing along each metric; the report is scored and color coded to indicate performance levels (i.e., red indicates very low performance, orange indicates low performance, yellow indicates intermediate performance, green indicates high performance, and blue indicates very high performance). In addition, the report provides users with information on the trends associated with each metric. For example, educators, parents, and community members are given the numeric score for how the province schools fared on the "safe and caring school" metric for that year, along with the scores for the previous year and the previous three-year average.

Figure 2: **Alberta Results Report Sample**

Accountability Pillar Overall Summary
Annual Education Results Reports - Oct 2013
Province: Alberta

Measure Category	Measure Category Evaluation	Measure	Alberta			Measure Evaluation		
			Current Result	Prev Year Result	Prev 3 Yr Average	Achievement	Improvement	Overall
Safe and Caring Schools	Excellent	Safe and Caring	89.0	88.0	88.1	Very High	Improved Significantly	Excellent
Student Learning Opportunities	Good	Program Studies	81.5	80.7	80.7	High	Improved Significantly	Good
		Education Quality	89.8	89.4	89.3	Very High	Improved Significantly	Excellent
		Drop Out Rate	3.5	3.2	3.9	High	Improved Significantly	Good
		High School Completion Rate (3yr)	74.8	74.1	72.7	High	Improved Significantly	Good
Student Learning Achievement (Grades K-9)	Issue	PAT: Acceptable	79.0	79.1	79.2	Intermediate	Declined	Issue
		PAT: Excellence	18.9	20.8	19.9	Intermediate	Declined Significantly	Issue
Student Learning Achievement (Grades 10-12)	Good	Diploma: Acceptable	84.6	83.1	82.5	High	Improved Significantly	Good
		Diploma: Excellence	21.7	20.7	20.1	High	Improved Significantly	Good
		Diploma Exam Participation Rate (4+ Exams)	56.6	56.2	54.9	High	Improved Significantly	Good
		Rutherford Scholarship Eligibility Rate (Revised)	61.3	61.5	59.4	High	Improved Significantly	Good
Preparation for Lifelong Learning, World of Work, Citizenship	Good	Transition Rate (6 year)	56.5	58.4	59.2	High	Maintained	Good
		Work Preparation	80.3	79.7	79.9	High	Improved	Good
		Citizenship	83.4	82.5	82.0	Very High	Improved Significantly	Excellent
Parental Involvement	Good	Parental Involvement	80.3	79.7	79.8	High	Improved Significantly	Good
Continuous Improvement	Excellent	School Improvement	80.6	80.0	80.0	Very High	Improved Significantly	Excellent

*Source: Alberta Education. Retrieved October 3, 2014, from https://education.alberta.ca/media/7632276/oct2013_*
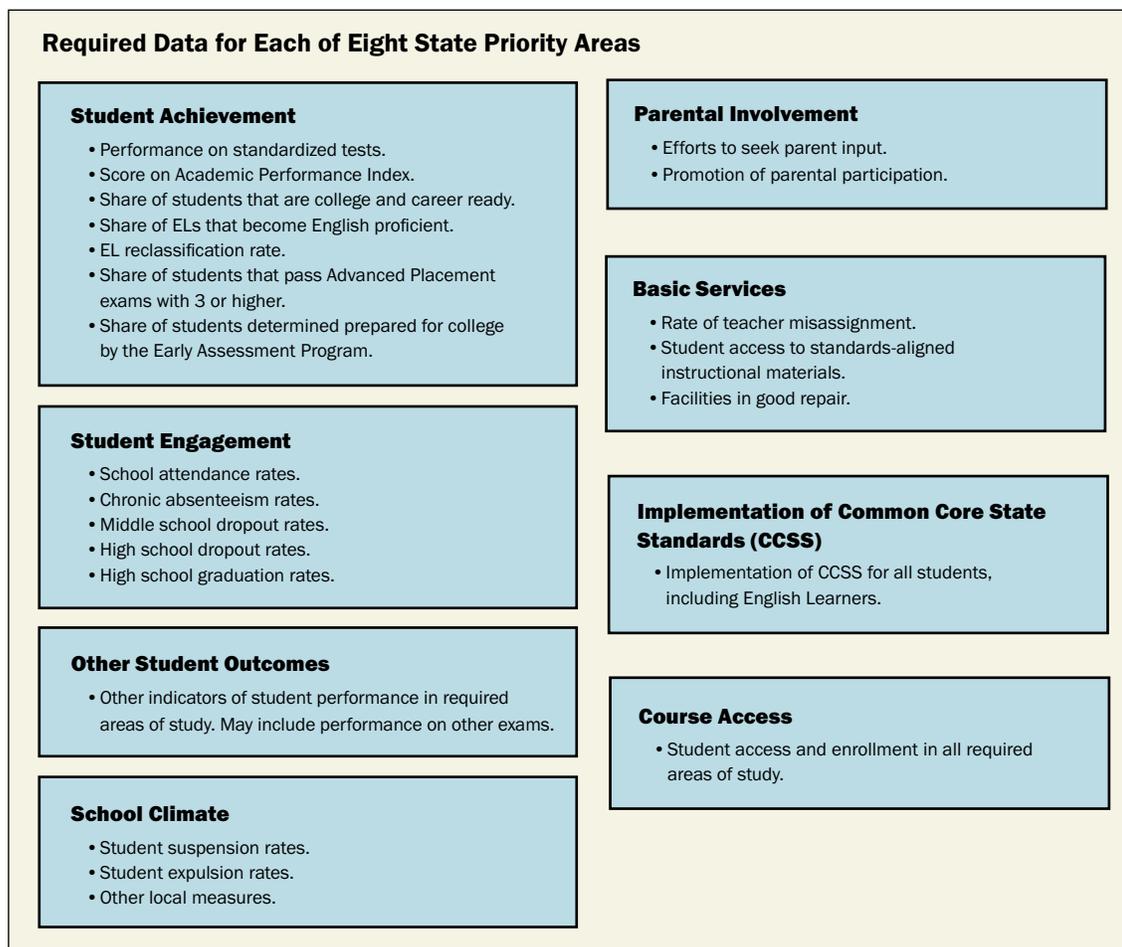*apori_201310_province_report.pdf.*

Because the Results Report is an electronic tool, users can click on each metric, which then drills down into more detailed information about the measure. For example, tables and graphs display a breakdown of how each role group (teachers, parents, students) responded to the statement that their school is safe, in addition to how parents have answered each survey item related to this measure for the last five years.

The seven sets of indicators do not roll up into a summative score; instead, the Results Report displays each individual measure and performance level. Administrators from Alberta note this is deliberate so that schools can focus on improving and prioritizing their efforts on specific areas of need that were highlighted when the performance data were disaggregated.[34]

California's State and Local Priorities

California's new accountability system is in many ways similar to Alberta's. The state has set eight priority areas, each of which includes several indicators regularly tracked by local districts (see Figure 3.) Indicators with comparable data are also tracked by the state. As part of their Local Control Accountability Plans, the districts set goals for areas they are working to improve and with county oversight annually assess their progress.

Figure 3: **Indicators for California's State Priority Areas**

Required Data for Each of Eight State Priority Areas

Student Achievement
- Performance on standardized tests.
- Score on Academic Performance Index.
- Share of students that are college and career ready.
- Share of ELs that become English proficient.
- EL reclassification rate.
- Share of students that pass Advanced Placement exams with 3 or higher.
- Share of students determined prepared for college by the Early Assessment Program.

Student Engagement
- School attendance rates.
- Chronic absenteeism rates.
- Middle school dropout rates.
- High school dropout rates.
- High school graduation rates.

Other Student Outcomes
- Other indicators of student performance in required areas of study. May include performance on other exams.

School Climate
- Student suspension rates.
- Student expulsion rates.
- Other local measures.

Parental Involvement
- Efforts to seek parent input.
- Promotion of parental participation.

Basic Services
- Rate of teacher misassignment.
- Student access to standards-aligned instructional materials.
- Facilities in good repair.

Implementation of Common Core State Standards (CCSS)
- Implementation of CCSS for all students, including English Learners.

Course Access
- Student access and enrollment in all required areas of study.

Source: Legislative Analysts Office (2013). An overview of the Local Control Funding Formula. Sacramento, CA: Mac Taylor.

The indicators include measures of:

- academic outcomes, including ELA and math test scores, English language proficiency, college-ready scores on AP tests or the state's EAP exam, and college and career readiness (e.g., the proportion of students who complete the A-G college preparatory curriculum and/or a state-approved career technical education sequence of courses and work-based experiences);

- opportunities to learn, including access to a full curriculum; access to materials, facilities, and teacher qualifications; implementation of new state standards; graduation rates; and middle and high school dropout rates; and

- engagement, including student attendance, chronic absenteeism, suspension and expulsion rates, and parent involvement.

Under ESSA, the state will design a combined state/federal reporting system based on a subset of these measures collected with comparable data statewide. These will be used to identify schools for intensive assistance. Local districts will add their own measures to guide their continuous improvement efforts. Rather than covering a few schools targeted for intervention, the state expects all schools to be working to improve the areas that are most high leverage for their overall progress.

CORE Districts' School Quality Improvement System

In California, a group of districts joined together to work as a consortium on a variety of school improvement efforts and secured an ESEA flexibility waiver from the U.S. Department of Education that shaped their collective accountability work. Building on the foundation of the California state priorities, the CORE districts add certain indicators—such as evidence of social-emotional learning and climate surveys from schools—and present a summary of key indicators in two domains:

1. Academic Achievement: growth and achievement on state English Language Arts and math tests, graduation rates, and on-track-to-graduate rate (grade 8).

2. Social-Emotional and School Culture and Climate: measures of social-emotional skills; suspension/expulsion rates; chronic absenteeism; culture/climate surveys from students, staff, and parents; and English learner re-designation rates.

These are the main pillars of the districts' School Quality Improvement System (SQIS). The CORE districts' main focus is on continuous improvement. To that end, the consortium developed a school report card that employs a data dashboard approach so that all stakeholders are able to view the school's performance on each individual measure (see Figure 4). The report card displays the performance information for each measure, detailing the previous year's score alongside the current year's score. It also calculates a change score, which represents the difference between the two scores.

Like the Alberta Results Report, the CORE districts' school report card employs color coding to indicate performance levels (i.e., green indicates above-average performance, orange indicates average performance, and red indicates below-average performance). Arrows are used to signify trends (increasing, declining, and no change). These features are designed to make the information actionable so that school improvement efforts are data based and well informed. Similar to the Alberta Results Report, the CORE districts' school report card disaggregates the data even further to paint a more nuanced picture of school performance, such as the amount of improvement that

is needed to advance an index level, how student subgroups are performing in relation to the all-students group, and how the school is performing in comparison to the district average and the CORE-wide average. An electronic version of the school report card/data dashboard is under development and expected to be released during the 2015–16 school year.

Figure 4: **CORE Districts School Quality Improvement Report Card**

SAMPLE HIGH SCHOOL

Public 3,175 students	CDS code: 01 61259 0111856	SD: 100%	AA: 8%	Ft: 0%
Principal Gerald Greenbrier	Sunny Valley Hillside District	EL: 19%	AVAN: 0%	Pt: 0%
125 Main St.		SWD: 9%	AS: 24%	WH: 4%
Valley Hill, CA 12345			HI: 63%	Two+: 1%

	Metric result 2014	**Metric Result 2015**	Change in Metric Performance from 2014 to 2015	**Index Level 2015**	Change in Index Level from 2014 to 2015
ACADEMIC DOMAIN (see pages 14 & 15 for metric descriptions)					
Academic Performance English Language Arts	—	**40% MEET STANDARDS**	—	**5**/10	—
Growth English Language Arts	Coming Fall 2016				
Academic Performance Math	—	**44% MEET STANDARDS**	—	**5**/10	
Growth Math	Coming Fall 2016				
Four Year Cohort Graduation Rate (2014 Cohort)	**84% GRADUATED**	**84% GRADUATED**	0%	**6**/10	→ 0
Five Year Cohort Graduation Rate (2013 Cohort)	**81% GRADUATED**	**86% GRADUATED**	5%	**7**/10	↗ 1
Six Year Cohort Graduation Rate (2012 Cohort)	**90% GRADUATED**	**87% GRADUATED**	-3%	**7**/10	↘ 1
SOCIAL-EMOTIONAL/CULTURE CLIMATE DOMAIN (see pages 14 & 15 for metric descriptions)					
Chronic Absenteeism	**24% CHRONICALLY ABSENT**	**21% CHRONICALLY ABSENT**	-3%	**5**/10	↗ 2
Suspension/Expulsion Rates	**8% SUSPENDED/ EXPELLED**	**8% SUSPENDED/ EXPELLED**	0%	**6**/10	→ 0
English Learner Re-designation	**10% RE-DESIGNATED**	**14% RE-DESIGNATED**		**8**/10	↗ 2
Social-Emotional Skills	Coming Fall 2016				
Culture and Climate	Coming Fall 2016				

Green = above average (Index Level 8, 9, and 10) **Orange** = average (Index Levels 4, 5, 6, 7) **Red** = below average (Index Levels 1, 2, 3)

Source: CORE districts website. Retrieved September 14, 2015, from http://coredistricts.org/wp- content/uploads/2015/09/Full_High-School-Mock-Report_updated-7_20_15.pdf.

The CORE schools and districts use the results from the broader set of state and local indicators to evaluate their progress each year and plan for ongoing improvements. The report card is a means to present some of the evidence in a dashboard for parents and the public. These data are also used to identify schools for additional assistance, as discussed next.

New York City's School Quality Guide

The New York City Department of Education (NYCDOE) developed the School Quality Snapshot and the School Quality Guide as a way to inform parents, students, and educators about each school's performance on a range of indicators that are aligned to their educational priorities.[35]

The School Quality Guide reports on seven domains of interest:

1. **Student Achievement** at the elementary school level, is measured by student growth and performance on state tests, 6th grade core courses pass rates by former 5th graders; at the middle school level, is measured by student growth and performance on state tests, core courses pass rates, high school credit earned by 8th graders, and 9th grade credit accumulation by former 8th graders; at the high school level, is measured by high school credit accumulation, data on student performance on Regents exams, 4-year and 6-year graduation rates, college and career readiness (i.e., percentage of students who complete approved college or career preparatory courses and assessments, percentage of students who met CUNY's standards for avoiding remedial classes, and percentage of students who graduate from high school and enroll in college or other postsecondary program within 6 months), and postsecondary enrollment.

2. **Rigorous Instruction** measures the extent to which curriculum and instruction are designed to engage students and foster critical thinking skills, and are aligned to the Common Core State Standards. It is measured by data from the NYC School Quality Review (SQR) (i.e., how interesting and challenging is the curriculum, how effective is the teaching and learning, and how well does the school assess student learning) and the average percent of positive responses to the NYC School Survey questions related to rigorous instruction.

3. **Collaborative Teachers** measures the extent to which teachers participate in opportunities to develop and grow their competencies and contribute to the continuous improvement of the school community; and is measured by data from the NYC SQR (i.e., how well do teachers work together) and from the NYC School Survey (e.g., the percentage of teachers who report that they work together to design instructional programs).

4. **Supportive Environment** measures the extent to which the school establishes a climate in which students feel safe, challenged to grow, and supported to meet high expectations. It is measured by data from the NYC SQR (i.e., how clearly are high expectations communicated to students and staff), the NYC School Survey questions related to school safety and feeling supported (e.g., percentage of students who feel safe in the hallways, bathrooms, locker room, and cafeteria); and administrative data on student attendance and movement of students with disabilities to the least restrictive environment.

5. **Effective School Leadership** measures the extent to which the school leadership inspires the school community with a clear instructional vision and effectively distributes leadership to realize this vision. It is measured by data from the NYC School Survey (e.g., percentage of teachers who say that the principal communicates a clear vision for the school, percentage of parents who feel that the principal works to create a sense of community in the school).

6. **Strong Family-Community Ties** measures the extent to which the school forms effective partnerships with families and outside organizations to improve the school. It is measured by data from the NYC School Survey (e.g., percentage of parents that report the school staff regularly

communicate with them about how staff can help their children learn; percentage of teachers who report that teachers at the school work closely with families to meet students' needs).

7. **Trust** measures the extent to which relationships between administrators, educators, students, and families are based on trust and respect. It is measured by data from the NYC School Survey (e.g., percentage of teachers who report that they trust the principal; percentage of parents who report that the school staff work hard to build trusting relationships with them; percentage of students who report that teachers treat them with respect).
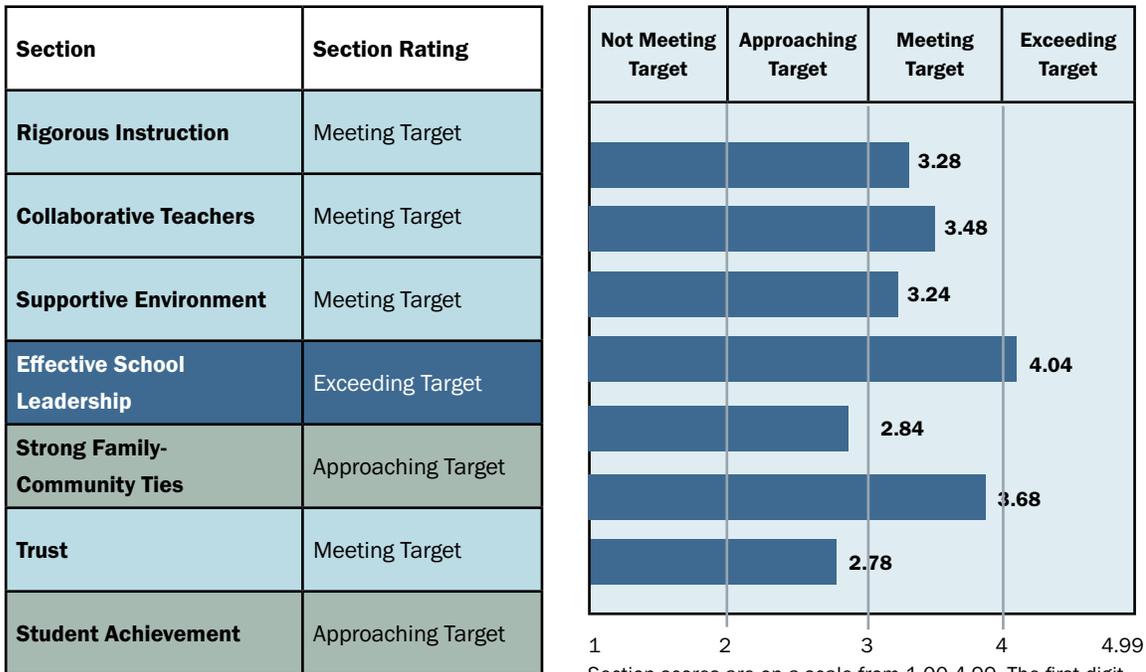
The set of measures contributes to a section score that determines the performance level for each domain. In the School Quality Guide, the performance levels for each domain range from not meeting target to approaching target to meeting target to exceeding target (see Figure 5).

Figure 5: **NYCDOE School Quality Guide, Summary Page**

Framework for Great Schools   Archer Elementary School

The Framework consists of six elements—Rigorous Instruction, Collaborative Teachers, Supportive Environment, Effective School Leadership, Strong Family-Community Ties, and Trust—that drive Student Achievement. The School Quality Guide shares ratings and data on each of the Framework elements, based on information from Quality Reviews, the NYC School Survey, student attendance, and movement of students with disabilities to less restrictive environments. The School Quality Guide also shares ratings and data on Student Achievement based on a variety of quantitative measures of student growth and performance.

Section	Section Rating
Rigorous Instruction	Meeting Target
Collaborative Teachers	Meeting Target
Supportive Environment	Meeting Target
Effective School Leadership	Exceeding Target
Strong Family-Community Ties	Approaching Target
Trust	Meeting Target
Student Achievement	Approaching Target

Not Meeting Target	Approaching Target	Meeting Target	Exceeding Target
		3.28	
		3.48	
		3.24	
			4.04
	2.84		
		3.68	
	2.78		

1   2   3   4   4.99

Section scores are on a scale from 1.00-4.99. The first digit corresponds to the section rating, and the additional digits show how close the school was to the next rating level.

State Accountability Status: Good Standing
This designation is determined by the New York State Department of Education. More information on New York State accountability can be found here: http://schools.nyc.gov/Accountability/tools/accountability/default.htm

Source: NYCDOE School Quality Reports. Retrieved March 15, 2016, from http://schools.nyc.gov/Accountability/tools/report/default.htm.

In efforts to set fair and thoughtful targets for the accountability metrics, the School Quality Guide aligns performance targets for each school based on a comparison group of similar students across the city.[36] Each student at a school is matched to the 50 most similar students from other schools throughout the city, based on prior test scores and demographic data. In addition, the School Quality Guide presents the NYC School Survey results so that the school's data may be compared with the city's range of scores (i.e., minimum score, average score, and maximum score). This information allows stakeholders to contextualize the school's survey data and gain a more nuanced understanding of how the school is progressing on the domains of interest.

As with the Alberta Results Report, the scores for each domain in the School Quality Guide are not rolled up into one summative score. Rather, the results of each measure are viewed separately and are more likely to reveal areas of strength as well as weakness and facilitate the use of data-based decision-making to inform school improvement efforts.

How Might Data Be Combined for Decision-making?

States and districts that have developed dashboards use the various measures to identify progress and needs in different aspects of school functioning, both for students overall and for identified groups of students, and to plan for appropriate, targeted improvement strategies. This planning is often further informed by additional diagnostic processes such as school or program quality reviews, described further in the next section.

The multiple measures approach supports more strategic interventions than those informed only by a unidimensional rating, ranking, or grade. It is far less clear what to work on if educators know only that their school has scored a 780 on a performance index or a C+ on a grading scheme, than if they know they are making strong progress on test scores but struggling with graduation rates for Latino students. With more detailed information, educators can focus their efforts on what really needs attention.

How can states preserve the benefits of a multiple measures approach under ESSA? The law requires states to develop processes for identifying and supporting the lowest performing schools (the "bottom 5 percent" of Title I schools) and those with sustained equity gaps. Although ESSA states that the set of academic measures must have greater weight than other non-academic measures in making the determination, this does not mean that a unidimensional index or grading scheme must be used as the foundation of the accountability system.

The law's requirements could be met in the context of a multiple measures system through the use of decision rules at the time the identification of schools must be made (at least once every three years). This does not require using a single dimensional ranking system for schools as the main foundation of the accountability system. In a system where the goal is effective support for improvement for all schools, a dashboard approach is much more likely to support productive diagnosis and planning than an approach that ranks all the schools against each other but provides little information about where they can improve.

Two examples are CORE's weighted measures approach, which is used under the districts' federal waiver to identify schools for intensive assistance, and a "decision rules" approach emerging in California's new accountability policy.

A Weighted Measures Approach

One approach to preserve a multiple measures dashboard is to apply weights to specific measures for the purpose of identifying lowest performing schools at the time of decision-making, without reducing the measures to a composite index for other purposes. An analogy for such a technique would be providing parents with a student's report card for each subject area throughout all grade levels, and then calculating an overall GPA at the end of high school for the college transcript. Colleges also receive the information about all those courses and grades over the students' high school years so that they can both evaluate specific strengths, weaknesses, and trends, and see how these roll up into a single measure.

The CORE districts' SQIS provides an illustrative example of this method. The SQIS emphasizes a multiple measures approach to accountability and includes the following indicators: academic achievement and growth; graduation rates; high school readiness of 8th graders; social-emotional and culture-climate factors which include chronic absenteeism, social-emotional skills, suspension/expulsion rates, surveys of the student, staff, and parent body on school culture and climate; and English language learner re-designation rates.

As part of the SQIS, the CORE districts developed the School Quality Improvement Index that applies weights to selected measures in the two domains in order to meet federal requirements for identifying low-performing schools. The academic achievement domain comprises 60% of the total and the social-emotional and school culture and climate domain comprises the remaining 40%. When schools need to be identified for intensive assistance, a score is calculated for each school, and those with the lowest scores are thereby identified (see Figure 6).

Once this identification is made, districts conduct school quality reviews to determine the nature of the supports needed. This remains the key tool for informing ongoing improvement for these and all the other schools in the network. The weighting of measures is used only at a moment in time to identify schools for assistance, and not as an ongoing strategy to rank schools against each other.

Figure 6: **Domain Weights for CORE School Quality Improvement Systems**
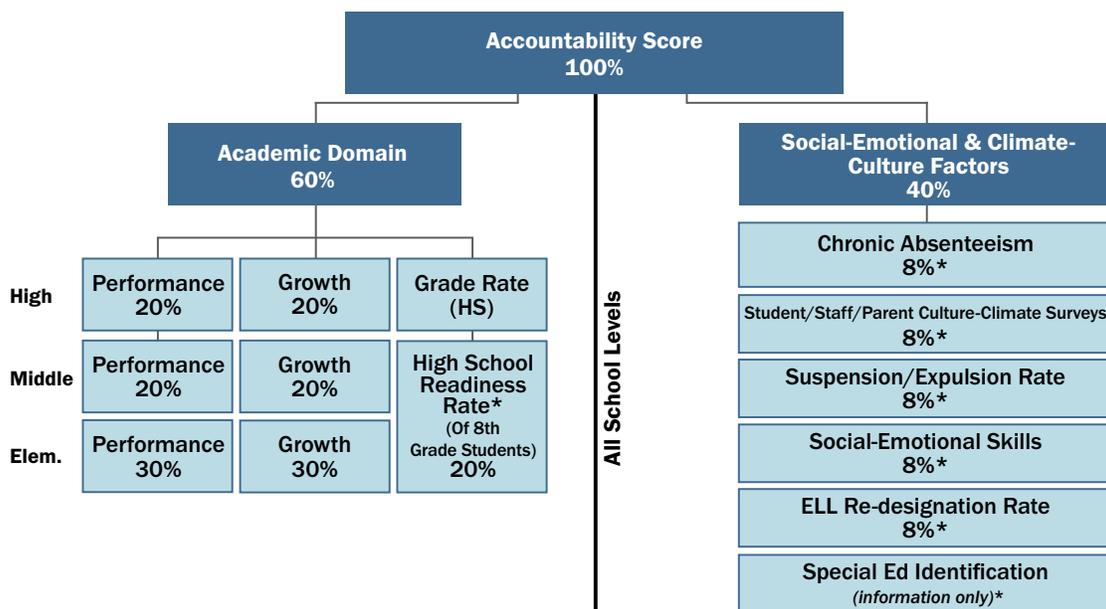
Source: CORE's Accountability System (N. Bookman, personal communication, January 27, 2015).

The data from the dashboard is designed to track students' progress related to each metric and to signal the types of interventions and continuous improvement strategies needed in the schools. Thus, when a school falls below certain thresholds (for example, a targeted rate for eighth graders who are on track to graduate), this triggers district attention, and improvement resources may be allocated to the school regardless of the school's overall accountability score.

> Another approach is to use decision rules to identify schools based on the multiple measures in the dashboard.

Use of Decision Rules

Another approach is to use decision rules to identify schools based on the multiple measures in the dashboard. In California, the Public School Performance Accountability Act of 1999 established criteria for determining low-performing and high-performing schools, along with a system of assistance and consequences.[37] Schools were identified for assistance if they did not meet their annual improvement targets for three years in a row. At a certain level of high attainment, schools had met a threshold that no longer required the same rate of annual growth overall, as long as they remained above the threshold.

In the context of a multiple measures dashboard, similar decision rules could be used, with academic outcome indicators carrying significant weight based upon their prominence in the decision process. For example, if a school or district—or a student population within a school or district—failed to make progress or reach a threshold in two of several academic outcomes indicators over a period of time, or failed to make progress on any three of several other indicators (e.g. school climate, opportunities to learn, engagement), the school or district could be identified for intensive assistance. The nature of the assistance would focus the areas of need represented by those indicators and any other key aspects of school or district functioning revealed by the review process. Other decision rules are, of course, possible.

In current conversations to shape California's new accountability system, Children Now, a statewide policy and advocacy organization, has proposed a method of interpreting data that allows simultaneous consideration of performance and improvement on any given indicator, overall and relative to other districts, by analysis of a scatter plot on a two-dimensional graph. In Figure 7, each circle represents a separate school district (the size of the circle reflects district enrollment); the location on the scatter plot represents performance on the x-axis, and growth on the y-axis.

Presented in this way, the information allows schools or districts to see themselves in relation to others and in relation to a standard as indicated by the green lines. For example, Big Tree School District can see that its graduation rate, as indicated by the location on the x-axis, is about 77%, which meets a minimal standard but not the state's ultimate goal of 90% (shown by the green line). This rate is growing slowly, having increased by about 1 percent—as indicated by its place on the y-axis—and is improving at a slower rate than most other districts at that graduation rate level, as shown by the position of the district in relation to others just above and below it.

The red lines on the graph delineate areas of concern with respect to performance or improvement. In the lower left quadrant, schools or districts that are both low-performing and not improving would be identified for intensive interventions and support.

Figure 7: **Graduation Rates and Growth in Graduation Rates for Districts**
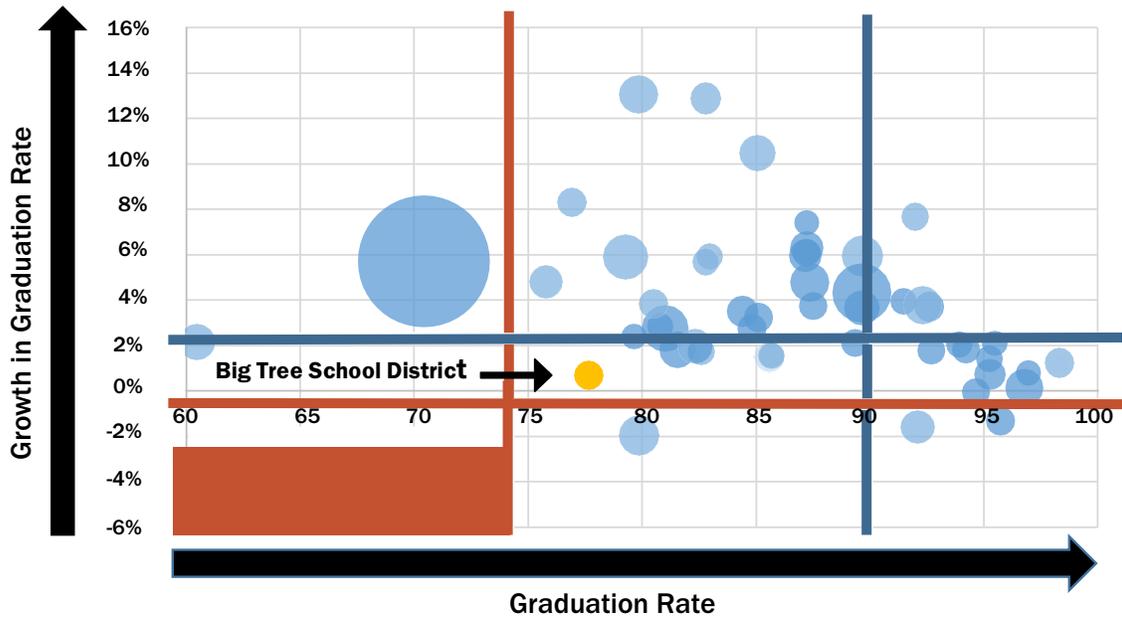
Source: Children Now

These data can be further disaggregated by student group, which permits a school or district to see performance and improvement in relation to a particular student population and, at the same time, in relation to other schools' or districts' success with that population. In Figure 8, our fictional Big Tree School District has improved its graduation rate for special education students by 5%. The district could identify other districts that have had stronger success and/or improvement with this population to learn more about what strategies could be effective.

Figure 8: **Grade Rate/Growth in Graduation Rate for Special Education**
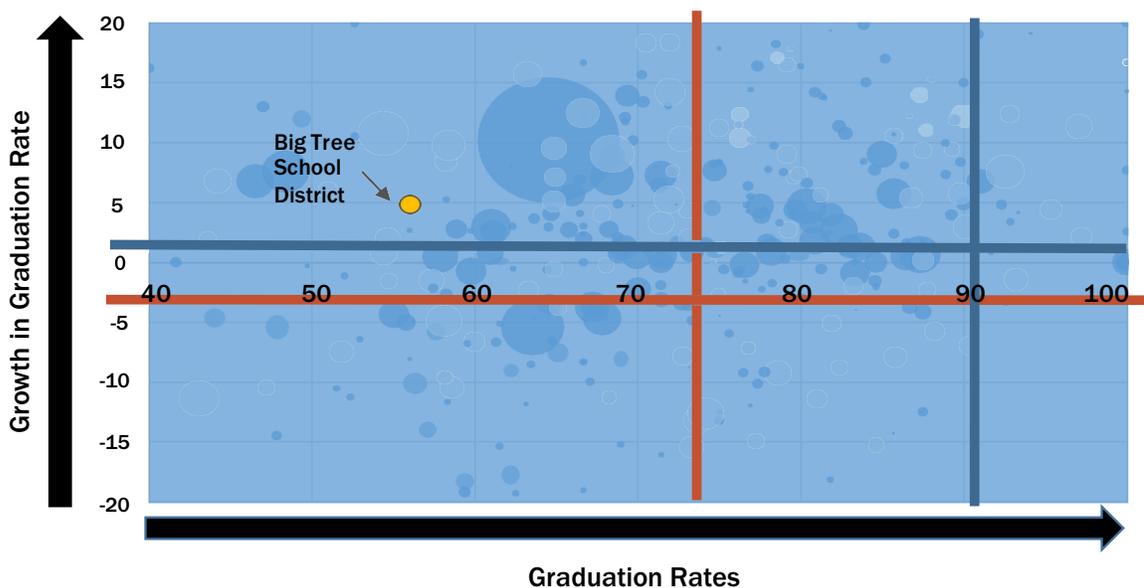
Source: Children Now

The state can look at data on each of the key indicators (overall and by student subgroup) and identify schools that fall in the lower left quadrant—signifying low performing and not improving—on multiple measures and for different subgroups of students. As this analysis is conducted, measures of academic outcomes can be more prominently considered.

A critical point is that schools and districts in a continuous improvement cycle can identify any areas where they want to improve and, ideally, access state resources and reach out to schools that have experienced success to help them improve on this indicator. A system like California's would produce these data for many indicators, such as graduation rates; assessments of ELA, math, science, and English language learner proficiency gains; student completion of college- and career-ready curriculum; attendance and chronic absenteeism; and suspensions and expulsions. The data could help schools and districts identify areas for focus, identify other schools making strong gains, and allow the state to recognize and study successful efforts to share new knowledge with others.

> Schools and districts in a continuous improvement cycle can identify any areas where they want to improve and, ideally, access state resources and reach out to schools that have experienced success to help them improve on this indicator.

The state might create learning and improvement resources around specific needs—for example, strengthening programs for English language learners—and identify districts or schools that need particular help in this area. It might assemble research evidence about what works and study successful schools and districts that have made strong gains. Building on this knowledge, the state might create supported learning networks to create stronger programs, access better materials, engage in professional development for teachers and leaders, work with instructional coaches, visit and study successful schools/districts, and monitor progress in English language learning. A similar process might occur with schools or districts that are struggling to make progress in mathematics, or those that are seeking to reduce high or discriminatory suspension rates.

Across domains, the state might offer assistance to more than 5% of its schools that receive intensive assistance in one or more arenas. But the assistance would be targeted and well supported, and conducted so as to build a learning community that takes advantage of the collaboration that can spur more progress. For schools or districts that are low-performing and failing to improve in many areas, a more comprehensive approach to intervention would be called for, and the state could create a set of resource and interventions to ensure that highly trained experts are made available to evaluate the dysfunctions, get to the root causes, and determine changes and strategic investments that are needed.

How Might Diagnostic Systems Be Developed?

To move accountability systems from a compliance orientation to one centered on continuous improvement and learning, redesigned school accountability systems must include rich sources of diagnostic information. In their efforts to improve student learning, decision makers need to know whether and how different learners are provided the opportunities and supports they need to be successful, including high quality instruction, a rigorous and meaningful curriculum, a positive school culture and climate, and strong school leadership. Without this context information, student performance scores are less actionable.

School reviews can generate the needed contextual, qualitative information to better understand the quality of teaching and learning and to promote continuous improvement. A number of states are designing and engaging in the following approaches to school reviews:

- annual planning and review;

- program review;

- school quality review; and

- diagnostic review.

Annual planning and review is a continuous improvement process that California has developed as part of its initiative to provide local education agencies with more control over school budgeting while strengthening local accountability. The goal of the annual planning and review is to inject more transparency into how funds are spent as well as the practices and strategies that the district will implement to improve teaching and learning in schools.[38]

A **program review** is a process whereby school staff document, analyze, and reflect on components of an instructional program to identify areas of strength and needed improvement. A program review serves multiple purposes: to improve the quality of teaching and learning, allow equal access to opportunities to learn for all students, promote performance-based assessments of student understanding, and ensure a coherent integration of relevant skills across all content areas throughout the school.[39]

The **School Quality Review (SQR)** is a formal process for evaluating teaching and learning. The SQR typically involves a school self-assessment, followed by a site visit by expert and peer reviewers who develop qualitative insights by observing teaching, reviewing student work, and talking with stakeholders. This review is often designed to complement the quantitative information provided through data dashboards. Moreover, the findings garnered from the SQR provide educators and administrators with actionable information to develop school improvement plans and assist stakeholders in prioritizing and targeting resources to provide learning supports and build local capacity.

Diagnostic reviews function similarly to SQRs and serve similar purposes. The difference, however, lies in the segment of schools to which the reviews are applied. Whereas SQRs have been employed traditionally as a continuous improvement model for all schools within a school system (e.g., NYC schools, Massachusetts charter schools), diagnostic reviews have been promoted for use primarily with low-performing schools within a larger school system to create a firmer connection between accountability determinations and appropriate and meaningful supports and interventions. These reviews are typically conducted by organizations external to the school system.

Detailed below are examples of state approaches to building continuous improvement systems. We highlight:

- California's annual planning and review process;

- Kentucky's program review process;

- SQR processes used in New York City, Vermont, Massachusetts, New Hampshire, and California; and

- Diagnostic reviews used in Kentucky, Massachusetts, and Connecticut.

Annual Planning and Review

California has incorporated into its new accountability system processes for engaging community members and other local stakeholders, including the ability to participate in the decision-making process for setting goals, determining budgets, and evaluating progress to guide ongoing improvement efforts. Community engagement is a centerpiece of the Local Control and Accountability Plan (LCAP) development process. The LCAP is a three-year plan that outlines the district's annual goals and the strategies that will be used to achieve the stated goals.[40] In addition, the district goals must address the state's eight priority areas, which can be categorized around three focus areas: conditions of learning, pupil outcomes, and engagement. The plans must be updated annually, and parents and community members are actively recruited to participate in the plan development process and provide input.

Each year, the district's implementation of the LCAP is reviewed through both self-assessment and county-level review. The LCAP rubric is being developed by the State Board of Education around the state priorities and additional local indicators. The rubric is designed to assess progress and prompt active reflection on areas of strength and needed growth. Because the LCAP is a planning tool, it allows district personnel, parents, and community members to continuously monitor the district's progress toward its goals and to promote continuous improvement.

As part of its LCAP process, the CORE network supports ongoing improvements by facilitating professional learning activities across districts. For schools identified as under-performing, the CORE process includes additional diagnosis through a SQR, followed by provision of needed resources and professional development as part of a commitment to build professional capacity in those schools. One of several strategies for transforming low-performing schools is based on a highly successful approach developed in Shanghai, China. The method pairs demographically similar high- and low-performing schools based on specific areas of strengths and weaknesses. Following an initial needs assessment of the low-performing school, educators from both schools share ideas and intervention practices to design and implement an improvement plan.

Program Review

Kentucky's redesign of its accountability system began in 2009, when the Kentucky legislature set the course for a new assessment and accountability system titled Unbridled Learning: College/Career-Readiness for All. The Unbridled Learning system is a balanced approach to public school accountability based on multiple measures of school effectiveness. The system comprises three strategic priorities: next-generation learners, next-generation instructional programs and support, and next-generation professionals.[41]

The state priority for next-generation instructional programs and supports is based on reviews that function as systematic analyses of an "instructional program, including instructional practices, aligned and enacted curriculum, student work samples, formative and summative assessments, professional development and support services, and administrative support and monitoring."[42] Intended to identify areas of strengths and needs, the reviews are conducted by committees made up of school personnel (i.e., subject-matter teachers, teachers from other disciplines, and school leaders) and representatives from parent, student, and business communities.

Program reviews have been developed for arts and humanities, practical living and career studies, writing, and K-3 early learning. A world language program review was piloted during the 2014–15 school year with high school students. High schools will be held accountable for the world language program review in the 2015–16 school year, and elementary and middle schools will be included in the following academic year. Currently, next-generation instructional programs and supports account for 23% of a school's overall school accountability score.

School Quality Review

In **New York City**, the SQR is used to examine the quality of instruction in schools and to provide the school community an opportunity to reflect on and self-evaluate its progress and improvement efforts.[43] The review focuses on instructional and organizational coherence as keys to improving student learning and evaluates the school's work as it relates to the following five quality indicators:

1. Rigorous, engaging and coherent curricula aligned to the Common Core Learning Standards (CCLS).

2. Research-based, effective instruction that yields high quality student work.

3. Curricula-aligned assessment practices that inform instruction.

4. Establishment of a culture of learning that communicates high expectations with supports.

5. Engagement of teacher teams in collaborative practice using the inquiry approach to improve classroom practice.[44]

The review process begins with the school completing a self-evaluation form that provides the contextual background for review and affords the staff an opportunity to reflect upon and assess the school's functioning. Although an earlier version of the review developed by the state featured teams of peer reviewers led by an expert reviewer, the current New York City version is less extensive, typically including only one reviewer.[45]

During the process, the reviewer observes classrooms and teacher team meetings; meets with the school administrator, students, parents, teachers, and the teacher union chapter leader; and examines curricular artifacts and other school-related documents. The reviewer must use specific tools such as the Quality Review rubric,[46] which describes the quality "look-fors" in each grade; the Classroom Visitation Tool,[47] an evidence-gathering document; and a Record Book, used to document and organize the evidence collected during the site visit. At the conclusion, the reviewer provides verbal feedback to the principal and the school leadership team on the school's preliminary rating, emphasizing areas of strength and areas for future focus, with specific examples gathered from the visit.

After a quality assurance process, a final report detailing the findings is published on the New York City Department of Education's website. A score derived from the quality review is reported on the school report card. The report is then used as a lens through which to better understand and interpret the quantitative data (e.g., student test scores, school climate survey, gap scores) while informing improvement efforts at the school.

> In Vermont ... the standards focus on assuring that all schools deliver proficiency-based learning, flexible pathways to graduation, safe school environments, high-quality staffing, and financially efficient practices.

In **Vermont**, the State Board of Education adopted the Education Quality Standards (EQS) in 2014[48] to ensure that all students have access to high-quality educational programming that is substantially equal throughout the state. The standards focus on assuring that all schools deliver to their students proficiency-based learning, flexible pathways to graduation, safe school environments, high-quality staffing, and financially efficient practices. To evaluate schools' delivery of the standards, the Vermont Agency of Education has begun developing the Education Quality Review (EQR), which is the tool with which the state and the public will measure student learning and school progress.[49]

The EQR consists of two complementary processes to drive continuous improvement: the Annual Snapshot Review and the Integrated Field Review (IFR). The annual snapshot gathers quantitative data along the five dimensions of school quality related to the EQS to examine a school's overall performance and assess the extent of parity for students among and between the districts/supervisory unions and schools. The IFR is intended to provide stakeholders with an in-depth review of school quality, illuminate potential reasons for what is revealed by the quantitative data, identify promising practices from the field, and provide support and interventions for school improvement.

Schools will participate in IFRs at least every three years. The teams conducting the reviews comprise a combination of students, educators, and Agency staff. The review team observes classrooms, reviews student work and school policies, and meets with parents, students, and staff to evaluate the school's performance along the same five dimensions of school quality guiding the Annual Snapshot Reviews. The on-site review concludes with the development of a written report that assesses performance, highlights a commendation and a focus area for each school quality dimension, and identifies resources to aid improvement.

As in **Kentucky** (see page 34), the results from the IFR will be integrated into the process for making accountability determinations for the supervisory unions/districts. The IFR process is being field tested during the 2015-16 school year and the EQS is expected to be implemented statewide the following year.

In **Massachusetts**, the Massachusetts Department of Elementary and Secondary Education (MDESE) employs an SQR process for reviewing charter schools for renewal. Each charter prepares a school accountability plan, which establishes its performance objectives and measures. The Department reviews the school's performance against the Charter School Performance Criteria[50]

based on this plan and an on-site visit, during which reviewers gather qualitative evidence of the school's performance on six criteria:

1. Mission and key design elements.

2. Access and equity.

3. Program delivery.

4. Culture and family engagement.

5. Capacity.

6. Governance.[51]

The frequency and length of the on-site visit is determined by the school's size, age, location, and status (whether it is on conditions or probation).[52] For example, charter schools in the first year of operation qualify for a year-one visit and are visited for one day only. Schools in year two or three of a charter term, or that are part of a network of charter schools, qualify for a full visit and are visited for one to three days. Schools on condition or probation receive a targeted visit, which lasts a half-day to a full day. Each charter school then receives a two-to-three day renewal inspection visit every five years.

Prior to the on-site visit, the school staff assembles key documents (e.g., school's master schedule, strategic plan, lesson plans, curricular units, summaries of key assessment data, descriptions of high quality instructional practices implemented in the school) and submits them to the Department. The review team consists of a Department staff member as the lead reviewer. Other Department staff and/or external volunteers may be engaged depending on the size, location, or special conditions of the school. On-site, the review team members conduct focus groups, interviews, observe classrooms,[53] conduct a document review to form judgments about the quality of teaching and learning in the school. Based on the evidence observed and collected, the review team creates initial findings that translate into ratings of the school's performance relative to the performance criteria. Criteria ratings range from "falls far below" to "partially meets" to "meets" to "exceeds."

At the end of the on-site visit, the review team shares its initial findings and ratings with the school staff. The team does not make suggestions or recommendations for improvement.[54] Instead, the team recounts the evidence-based or observed findings and ratings describing the school's performance in relation to the Charter School Performance Criteria. After the visit, the reviewers prepare a formal report that becomes a permanent part of the school's record. The site visit report is used to guide decision-making and ongoing improvement and is integral to the charter renewal process.

As part of the PACE system in **New Hampshire**, a district peer-review and auditing process has been incorporated into the system to ensure validity and reliability and to build local capacity. Peer review teams of external practitioners conduct school visits and examine the evidence submitted by the district. The peer review process supports continuous improvement as feedback based on common criteria is reported back to the districts. Eventually, the peer review audits will become a requirement of the approval process for districts seeking to implement the PACE system of accountability.

Diagnostic Review

- **Kentucky's diagnostic reviews of priority schools.** In 2014, the Kentucky Board of Education outlined intervention options for low-performing priority schools and districts.[55] A diagnostic review is scheduled within 90 days of a school being identified as a priority school, to evaluate the functioning of the school council and the leadership capacity of the principal. The diagnostic review may include the following activities:

  - Analysis of state and local education data.

  - Review of comprehensive school improvement plans and other planning documents.

  - Interviews with students, parents, all school council members, school and district personnel, and community members.

  - Direct observation.

  - Administration of school climate and culture surveys.

  - Review of school council minutes and agendas.

  - Review of family and community involvement strategies.

The Kentucky Department of Education contracts with AdvancEd, to lead the diagnostic reviews of its priority schools. The AdvancEd Standards for Quality Schools focus on five areas:

1. Purpose and direction.

2. Governance and leadership.

3. Teaching and assessing for learning.

4. Resources and support systems.

5. Using results for continuous improvement.[56]

The process includes an internal review conducted by school and district staff along with parents, students, and community leaders, and a school site visit, conducted by an AdvancEd team of external reviewers. The internal review results in the development of a school profile that portrays the school's current reality based on multiple sources of data, such as student assessment scores, student demographic data, programmatic data, and stakeholder perception/ survey data. The external review assesses the school's quality based on an on-site review of classroom observations and interviews with school staff, board members, district staff, parents, students, and community members. The external reviewers produce a written report that delineates the school's strengths as well as recommendations for improvement. The diagnostic report is the basis for the development of the priority school's continuous improvement plan.

After the diagnostic review, a written report is submitted to the Commissioner of Education with a recommendation as to whether the school council and principal have the capacity to lead the school to recovery or should be replaced. Following the initial diagnostic review, a follow-up review is to be conducted at least once every two years.

- **Massachusetts reviews of low-performing schools.** In similar fashion, the MDESE contracts with the American Institutes for Research (AIR) to conduct Monitoring Site Visits (MSV) at low-performing schools (Levels 4 and 5 in the state rating system) and School Redesign Grant recipient schools. The visits provide these schools with targeted information and feedback on the schools' progress toward meeting their improvement goals.[57] A school's progress is evaluated against four evidence-based turnaround practice areas:

  1. Leadership, shared responsibility, and professional collaboration.

  2. Intentional practices for improving instruction.

  3. Student-specific supports and instruction to all students.

  4. School culture and climate.[58]

Specific indicators have been identified for each turnaround practice area. For example, the leadership, shared responsibility, and professional collaboration area comprises seven indicators:

- Use of autonomy.

- High expectations.

- Vision/theory of action.

- Monitoring school progress.

- Instructional leadership and improvement.

- Time use.

- Communication with staff.

Each indicator is assessed along a four-point implementation continuum: (1) no evidence, (2) developing, (3) providing, and (4) sustaining. When a school is performing at the sustaining level across all of the indicators within a turnaround practice area, the school is viewed as demonstrating a high level of coherence that can lead to improvements in student outcomes.[59]

The site visit typically takes two days to complete by AIR staff. Prior to the visit, the school staff collects and sends to the AIR staff preparatory materials, such as the master schedule, a staff list, the daily schedule, and a map of the school. The first day of the on-site visit generally consists of interviews and focus groups with school leaders, staff, students (only in middle and high schools), district staff, and identified external partners. The second day visit generally consists of classroom observations. The AIR staff employs the Classroom Assessment Scoring System (CLASS) observation tool, with each classroom observed for up to 20 minutes. Within four business days of the visit, the findings from the observations are shared in the Schoolwide Instructional Observation report with school and district leaders as well as with the MDESE. The overall MSV report, delivered within four weeks, incorporates data from interviews and focus groups and the classroom observations. The report provides an overview, specific information, and evidence related to turnaround practices and the school's progress on the indicators.

- **Connecticut's instructional audits for schools and districts in need of improvement.**
  Connecticut contracts with Cambridge Education, an organization that provides consultancy and training services in education, to conduct instructional and fiscal audits of schools and districts identified as in need of improvement in Year 3 or greater under the NCLB legislation.[60] The instructional audits are a central component in the state's strategy for promoting continuous school improvement. They evaluate a school's progress against five quality criteria:[61]

  1. Student achievement in the core subject areas.

  2. Quality of instruction provided through teaching, learning, and curriculum.

  3. Students' personal character development.

  4. Effective leadership and management.

  5. Partnerships with parents and community.

  To conduct the audits, Cambridge Education works collaboratively with staff from the Connecticut State Department of Education (CSDE) to review school and district self-evaluations and conduct an on-site visit. The on-site visit includes observations of classroom teaching and interviews with school and district staff, students, parents, and community members. At the conclusion of the visit, preliminary verbal findings about the school's or district's strengths and areas of development are shared with the leadership team. After the site visit, a written report is drafted by Cambridge Education staff and reviewed for accuracy by school and district staff and then by the CSDE staff. A final report is submitted to the CSDE, which then releases it to the school or district. The report becomes the grist with which school and district staff, in collaboration with local stakeholders, revise the district improvement plans.

How Might Evidence-based Interventions Be Evaluated?

Following the review of information and a diagnostic process, states have an obligation to help schools improve. ESSA requires schools and districts to develop and implement evidence-based interventions to improve student outcomes. As explained earlier, an educational intervention is evidence based when at least one high-quality study demonstrates that the intervention is statistically significantly likely to improve student achievement.

A large body of educational research has explored practices that are effective (and ineffective) for improving student outcomes. A key issue is that strategies for improving performance have generally proved successful only when implemented effectively and in contexts that can benefit from what they have to offer. Understanding how to evaluate this research can empower state and local policymakers to adopt educational interventions that best address the unique context of their local education system.

Below, we discuss four commonly used interventions that research has found to improve student outcomes under certain circumstances, seeking to illuminate the conditions under which they have been found to be successful. In a future publication, we will review a broader range of potential interventions more deeply.

High-quality Professional Development

Clearly, changing curriculum and teaching practices require investments in teachers' professional learning, and some schools have shown significant achievement gains by making such investments strategically.[62] However, not all professional development (PD) is designed in ways that produce these effects.

A key feature of effective PD is that teachers work together on a particular set of practices over a sustained period of time. Of nine well-designed experimental or quasi-experimental investigations, Yoon and colleagues found that 14 or fewer hours of professional development on a given topic showed no significant effects on student learning.[63] The professional development that showed positive and significant effects on student achievement averaged 49 hours.[64]

The greatest improvements in student achievement have been found to be associated with PD approaches that:

- focus on deepening teachers' content knowledge and instructional practices;[65]

- function as a coherent part of a school's improvement efforts, aligned with curriculum, assessments, and standards, so that teachers can implement the knowledge and practices they learn in their classrooms;[66]

- occur in collaborative and collegial learning environments in which teachers participate in professional learning and together grapple with issues related to new content and instructional practices;[67]

- provide authentic activities rooted in teachers' inquiry and reflection about practice within the context of the curriculum and students they teach;[68]

- link to analysis of teaching and student learning, including the formative use of assessment data;[69] and

- are supported by coaching, modeling, observations, and feedback.[70]

- States and districts will want to be informed by research that highlights the critical components of PD most likely to markedly improve teachers' skills and students' outcomes.

Class Size Reduction

Reducing class size can help improve student outcomes under some circumstances. However, the effects appear to vary depending on the age and character of the students and the extent of class-size reduction pursued. And they assume that other variables, such as the quality of teachers and curriculum, remain constant.

For example, a meta-analysis of 77 studies exploring the effects of class size found that smaller class sizes were associated with improved student achievement, with the greatest effects when certain smaller class thresholds were reached. For example, reducing a class size of 40 students to a class of no more than 20 students, or a class of 25 students to a class of 10–15 students produced the greatest gains in student achievement.[71]

Similarly, the well-known experimental study of Tennessee's Project Student Teacher Achievement Ratio (STAR) found that reducing class sizes below certain threshold levels in kindergarten through third grade improved student achievement, with benefits persisting for at least five years after the student was assigned to a smaller class.[72] In particular, small classes of fewer than 18 students made greater gains in their achievement on standardized tests than students in regular-sized classes (22–25 students). Importantly, the effects of being in a small class were nearly twice as large for students of color as for their white peers. Test score gains were greatest for children in kindergarten and 1st grade with persistent long-term effects on a variety of academic outcomes in middle and high school.[73]

Studies of Wisconsin's statewide class size reduction experiment found that reducing student-teacher ratios in kindergarten through 3rd grade to fewer than 15 students per teacher (as compared to ratios of 21:1 and 25:1) was associated with improved student achievement. The largest benefits from smaller class sizes were experienced by African American students and students in urban districts with large proportions of low-income students.[74]

In sum, positive results, especially for low-income students and students of color, have been found in the literature when class-size reduction programs are well-designed, meet a relatively low threshold of class size (in the vicinity of 15 to 18 students), and are implemented in the early grades.

Community Schools and Wraparound Services

A community school is both a physical place and a set of partnerships between the school and other community resources.[75] Community schools take on a results-focused integrated approach that links high-quality academics with health and social services, youth and community development, and community engagement. Particularly in schools serving low-income students, community school models and wraparound services have been found to improve student outcomes.[76]

The rationale for a community school is that students need more than just high-quality instruction to achieve academic and personal success. Children need access to housing, food, and health care, as well as social and learning supports. In addition, parents and the broader community need coordinated, one-stop services so that they can easily receive assistance that ultimately supports children's development. Consequently, many community schools offer onsite clinics that provide physical and mental health care, social welfare services, before- and after-school care, tutoring and mentoring, preschool, a focus on social-emotional learning and positive discipline approaches, and parent and community engagement.[77] In addition, many of these services are open to all community members during the day, evening, and weekend.[78]

Research about the effectiveness of community schools and other approaches that ensure a wide range of services for children often finds that such schools are associated with improved student outcomes, especially for the most socioeconomically disadvantaged students.[79] Of course, it is not the "community schools" label that makes a difference: A critical mass of key features must be in place and well implemented to derive these effects.

In addition to overarching studies, certain features of such schools have been studied individually, showing positive influences on outcomes. For example, research has found significant student learning gains as a result of expanded learning time, including an extended school day, time for tutoring and homework help,[80] and summer learning time.[81]

The frequency of parent, family, and community engagement is positively associated with improved student academic achievement, lower rates of grade retention, fewer years that students spend in special education, and gains in English language development exam scores for English language learners.[82] Wraparound academic, health, and social services are associated with improved academic outcomes, especially for the most vulnerable students.[83] Social-emotional learning supports and positive behavioral interventions, such as restorative justice practices, are associated with increased student academic success and lower rates of suspensions, expulsions, and dropouts.[84]

> The frequency of parent, family, and community engagement is positively associated with improved student academic achievement, lower rates of grade retention, fewer years that students spend in special education, and gains ... for English language learners.

One approach, the Turnaround for Children (TFC) model, addresses these goals in underperforming schools through capacity building and culture change. TFC's model is an integrated set of practices and supports that address students' social, emotional, and physical well-being while also transforming school culture. TFC sets up a partnership between the school and a community-based mental health provider to create a focus on child development aimed at reversing poverty-induced traumas that impede learning. It then deploys a team consisting of a social work consultant (veteran clinical social worker), instructional coach (master teacher), and program director (experienced school administrator) to:

- build a high-capacity student support system that provides help to all children, including those with intense needs, either in their schools or in partnership with community-based mental health providers;

- train all teachers in proven classroom strategies that foster a safe, engaging learning environment and strong student-teacher relationships; and

- work with school leaders to drive school-wide improvement aligned to Common Core State Standards and district guidelines with the intention of creating a high-performing culture that involves the entire school community.

Data from TFC schools in New York City shows dramatic gains in math and English language arts scores, as well as greater safety, a decrease in suspensions, and a decrease in teacher absences and turnover.[85]

High School Redesign

The effective redesign of secondary schools is another intervention strategy supported by research. Although school size and structural features are potential tools to help schools support student attachment and learning, the results they produce depend in substantial part on how these elements are implemented. Effective redesigned schools share a number of features that influence student achievement, including personalization, a shared school mission focused on high-quality student learning, a strong core curriculum for all students, high-quality "authentic" instruction,

and a professional community.[86] The redesign strategies discussed below hold promise for helping schools achieve improved outcomes, but success ultimately depends on how each element is implemented.

- **Small size and personalization.** A number of studies have found that, all else being equal, schools have higher levels of achievement when they create smaller, more personalized communities of teachers and students in which teachers work together and students see a smaller number of teachers over a longer period of time. This allows teachers and students to come to know one another well.[87] For example, a study of 820 high schools in the National Education Longitudinal Study, which controlled for student characteristics, found that those schools that restructured to personalize education and develop collaborative learning structures produced significantly higher achievement gains that were also distributed more equitably across more and less advantaged students.[88] Other studies have found improved student and teacher relationships and increased student engagement, as well as improved student achievement, as a result of these strategies.[89]

  Personalization can be accomplished not only through smaller classes but also through longer stretches of time spent by adults with the same groups of students—through longer class periods associated with smaller pupil loads for teachers, fewer courses per term, advisories (classes in which teachers meet regularly with students to advise and support students with their work, sometimes staying together over multiple years), and looping (teachers stay with the same group of students for more than a year).[90] Another strategy for enhancing personalization is teaming, in which a few teachers share the same group of students and regularly discuss students' progress in addition to connecting curriculum across content areas.

- **Shared school mission.** A problem commonly reported among less successful schools is goal diffusion, as fragmentation and managerial distractions cause schools to lose focus on teaching and learning. A common theme running through the research on successful schools is having a clear, shared focus on student learning with common norms and practices across classrooms.[91] Faculty communication, community ownership, and a common purpose and curricular focus—all associated with developing a shared school mission—facilitate greater participation by marginal students.[92] Other research confirms that developing common goals, norms, and practices with a strong focus on teaching and learning leads to greater student engagement and student outcomes, especially for underserved students.[93] This signals the importance of skilled school leaders who can help staff, parents, and students develop a shared sense of mission that translates into common norms, beliefs, and practices.

- **Strong, common academic curriculum.** Students attending schools that emphasize academic rigor and provide a more common curriculum with less tracking are more likely to make greater gains in their academic achievement.[94] Moreover, students attending such schools have lower rates of absenteeism and stronger graduation rates.[95] In fact, when students of similar backgrounds and initial achievement levels are exposed to more and less challenging curricula, those given the richer curriculum opportunities outperform those placed in less challenging classes.[96] As a corollary, students achieve at lower levels and exhibit more behavioral problems when they are tracked into classes that are academically unchallenging.[97] Importantly, schools that have successfully created a common curriculum for students of varying levels of initial achievement have offered other supports and interventions alongside the curriculum, such as during- and after-school help with homework and tutoring.[98]

- **Authentic instruction.** A number of studies have found positive influences on student achievement from what researchers call authentic instruction—that is, meaningful instruction, curriculum, and assessment that requires students to construct and organize knowledge, consider alternatives, apply disciplinary processes to content central to the discipline, and communicate effectively to audiences beyond the classroom and school.[99] This is the kind of instruction called for in the new Common Core State Standards. For example, a study of more than 2,100 students in 23 restructured schools found significantly higher achievement on intellectually challenging performance tasks for students who experienced authentic instruction.[100]

  Authentic instruction generally occurs through performance-based activities, including experiments, mathematical modeling, social science inquiry, and other projects requiring in-depth study, extensive writing, or public presentations. These activities can create high expectations throughout a school and encourage mutual teacher and student accountability for meeting expectations.[101]

- **Professional community.** Many researchers have identified the collaboration associated with a professional community of teachers as a key element of successful schools.[102] Bryk, Camburn, and Louis define a professional community as teachers' focus on student learning, collective responsibility for school improvement, de-privatized practice, reflective dialogue, and staff collegiality and collaboration.[103] A professional community encourages teachers to take responsibility for student learning and provides them with tools to do so, through collaboration around learning problems and effective teaching practices. In their study of 24 restructured schools, Newmann and colleagues found that having a strong professional community of practice is one of three commonalities among schools achieving high levels of student learning.[104] Other research suggests that a collegial professional environment for teachers produces stronger achievement and generates greater collective responsibility for school improvement and student learning.[105]

Conclusion

The Every Student Succeeds Act provides an important opportunity to create new accountability strategies that seek to view students and schools more holistically. Taking advantage of this opportunity will require clarity about what the act permits and requires, as well as creativity in developing new measures, processes for school diagnosis and improvement, and evidence-based interventions that support deeper learning in contexts that further equity goals.

Endnotes

1. See, e.g., U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), Averaged freshmen graduation rates, http://nces.ed.gov/ccd/data_tables.asp and Adjusted Cohort Graduation Rates, and http://nces.ed.gov/ccd/data_tables.asp and http://nces.ed.gov/ccd/tables/ACGR_2010-11_to_2012-13.asp.

2. McMurrer, J. (2007). *Choices, changes, and challenges: Curriculum and instruction in the NCLB era.* Washington, DC: Center for Education Policy.

3. 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.

4. 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.

5. Darling-Hammond et al. (2014).

6. See CQ House Action Report No. 114-3/Dec. 1, 2015: Schools under such plans would be eligible for "comprehensive support and improvement" funding to take appropriate actions under the approved plans for up to four years as determined by the state. If a school has not addressed all its issues within the state-determined time period, a state could take more rigorous action. ESSA also permits local school districts to allow students in these schools to transfer to other public schools, starting with the lowest-achieving children from low-income families, unless prohibited by state law.

7. The strength of the evidence required is greater for interventions used in the lowest performing schools.

8. Darling-Hammond, L., & Falk, B. (2013). *Teacher learning through assessment: How student performance assessments can support teacher learning.* Washington, DC: Center for American Progress.

9. Conley, D. T., & Darling-Hammond, L. (2013). *Creating systems of assessment for deeper learning.* Stanford, CA: Stanford Center for Opportunity Policy in Education.

10. For a review, see Darling-Hammond, L., & Adamson, F. (2014). *Beyond the bubble test: How performance assessments support 21st century learning.* San Francisco: Jossey-Bass.

11. Measured Progress. (2009). *Commonwealth Accountability and Testing System: 2007-08 Technical Report.* Version 1.2. Retrieved February 20, 2010, from http://www.education.ky.gov/KDE/Administrative+Resources/Testing+and+Reporting+/Kentucky+School+Testing+System/Accountability+System/Technical+Manual+2008.htm.

12. For a review, see Darling-Hammond, L., & Rustique-Forrester, E. (2005). The consequences of student testing for teaching and teacher quality. In J. Herman & E. Haertel (Eds.), *The uses and misuses of data in accountability testing* (pp. 289–319). Malden, MA: Blackwell.

13. Newmann, F. M., Marks, H. M., & Gamoran, A. (1995). Authentic pedagogy: Standards that boost performance. *American Journal of Education,* 104 (4), 280-312; Lee, V. E., Smith, J. B., & Croninger, R. G. (1995, Fall). Another look at high school restructuring: More evidence that it improves student achievement and more insight into why. *Issues in Restructuring Schools.* Issue report no. 9, pp. 1-9. Madison, WI: Center on the Organization and Restructuring of Schools, University of Wisconsin.

14. The bank can be accessed at http://performanceassessmentresourcebank.org/.

15. New Hampshire Department of Education. (2014, November 21). *New Hampshire Performance Assessment of Competency Education: An accountability pilot proposal to the United States Department of Education.* Concord, NH: Author. Retrieved on March 13, 2015, from http://www.education.nh.gov/assessment-systems/documents/pilot-proposal.pdf.

16. California Office to Reform Education. (2015, February). *School Quality Improvement Index – Short Metric Descriptions.* Retrieved January 8, 2015, from http://coredistricts.org/wp-content/uploads/2015/02/School-Quality-Improvement-Index- Short-Metric-Descriptions-updated-2.18.15.pdf.

17. Bae, S., & Darling-Hammond, L. (2014). *Recognizing college and career readiness in the California school accountability system*. Stanford, CA: Stanford Center for Opportunity Policy in Education.

18. South Carolina Education Oversight Committee. (2015). *2014-2015 Accountability manual: The annual school and district report card system for South Carolina public schools and school districts*. Retrieved on September 29, 2015, from http://www.eoc.sc.gov/Reports%20%20Publications/Acct%20Manual%202014-15/FINAL%202014-15%20Manual.pdf.

19. Booher-Jennings, J. (2005). Below the bubble: "Educational triage" and the Texas accountability system. *American Educational Research Journal*, 42(2), 231-268.

20. Rothman, R. (2015). *Data dashboards: Accounting for what matters*. Washington, DC: Alliance for Excellent Education.

21. Bryk, A. S., Sebring, P. B., Allensworth, E., Easton, J. Q., & Luppescu, S. (2010). *Organizing schools for improvement: Lessons from Chicago*. University of Chicago Press.

22. Hirsch, E. D. (1987). *Cultural literacy: What every American needs to know*. New York, NY: Houghton Mifflin Company.

23. Goldhaber, D., Lavery, L., & Theobald, R. (2015). Uneven playing field? Assessing the teacher quality gap between advantaged and disadvantaged students. *Educational Researcher*, 44(5), 293-307; Lankford, H., Loeb, S., & Wyckoff, J. (2002). Teacher sorting and the plight of urban schools: A descriptive analysis. *Educational Evaluation and Policy Analysis*, 24(1), 37-62.

24. Darling-Hammond, L. (2014). *The flat world and education: How America's commitment to equity will determine our future*. NY: Teachers College Press.

25. Adamson, F., & Darling-Hammond, L. (2010). Inequitable Distribution of Teachers. *Education Policy Analysis Archives*.

26. Gottfried, M. A. (2010). Evaluating the relationship between student attendance and achievement in urban elementary and middle schools an instrumental variables approach. *American Educational Research Journal*, 47(2), 434-465.

27. Plank, S., Farley-Ripple, E., Durham, R., & Norman, O. (2009). *First grade and forward: A seven-year examination within the Baltimore City Public School System*. Baltimore, MD: Baltimore Education Research Consortium.

28. Rothman (2015).

29. Balfanz, R., & Byrnes, V. (2012). *Chronic absenteeism: Summarizing what we know from nationally available data*. Baltimore, MD: Johns Hopkins University Center for Social Organization of Schools.

30. Watanabe, T. (2015, January). Most major California school districts pledge to reduce suspensions. *Los Angeles Times*. Retrieved from http://www.latimes.com/local/lanow/lame- ln-school-suspensions-20150128-story.html.

31. Klein, A. (2015, June). Ed. Dept. Announces New Grant for Social and Emotional Learning. *Education Week*. Retrieved from http://blogs.edweek.org/edweek/campaign-k-12/2015/06/ed_dept_announces_new_grant_fo.html.

32. California Office to Reform Education Social-Emotional Skills. (2014, September 24). In CORE Districts School Quality Improvement System (Waiver). Retrieved on January 26, 2015, from http://coredistricts.org/wp-content/uploads/2014/10/SE-CC-Domain-Social-Emotional-Skills-updated-10.3.14-1.pdf.

33. Bryk et al. (2010).

34. Mellor, M., & Griffith, D. (2015). *Multimetric accountability systems: A next-generation vision of student success.* Alexandria, VA: ASCD.

35. New York City Department of Education. (2015, October 30). *Final changes to the School Quality Reports for 2014-15.* Retrieved March 15, 2016, from http://schools.nyc.gov/NR/rdonlyres/4F6B6F94-9DE0-4AD6-8F34-7D900E13E5C0/0/201415FinalChanges10302015.pdf.

36. New York City Department of Education. (2015).

37. Bitter, C., & O'Day, J. (2006). California's accountability system. In *Crucial issues in California education 2006: Rekindling Reform* (pp. 51-74). Berkeley, CA: Policy Analysis for California Education.

38. Local Control Funding Formula Overview. (2015, December 11). In California Department of Education. Retrieved on January 13, 2016, from http://www.cde.ca.gov/fg/aa/lc/lcffoverview.asp.

39. Program Reviews. (2015, March 20). In Kentucky Department of Education. Retrieved on March 24, 2015, from http://education.ky.gov/curriculum/pgmrev/Pages/default.aspx.

40. Local Control Funding Formula Overview. (2015).

41. Kentucky Department of Education. (2012). *Unbridled learning accountability model (with focus on the next-generation learners component).* Retrieved on March 5, 2015, from http://education.ky.gov/comm/UL/Documents/WHITE%20PAPER%20062612%20final.pdf.

42. Program Reviews, para. 2 (2015).

43. New York City Department of Education (n.d.1). *Reviewer handbook for the quality review 2013-2014.* Retrieved on August 12, 2014, from http://schools.nyc.gov/NR/rdonlyres/5CC3B3DF-49BF-4BA6-AFD0-0FE6BFBDC3E1/0/ReviewerHandbook201314CFarina.pdf.

44. New York City Department of Education (n.d.2). *Quality review 2013-14: Big ideas by indicator and sub-indicator.* Retrieved on August 12, 2014, from http://schools.nyc.gov/NR/rdonlyres/B94E5CCB-1FEB-4CED-992B-7239A7C70577/0/QualityReviewBigIdeas201314.pdf.

45. Reviews of larger schools of 1,500 students or more and District 75 schools may engage an additional reviewer.

46. New York City Department of Education *2015-2016 Quality review rubric.* Retrieved from http://schools.nyc.gov/NR/rdonlyres/8C11A001-7E78-469D-996F-B0C3703CEA81/0/201314QualityReviewRubric.pdf.

47. See the NYCDOE *Reviewer handbook for the quality review* for an example of the Classroom Visitation Tool. Retrieved from http://schools.nyc.gov/NR/rdonlyres/5CC3B3DF-49BF-4BA6-AFD0-0FE6BFBDC3E1/0/ReviewerHandbook201314CFarina.pdf.

48. Vermont State Board of Education. (2014). *Education quality standards: State board rule 2000.* Retrieved on October 2, 2015, from http://education.vermont.gov/documents/EDUFinalEQS_AsAdopted.pdf.

49. Fowler, A. (2015). *Education quality reviews: the whole picture.* Barre, VT: Vermont Agency of Education. Retrieved on October 2, 2015, from http://education.vermont.gov/documents/edu-oped-education-quality-reviews.pdf.

50. See http://www.doe.mass.edu/charter/guides/PerformanceCriteria.pdf for detailed descriptions of the performance criteria.

51. Massachusetts Department of Elementary and Secondary Education. (2014a). *Charter school performance criteria.* Malden, MA: Author. Retrieved on April 27, 2015, from http://www.doe.mass.edu/charter/guides/PerformanceCriteria.pdf.

52. Massachusetts Department of Elementary and Secondary Education. (2014b). *Charter school site visit protocol*. Malden, MA: Author. Retrieved on April 27, 2015, from http://www.doe.mass.edu/charter/guides/SiteVisit.pdf.

53. See http://www.doe.mass.edu/charter/guides/ObservationForm.pdf for the observation form used by the MDESE to conduct classroom observations.

54. Massachusetts Department of Elementary and Secondary Education. (2014b).

55. See http://www.lrc.ky.gov/kar/703/005/260.htm for more detailed descriptions of 703 KAR 5:260 legislation.

56. See http://palmbeachschools.org/schoolimprovement/documents/StandardsforQualitySchools.pdf for a more detailed explanation of AdvancEd's Standards for Quality Schools.

57. American Institutes for Research & Massachusetts Department of Elementary and Secondary Education. (2015). *Massachusetts Level 4 and Level 5 School Monitoring Site Visit protocol*. Retrieved on April 7, 2015, from http://www.doe.mass.edu/apa/sss/turnaround/MSV-Protocol.pdf.

58. American Institutes for Research. (2015). *Massachusetts Monitoring Site Visits: Turnaround practices and indicators*. Washington, DC: Author. Retrieved on April 7, 2015, from http://www.doe.mass.edu/apa/sss/turnaround/MSV-TurnaroundIndicators.pdf. See also http://www.doe.mass.edu/apa/sss/turnaround/MSV-TurnaroundIndicators.pdf for a more detailed explanation of the specific indicators for each turnaround practice.

59. American Institutes for Research. (2015).

60. School and District Assessments. (2009, November 4). In Connecticut State Department of Education. Retrieved on January 11, 2016, from http://www.sde.ct.gov/sde/cwp/view.asp?a=2700&q=322194.

61. Connecticut State Department of Education. (n.d.). *Connecticut quality review: Notes of guidance 2007-2008*. Hartford, CT: Author. Retrieved on April 27, 2015, from http://www.sde.ct.gov/sde/lib/sde/pdf/cali/assess_cambridgeframeworkdocument.pdf.

62. Cohen, D. K., & Hill, H. C. (2000). Instructional Policy and Classroom Performance: The Mathematics Reform in California. *Teachers College Record*, 102(2), 294-343. See also Corcoran, T. B., Shields, P. M., & Zucker, A. A. (1998). *The SSIs and professional development for teachers*. Menlo Park, CA: SRI International; Elmore, R. F. (1997*). Investing in teacher learning: Staff development and instructional improvement in Community School District #2, New York City*. New York, NY: National Commission on Teaching & America's Future.

63. Yoon, K. S., Duncan, T., Lee, S. W. Y., Scarloss, B., & Shapley, K. L. (2007). Reviewing the evidence on how teacher professional development affects student achievement. *Issues and Answers Report*, REL 2007-No. 033. Washington, D.C.: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. http://ies.ed.gov/ncee/edlabs.

64. Yoon et al. (2007). See also Cohen, D. K., & Hill, H. C. (2000); Dunst, C.J., Bruder, M., & Hamby, D.W. (2015). Metasynthesis of in-service professional development research: Features associated with positive educator and student outcomes, *Educational Research and Reviews, 10(12), 1731-1744*; Lumpe, A., Czerniak, C., Haney, J., & Beltyuova, S. (2012). Beliefs about teaching science: The relationship between elementary teachers' participation in professional development and student achievement, *International Journal of Science Education 34(2), 153-166;* Wei, R.C., Darling-Hammond, L., & Adamson, F. (2010). *Professional development in the United States: Trends and challenges*. Dallas, TX. National Staff Development Council; Weiss, I. R., & Pasley, J. D. (2006). Scaling Up Instructional Improvement through teacher professional development: Insights from the local systemic change initiative. *CPRE Policy Briefs*. RB-44. Consortium for Policy Research in Education; Desimone, L. M., Porter, A. C., Garet, M. S., Yoon, K., & Birman, B. F. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study, *Educational Evaluation and Policy Analysis* 24(2), 81-112.

65. Carpenter, T. P., Fennema, E., Peterson, P. L., Chiang, C. P., & Loef, M. (1989). Using knowledge of children's mathematics thinking in classroom teaching: An experimental study. *American Educational Research Journal*, 26(4), 499-531. See also Saxe, G. B., Gearhart, M., & Nasir, N. S. (2001). Enhancing students' understanding of mathematics: A study of three contrasting approaches to professional support. *Journal of Mathematics Teacher Education*, 4(1), 55-79.

66. Wei et al. (2010); Penual, W. R., Fishman, B. J., Yamaguchi, R., & Gallagher, L. P. (2007). What makes professional development effective? Strategies that foster curriculum implementation. *American Educational Research Journal,* 44(4): 921-958; Desimone et al. (2001); Supovitz, J. A., Mayer, D. P., & Kahle, J. B. (2000). Promoting inquiry-based instructional practice: The longitudinal impact of professional development in the context of systemic reform. *Educational Policy*, 14(3), 331-356.

67. Darling-Hammond, L., & McLaughlin, M. W. (1995). Policies that support professional development in an era of reform. *Phi Delta Kappan*. 76(8), 597-604. See also Darling-Hammond, L., Wei, R. C., Andree, A., Richardson, N., & Orphanos, R. (2009). *Professional learning in the learning profession*. Washington, DC. National Staff Development Council; Knapp, M.S. (2003). Professional development as policy pathway. *Review of Research in Education*, 27(1), 109-147.

68. Dunst et al. (2015);Wei et al. (2010).

69. Dunst et al. (2015); Wei et al. (2010).

70. Dunst et al. (2015); Cavanaugh, B. (2013). Performance feedback and teachers' use of praise and opportunities to respond: A review of the literature. *Education and Treatment of Children* 36(1), 111-136; Wei et al. (2010).

71. Glass, G. V., & Smith, M. (1979). Meta-Analysis of Class Size and Achievement. *Educational Evaluation and Policy Analysis*, 1(1), 2-16.

72. Mosteller, F. (1995). The Tennessee study of class size in the early school grades. *The future of children*, 113-127; Nye, B., Hedges, L. V., and Konstantopoulos, S. (1999). The long-term effects of small classes: A five-year follow-up of the Tennessee class-size experiment. *Evaluation and Policy Analysis*, 21(2), 127-142.

73. Kim, J. (2006/2007). The relative influence of research on class-size policy. *Brookings Papers on Education Policy*, 273-295.

74. Kim, J. (2006/2007).

75. Institute for Educational Leadership. *What is a community school?* http://www.communityschools.org/aboutschools/what_is_a_community_school.aspx.

76. Blank, M., Jacobson, R., & Pearson, S. (2009). Well-conducted partnerships meet students' academic, health, and social service needs. *American Educator*, 33, 30-36.

77. Harkavy, I., & Blank, M. (2002). Community school: A vision of learning that goes beyond testing. *Education Week*, 52. Available at http://www.edweek.org. See also Daniel, J. (2015). Community schools as an effective reform strategy. Congressional forum: Closed for learning: The impact of school closures on students and communities. [PowerPoint slides]. Obtained from author.

78. Harkavy & Blank. (2002).

79. Adams, C. M. (2010). *The community school effect: evidence from an evaluation of the Tulsa Area Community School Initiative.* The Oklahoma Center for Educational Policy. Retrieved January 5, 2016, from http://www.csctulsa.org/files/file/Achievement%20Evidence%20from%20an%20Evaluation%20of%20TACSI.pdf. See also Castrechini, S., & London, R. A. (2012). *Positive student outcomes in community schools*. Washington, DC: Center for American Progress; Dobbie, W., & Fryer, Jr., R. G. (2011). Are high-quality schools enough to increase achievement among the poor? Evidence from the Harlem Children's Zone. *American Economic Journal: Applied Economics*, 158-187.

80. Farbman, D. (2015). *The case for improving and expanding time in school: a review of key research and practice.* Retrieved January 4, 2016, from http://www.timeandlearning.org/sites/default/files/resources/caseformorelearningtime.pdf. See also Dobbie et al. (2011); Hoxby, C. M., Muraka, S., & Kang, J. (2009). How New York City's charter schools affect achievement. Cambridge, MA; *New York City Charter Schools Evaluation Project,* 1-85.

81. McCombs, J., Augustine, C. H., & Schwartz, H. L. (2011). *Making summer count: How summer programs can boost children's learning.* Santa Monica, CA: RAND Corporation.

82. Castrechini & London. (2012); Miedel, W.T. & Reynolds, A.J. (2000). Parent involvement in early intervention for disadvantaged children: Does it matter? *Journal of School Psychology*, 37(4), 379-402.

83. Anderson, K., & Emig, C. (2014). *Integrated student supports: a summary of the evidence base for policymakers*. Retrieved on January 4, 2016, from http://www.childtrends.org/wp-content/uploads/2014/02/2014-05ISSWhitePaper1.pdf. See also Basch, C.E. (2011). Healthier students are better learners: A missing link in school reforms to close the achievement gap. *Journal of School Health*, 81(10), 593-598; Vinciullo, F. M., & Bradley, B. J. (2009). A correlational study of the relationship between a coordinated school health program and school achievement: A case for school health. *The Journal of School Nursing, 25(6), 453-65.*

84. Drysfoos, J. G. (2000). *Evaluation of community schools: Findings to date*. Coalition for Community Schools. Retrieved January 4, 2016, from http://www.communityschools.org/assets/1/AssetManager/Evaluation%20of%Communit y%20Schools_joy_dryfoos.pdf. See also Bradshaw, C. P., Mitchell, M. M., & Leaf, P. J. (2010). Examining the effects of schoolwide positive behavioral interventions and supports on student outcomes results from a randomized controlled effectiveness trial in elementary schools. *Journal of Positive Behavior Interventions,* 12(3), 133-148; Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, 82(1): 405–432; Horner, R. H., Sugai, G., Smolkowski, K., Eber, L., Nakasato, J., Todd, A. W., & Esperanza, J. (2009). A randomized, wait-list controlled effectiveness trial assessing school-wide positive behavior support in elementary schools. *Journal of Positive Behavior Interventions*, 11(3), 133-144.

85. Cantor, P. A., Smolover, D. S., & Stamler, J. K. (2010). Innovation designs for persistently low-performing schools: Transforming failing schools by addressing poverty-related barriers to teaching and learning. In *Transforming America's Education Through Innovation and Technology*, 25(4). Retrieved from http://www.aspeninstitute.org/sites/default/files/content/upload/2010_Education_CR- Whistler.pdf.

86. Darling-Hammond, L., Ross, P., & Milliken, M. (2006/2007). High school size, organization, and content: What matters for student success? *Brookings Papers on Education Policy*, 9, 163-203.

87. Darling-Hammond et al. (2006/07).

88. Lee, V. E., & Smith, J. B. (1995). Effects of high school restructuring and size on early gains in achievement and engagement. *Sociology of Education* 68(4): 241-70.

89. Friedlaender, D., Burns, D., Lewis-Charp, H., Cook-Harvey, C. M., & Darling-Hammond, L. (2014). *Student-centered schools: Closing the opportunity gap*. Stanford Center for Opportunity Policy in Education (2014). See also Shear, L., Soung, A., House, B., Martinez, B., & Smerdon, B. (2005). *Creating cultures for learning: Supportive relationships in new and redesigned high schools*. Washington, DC; American Institutes for Research; Wasley, P. A., Fine, M., Gladden, M., Holland, N. E., King S.P., Mosak, E., & Powell, L. C. (2000). *Small schools: Great strides. A study of new small schools in Chicago*. New York, NY: Bank Street College of Education.

90. Lee & Smith. (1995). See also Gambone, M. A., Klem, A. M., Moore, W. P., & Summers, J. A. (2002). *First things first: Creating the conditions and capacity for community-wide reform in an urban school district*. New York, NY: Manpower Demonstration Research Corporation.

91. Darling-Hammond et al. (2006/07).

92. Howley, C. B., & Harmon, H. L., eds. (2000). *Small high schools that flourish: Rural context, case studies, and resources*. Charleston, West Virginia: Appalachia Educational Laboratory.

93. Darling-Hammond et al. (2006/07). See also Wasley et al. (2000); Bryk, A. S., Lee, V. E., & Holland, P. B. (1993). *Catholic Schools and the Common Good*. Harvard University Press.

94. Darling-Hammond et al (2006/07). See also Lee, V. E., Croninger, R. G., & Smith, J. B. (1997) Course-taking, equity, and mathematics learning: Testing the constrained curriculum hypothesis in U.S. secondary schools. *Educational Evaluation and Policy Analysis*, 19(2), 99-121; Bryk et al (1993).

95. Darling-Hammond et al. (2006/07).

96. Oakes, J. (2005). *Keeping track: How schools structure inequality* (2nd ed.) Yale University Press.

97. Oakes. (2005).

98. Darling-Hammond et al. (2006/07). See also Friedlaender et al. (2014); Mitchell, K., Shkolnik, J., Song, M., Uekawa, K., Murphy, R., Garet, M., & Means, B. (2005). *Rigor, relevance, results: The quality of teacher assignments and student work in new and conventional high schools*. Evaluation of the Bill & Melinda Gates Foundation's high school grants. Washington, DC: American Institutes of Research.

99. Bransford, J. D., Brown, A. L., & Cocking, R. R. (1999). *How People Learn: Brain, Mind, Experience, and School*. National Academy Press.

100. Newmann et al. (1995).

101. Friedlaender et al. (2014). See also Newmann et al. (1995).

102. Darling-Hammond et al. (2006/2007).

103. Bryk, A., Camburn, E., & Louis, K. (1999). Professional community in Chicago elementary schools: Facilitating factors and organizational consequences. *Educational Administration Quarterly* 35:751-81.

104. Newmann et al. (1996).

105. Kraft, M. A. & Papay, J. P. (2014). Do supportive professional environments promote teacher development? Explaining heterogeneity in returns to teaching experience. *Educational Evaluation and Policy Analysis*, 36(4): 476-500. See also Bryk et al. (1999).

About the Authors

Linda Darling-Hammond is Charles E. Ducommun Professor of Education Emeritus at Stanford University and President of the Learning Policy Institute. She has conducted extensive research on issues of accountability, assessment, and school improvement and has worked with many federal, state, and local agencies, as well as schools, on these issues. Among her recent books on these topics are *The Flat World and Education: How America's Commitment to Equity will Determine our Future* and *Beyond the Bubble Test: How Performance Assessments Support 21st Century Learning*.

Soung Bae is a Senior Research and Policy Analyst at the Stanford Center for Opportunity Policy in Education. Her research interests focus on teacher learning and development, organizational learning, and the role of educational policy on teachers' work. Most recently, her research examined policies related to school accountability, college and career readiness, and school improvement.

Channa M. Cook-Harvey, is a Senior Researcher at the Learning Policy Institute. Previously she was a high school English teacher and principal, as well as a research associate at the Stanford Center for Opportunity Policy in Education where she studied student-centered learning and social-emotional learning for high school students.

Livia Lam is a Senior Policy Advisor at the Learning Policy Institute. She previously served in the Washington, D.C., office of Senator Maria Cantwell and as a senior labor policy advisor on the U.S. Senate Committee on Health, Education, Labor, and Pensions and the U.S. House of Representatives Committee on Education and the Workforce. Lam also was the deputy director of intergovernmental affairs at the U.S. Department of Labor.

Charmaine Mercer is Director of the Washington, D.C., office and a Senior Researcher at the Learning Policy Institute. Previously she worked at the Alliance for Excellent Education as the Vice President for policy and advocacy in standards, assessments, and deeper learning. Mercer also served in the Office of Planning, Evaluation, and Policy Development at the U.S. Department of Education and the Congressional Research Service. She was a legislative staffer on the U.S. House of Representatives Committee on Appropriations' Labor, Health and Human Services, Education, and Related Agencies Subcommittee and on the House Committee on Education and Labor's K–12 education team.

Anne Podolsky is a Researcher and Policy Analyst at the Learning Policy Institute. As an education lawyer and teacher by training, she has served in legal, research, and policy roles with a variety of organizations, including the John W. Gardner Center for Youth and Their Communities, the New York State Education Department, the Children's Advocacy Institute, and Palantir Technologies.

Elizabeth Leisy Stosich is a Research and Policy Fellow at the Stanford Center for Opportunity Policy in Education. Her research focuses on the school conditions and learning opportunities that support teachers in engaging in ambitious instruction, including opportunities for teachers to use curriculum, assessments, and inquiry-based practices to meet the goals of standards-based accountability policies. Her research interests include standards-based accountability policies, school improvement, teacher collaboration, and teacher preparation and development.