



Using Teacher Evaluation Reform and Professional Development to Support Common Core Assessments

Peter Youngs

February 2013

Center for American Progress



Using Teacher Evaluation Reform and Professional Development to Support Common Core Assessments

Peter Youngs

February 2013

Contents

- 1 Introduction and summary**
- 7 The Common Core State Standards Initiative**
- 11 Previous large-scale reform efforts and past teacher evaluation practices**
- 13 Insights from medicine: Standard of care**
- 15 New advances in teacher evaluation meet the Common Core State Standards Initiative**
- 25 Teacher preparation and licensing**
- 27 Implementing the Common Core Standards and state teacher evaluation reform**
- 31 Policy recommendations and conclusion**
- 34 About the author and acknowledgements**
- 35 Endnotes**

Introduction and summary

The Common Core State Standards Initiative, in its aim to align diverse state curricula and improve educational outcomes, calls for K-12 teachers in the United States to engage all students in mathematical problem solving along with reading and writing complex text through the use of rigorous academic content. Until recently, most teacher evaluation systems in this country did not measure or promote the ability of teachers to practice in these ways.

This report discusses efforts to develop and implement Common Core standards and assessments in the 45 states and the District of Columbia that are initiative members, and outlines how past attempts to enact standards-based reform have been impeded by limitations in teacher evaluation. It also draws on the notion of “standard of care,” from the field of medicine, to note that advances in our understanding of subject matter, pedagogy, how students learn, and technology call for teachers to continually acquire new knowledge and to refine their instructional practices by participating in comprehensive professional development on a regular basis.

Several new approaches to evaluating teachers hold promise for promoting the type of ongoing teacher learning and changes in instruction that would be associated with a professional standard of care in K-12 teaching. These approaches include classroom observation protocols, student surveys, value-added models, and teacher performance assessments. This report details these approaches and explains their potential to strongly support the enactment of the Common Core standards and assessments. At the same time, it also considers a number of challenges connected with implementing each of these. One challenge common to all four approaches is the need for principals to participate in professional development related to the appropriate use of each approach.

Currently, two multistate consortia are developing rigorous student assessments anchored in the Common Core State Standards in mathematics and English language arts, or ELA. In mathematics, the assessments are being designed to measure students’ conceptual understanding, procedural skills and fluency, and

application and problem-solving skills. In English language arts, the assessments will measure students' ability to read and comprehend complex text across the curriculum, write effectively, and conduct and report on research, in addition to measuring their knowledge of vocabulary, grammar, and punctuation. Advocates of this approach argue that this move to Common Core standards and assessments across states and school districts is likely to lead to more rigorous, content-rich instruction and improved student achievement. Skeptics, however, question whether the Common Core reforms, and their eventual effects on teaching and learning, are likely to differ much from past efforts to enact standards-based reform and high-stakes testing policies.

In light of less-than-successful past reform efforts the question is: How are current reforms in teacher evaluation likely to affect the implementation of the Common Core standards and assessments? The medical profession and its notion of “standard of care” can be useful in considering this question. In medicine, the standard of care is a treatment guideline, be it general or specific, which defines appropriate medical treatment based on scientific evidence and collaboration between medical professionals involved in the treatment of a given condition. A key aspect of this definition of standard of care is that appropriate medical practice is based on scientific evidence.

When the notion of standard of care is applied to education and K-12 teaching, it points to the need for all teachers to regularly acquire new knowledge of content, pedagogy, learning theory, and technology by participating in comprehensive professional development with the goal of enacting appropriate and effective instructional practices that will promote student learning. In the past, however, standards-based reform and other improvement efforts faltered in part because teacher evaluation systems failed to meaningfully assess instruction or promote teacher knowledge acquisition.¹ Instead, past teacher evaluation systems tended to focus on a narrow range of teaching practices during classroom observations with virtually all teachers receiving satisfactory ratings. The upshot: teachers generally had little incentive to acquire new knowledge or refine their instructional practices.

To address these shortcomings, several new approaches to teacher evaluation focus much more on instruction, subject matter, and/or teachers' effects on student learning than did past teacher evaluation practices. Classroom observation protocols such as the Framework for Teaching and the Classroom Assessment Scoring System, or CLASS, which was recently explored in the CAP report “Implementing Observations Protocols: Lessons for K-12 Education from the Field of Early Childhood,”² represent important advances over previous observa-

tion instruments. In particular, these protocols distinguish among teachers at various levels of proficiency, provide detailed feedback to teachers, have documented reliability, and have demonstrated empirical relationships with student learning. Similarly, the Tripod Student Perception Surveys, which ask students for their views of the instruction offered by their teachers, also differentiate among teachers at various levels of performance, provide detailed feedback, have evidence of reliability, and are empirically related to student learning.

Another assessment tool, teacher value-added models, or VAMs, differ from observation protocols and student surveys in that they use student achievement data to evaluate teachers' effectiveness, but do not collect or utilize data on instruction. Compared to other innovations in teacher evaluation, value-added models face a host of unique challenges, including issues related to the stability of value-added scores over time, nonrandom assignment of students to teachers, and whether the constructs measured by student assessments are consistent, or "vertically scaled," which means they were intentionally designed around the same constructs or topics across grades.³ At the same time, advocates maintain that value-added models can be combined with other measures, such as observation protocols and student surveys, to assess and promote effective teaching.

Then there are teacher performance assessments such as the National Board for Professional Teaching Standards, or NBPTS, assessments and the edTPA, the newest generation of Teacher Performance Assessment, which differ from the Framework for Teaching and CLASS protocols in that they are content-specific; they focus on multiple lessons from the same unit of instruction; and they feature video clips of instruction, student work samples, and written analyses by teachers of their instruction. Compared to other approaches to teacher evaluation, teacher assessments may have the greatest potential to promote the types of knowledge, including pedagogical content knowledge and instructional practices associated with the Common Core Standards. But these teacher performance assessments are also much more time- and resource-intensive than the other approaches, and districts would confront several challenges in enacting them for use in annual teacher evaluations.

In sum, each of these new approaches to teacher evaluation have the potential to foster the types of rigorous instructional practices called for by the Common Core standards and assessments, but policymakers, district and school administrators, and teachers face a number of challenges in implementing them. In light of these challenges, this report includes several recommendations for how observation protocols, student surveys, value-added models, and teacher performance assess-

ments can be implemented and successfully utilized, along with comprehensive professional development for teachers and principals needed to support the Common Core State Standards Initiative and to measure and promote rigorous instruction. In particular, in enacting these new approaches, districts should consider taking the following steps:

- Utilize school-based instructional coaches in English language and mathematics to provide ongoing professional learning opportunities to teachers related to the Common Core standards and assessments. In addition, districts can support principals' efforts to connect teachers to relevant external professional development based on classroom observation and student survey ratings.
- Ensure the validity and reliability of classroom observation protocols by implementing a standardized approach to training principals and other evaluators, and monitoring their ratings. In addition, districts can train principals to provide timely, meaningful feedback to teachers based on observational data.
- Provide training to principals to ensure that student surveys are administered in a uniform way across schools and classrooms, and work out specific procedures for administering them to young children and students with disabilities. To address teachers' resistance to the use of student surveys, districts can educate them about the value of student survey data. Further, districts can train principals to provide timely, meaningful feedback to teachers based on student survey data.
- Use multiple years of value-added model data in evaluating individual teachers. High-stakes decisions such as dismissal, career ladder promotion, or merit pay should focus only on those teachers who consistently receive bottom-quartile or top-quartile value-added model scores over multiple years.
- Combine aspects—or scaled-down versions—of the National Board for Professional Teaching Standards and edTPA assessments with one or more of the other approaches discussed in this report.

In order to fully answer the question of how the Common Core assessments can be used to promote rigorous classroom instruction, it is important to have some background information on the Common Core State Standards Initiative—particularly its vision of reform, including potential roles for principal leadership and for teacher professional learning in enacting this vision, which this report provides. In addition to exploring that background, this report discusses current efforts by

two multistate consortia to develop content frameworks and student assessments aligned with the Common Core Standards and considers why past efforts aimed at standards-based reform and high-stakes testing have been limited by the nature of teacher evaluation systems.

This report also introduces the notion of standard of care and explains why advances in our understanding of academic content, teaching, learning theory, and technology require teachers to regularly acquire new knowledge and revise their practices by participating in comprehensive professional development.

The approaches to teacher evaluation presented here, when combined with comprehensive professional development for teachers and school leaders and changes in the organization and capacities of school districts, have the potential to support the types of teacher knowledge acquisition and changes in instructional practices called for by the Common Core standards and assessments.

The Common Core State Standards Initiative

The recent efforts to develop Common Core standards and assessments have been led by states and coordinated by the National Governors' Association and the Council of Chief State School Officers, or CCSSO. Teachers, school administrators, and subject-area experts helped devise draft versions of the Common Core Standards in mathematics and English language arts. Several national organizations then provided feedback to the National Governors' Association and the Council of Chief State School Officers on the draft standards, including groups representing teachers, school administrators, postsecondary educators, civil rights groups, English language learners, and students with disabilities. In addition, the draft standards were made available for public comment and close to 10,000 comments were received.⁴

The final versions of the Common Core mathematics and English language arts standards were released in June 2010. The Common Core mathematics standards emphasize the need for students to develop conceptual understanding of mathematics topics, skills and fluency in mathematical procedures, and the ability to apply knowledge and skills in solving mathematics problems. The Common Core English language arts standards focus on students' ability to read complex texts, write effectively, conduct and report on research, and speak and listen in addition to measuring their knowledge of vocabulary, grammar, and punctuation. And through Race to the Top—part of the 2009 American Recovery and Reinvestment Act—the Department of Education provided incentives for states to adopt the Common Core Standards. As of January of this year, 45 states and the District of Columbia had adopted the standards.⁵

In order for the Common Core Standards to influence classroom instruction in the United States in a widespread and sustained way, schools and districts will need to provide opportunities for teachers to participate in high-quality, comprehensive professional development. Such professional development should have a strong content focus, engage teachers as active learners, be of sufficient duration, and involve participation with colleagues.⁶ In particular, it should help teachers

For the Common Core Standards to have a significant impact on instruction, principals will need to set and maintain high expectations for teachers' practices and regularly visit their classrooms.

acquire both content knowledge and pedagogical content knowledge. The latter is defined as knowledge of examples, representations, and explanations that are appropriate for a given concept and specific students; the ability to make productive use of student thinking; and knowledge of likely student misunderstandings.⁷ Many argue that teachers will need strong pedagogical content knowledge in order to enact the instructional practices called for by the Common Core Standards, which is why teachers need sustained professional learning opportunities that address content and content pedagogy.

In addition, for the Common Core Standards to have a significant impact on instruction, principals will need to set and maintain high expectations for teachers' practices and regularly visit their classrooms. This requires school leaders to be knowledgeable about both subject matter and instruction, while being able to converse with teachers about evidence of effective teaching. For many principals, this will necessitate participating in high-quality, comprehensive professional development in order to acquire new knowledge and prepare them to take on new roles.⁸ Further, as discussed below, the Common Core Standards will need to be accompanied by student assessments that measure and promote in-depth subject matter knowledge and high-level skills.

Many educational leaders and scholars have expressed support for the Common Core State Standards Initiative and its vision of reform. One such leader, Randi Weingarten, the president of the American Federation of Teachers, has been a leading advocate of the Common Core Standards. She writes that Common Core Standards “represent the best effort so far to transform today’s patchwork quilt of 50 state standards into one set of strong consistent expectations for what all students should know and learn” and she called for states to “align curriculum, assessments, and professional development for teachers to the standards.”⁹ Similarly, Chester Finn, president of the conservative-leaning Thomas B. Fordham Institute, notes that the Common Core Standards are “good, solid—indeed very ambitious—academic standards for primary and secondary schooling” and he adds that in order “to gain traction, they’ll need to be joined by solid curricula, effective instruction, and quality testing.”¹⁰

In fact, the Common Core State Standards Initiative has the potential to address two characteristics of schooling in the United States that some scholars argue pose major obstacles to both student and teacher learning. First, the lack of common academic standards leads to unequal coverage of core subject matter content across classrooms and schools and creates unequal educational opportunities

for many students.¹¹ Some contend that moving to a more rigorous, common curriculum in mathematics and English language arts will reduce inequities in educational opportunity while promoting high-level learning for all students.¹² Second, the lack of common standards inhibits efforts by teachers and principals to work collaboratively to analyze and improve instruction.¹³ As David K. Cohen, a professor at the University of Michigan, observes, such standards, when combined with well-aligned student assessments, can help educators “to define quality in students’ work and valid evidence of quality” and can give educators “a common vocabulary with which they can work with each other to identify, investigate, discuss, and solve problems of teaching and learning.”¹⁴

Content frameworks and student assessments

While Weingarten, Finn, and others have offered support for the Common Core State Standards Initiative, they rightly note that the new standards need to be accompanied by student assessments that are structured to improve the quality of teaching and student learning. Through Race to the Top, the Department of Education has provided funding to two multistate consortia that are currently developing student assessments tightly linked to the Common Core Standards. One of these consortia is the Partnership for Assessment of Readiness for College and Careers, or PARCC, and the other is the Smarter Balanced Assessment Consortium, or SBAC.

Both consortia plan to create summative student assessments for implementation by states in the 2014–15 school year. These assessments feature a combination of short- and extended-response items, performance tasks, and technology-enhanced tasks. In 2012 both the Partnership for Assessment of Readiness for College and Careers and the Smarter Balanced Assessment Consortium released prototype assessment items and tasks, as well as content framework/specifications designed to serve as a link between the Common Core Standards and their respective assessments, and to increase the likelihood that the assessments will represent the full range of content featured in the standards.

The Partnership for Assessment of Readiness for College and Careers’ Model Content Frameworks in Mathematics and the Smarter Balanced Assessment Consortium’s Content Specifications in Mathematics support the Common Core focus on conceptual understanding, procedural skill and fluency, and application and problem solving in mathematics in several ways.¹⁵ First, these mathemat-

ics frameworks both delineate opportunities for students to focus in-depth on particular mathematics topics and skills. Second, they identify connections among mathematics topics and skills within and across grade levels. Third, they make clear linkages between mathematics content and student mathematical practices. And finally, both mathematics frameworks feature examples of culminating standards or summative assessment targets.

Similarly, the partnership's Model Content Frameworks in English language arts and the Smarter Balanced Assessment Consortium's Content Specifications in English language arts address the Common Core emphasis on reading complex texts, writing effectively, conducting and reporting research, speaking and listening effectively, and language use.¹⁶ Both frameworks, for example, delineate specific reading demands, writing emphases, and research tasks for each grade; identify connections among standards regarding reading, writing, and research; make linkages across grades; and articulate literacy standards with regard to instruction in other subjects.

The prototype assessment items and performance tasks recently released by The Partnership for Assessment of Readiness for College and Careers¹⁷ and the Smarter Balanced Assessment Consortium¹⁸ are designed to convey the rigor and complexity of the summative assessments that will be released in the 2014–15 school year. Both consortia have encouraged educators to draw on the prototypes in planning changes in instruction that will help students meet the expectations of the Common Core standards and assessments. But the Partnership for Assessment of Readiness for College and Careers and the Smarter Balanced Assessment Consortium also note that the prototypes are not to be used as sample tests because the assessments are still undergoing an extensive review process, including field testing.

When it comes to any reform effort, there are a number of overarching questions including: Will these new reforms prove any more successful than previous efforts to improve classroom instruction and ultimately student outcome? And, more specifically, what is the likelihood that the Common Core standards and assessments will lead to extensive changes in mathematics and English language arts instruction in the coming years? Let's explore both these questions further.

Previous large-scale reform efforts and past teacher evaluation practices

Several scholars have compared the Common Core State Standards Initiative to past approaches to standards-based reform and have made some interesting observations. In the context of standards-based reform efforts around teacher evaluation in the 1990s, for example, was one result that demonstrated that tenured teachers in many districts had little incentive to acquire new knowledge or change their instruction.¹⁹ Instead, norms supporting privatized practice were customary in most schools; that is, teachers were left alone by administrators to practice as they chose. Thus, while standards-based reform efforts at that time called for teachers to participate in professional learning, set high expectations for all students, and alter their instruction in significant ways, prevailing teacher evaluation practices sent teachers a much different message: The status quo was acceptable and their annual evaluations were very likely to be the same from one year to the next regardless of any efforts they made to improve their teaching—certainly not a formula for driving change.

In the 2012 Brown Center Report on American Education, Senior Fellow Tom Loveless argued that the Common Core State Standards Initiative is likely to have similar—that is to say, negligible—effects on student achievement as the Goals 2000 report—part of the work done under President Clinton’s “The Goals 2000: Educate America Act” created to lead to improved student learning—did in the 1990s. Loveless draws on data from the National Assessment of Educational Progress, or NAEP, to contend that in the past, state scores on the assessment were not related to the quality of state standards nor their rigor—as indicated by where the passing score is set for proficiency levels on state assessments. While some view the Common Core State Standards Initiative as a way to reduce inequalities in curricula and student achievement across states, Loveless also notes that National Assessment of Educational Progress score differences are much greater within states than across states.²⁰

In another study, Bill Schmidt and Richard Houang, both from Michigan State University, conducted a content analysis of the Common Core Standards in mathematics in order to compare them to mathematics standards in high-performing nations. The authors conclude that the U.S. Common Core mathemat-

ics standards are both focused and coherent. Their measure of focus refers to whether various mathematics topics are covered in a relatively low number of grades (greater focus) or in a relatively high number of grades (less focus) across grades 1 through 8. Their measure of coherence refers to whether the sequence in which various mathematical topics are covered is consistent with the logical structure of mathematics. Finally, Schmidt and Houang argue that the Common Core Standards in mathematics are rigorous “as indicated by the consistency in topic coverage” between these standards and the mathematics standards in high-performing nations “especially at eighth grade.”²¹

While comparisons to National Assessment of Educational Progress scores, existing state content standards, and content standards in other nations are useful, this report notes that past practices in teacher evaluation in the United States help explain the limited effects of previous attempts at standards-based reform. For several decades, evaluations of practicing K-12 teachers typically consisted of classroom observations by a principal or another administrator followed by short meetings. The administrator would often use a checklist of behaviors to assess the teacher’s performance. In many school districts, untenured teachers were observed two or more times per year while tenured teachers were usually observed just once a year or less.

Practitioners and policymakers raised three specific concerns about such approaches to teacher evaluation. First, in most districts, administrator ratings lacked variability—in other words, virtually all teachers were judged to be performing at acceptable levels. Second, the observation instruments and observers typically failed to address teachers’ content knowledge, knowledge of content pedagogy, or the effects on student learning. Instead, these instruments primarily focused on generic teacher behaviors. Third, observations by principals and other administrators were usually carried out in a cursory way and had very little long-term impact on classroom instruction.²²

In light of concerns raised by Loveless and others, and based on the historical record, one has to wonder is the Common Core State Standards Initiative likely to experience a similar fate as standards-based reform in the 1990s? Some believe that advances in teacher evaluation have the potential to strongly support enactment of the Common Core standards and assessments, as well as widespread changes in instructional practice. Before describing these advances and considering their possible impact on the Common Core State Standards Initiative, let’s first take a closer look at the notion of standard of care and explain why standards-based reform efforts, such as the Common Core, necessitate changes in teacher evaluation.

Insights from medicine: Standard of care

In medicine, a standard of care is a treatment guideline, general or specific, which defines appropriate treatment based on scientific evidence and collaboration between medical professionals involved in the treatment of a given condition. In legal terms, it refers to the level at which an ordinary, sensible medical professional with the same training would practice under similar circumstances. The notion of standard of care is relevant in malpractice cases when a physician's decisions and/or practices are called into question.²³ In several recent cases, "courts have frequently upheld that the standard of care is what a minimally competent physician in the same field would do in the same situation, with the same resources."²⁴

How might this concept for doctors apply to K-12 teachers? A key aspect of this definition is that appropriate medical practice is based on scientific evidence. In the field of medicine, however, knowledge of health, illness, and appropriate treatments is constantly evolving. As a result, medical professionals must stay abreast of developments in their general fields of practice and specialty areas. Evidence in support of new medical treatments frequently comes from randomized controlled trials. But such evidence can also emerge from meta-analyses and other reviews that feature a wide range of medical studies, including quasi-experiments and natural variation studies.

In the field of K-12 education, rapid advancements in our understanding of subject matter, how children and adolescents learn, instruction, classroom assessment, and technology call for teachers to remain well-informed about developments in their general fields, as well as their areas of specialization. At the same time, university-based preparation programs and alternative certification programs must provide comparable, high-quality training to all teaching candidates and state licensure policies must establish guidelines for such training.

In addition, practicing teachers must have opportunities to participate in high-quality, comprehensive educational opportunities to improve their professional practice and have the chance to read and discuss research on a regular basis.

Teachers also need access to school-based instructional coaches, principals, and teacher colleagues as they attempt to integrate new knowledge and instructional strategies into their classroom practices. As in medicine, evidence regarding new educational practices will sometimes come from randomized controlled trials. But such evidence will also arise from meta-analyses and other reviews consisting of quasi-experimental studies and natural variation studies.

In sum, applying a standard of care in K-12 education would require all teachers to continually acquire new knowledge and make alterations to their instructional practices, in the way that highly effective teachers already do while at the same time specifying a minimum threshold of teaching practice. While the theory of action underlying standards-based reform in the 1990s placed similar expectations on teachers, this policy was weakened to some extent by prevailing teacher evaluation practices at the time. That is to say, given that administrators observed most tenured teachers once a year or less, most evaluations were perfunctory and nearly all teachers received satisfactory ratings or better. As a result, experienced teachers had little incentive to acquire additional knowledge or change their instruction, as called for by Goals 2000.

Several new approaches to teacher evaluation, more so than in the past, focus increasingly on subject matter, instruction, and/or teachers' effects on student learning. In the next section, this report describes four specific evaluation approaches and considers how the use of each is likely to affect the enactment of Common Core standards and assessments.

New advances in teacher evaluation meet the Common Core State Standards Initiative

In recent years concerns with shortcomings in traditional teacher evaluation systems have led district, state, and federal policymakers to focus on a number of new approaches to evaluating teachers. These include classroom observation protocols, student surveys, value-added models, and teacher performance assessments. The brief descriptions of each of the four approaches that follow consider ways in which each has the potential to support the Common Core State Standards Initiative—including the move to more rigorous, content-rich instruction—and discuss challenges that districts will face in enacting them for use in annual teacher evaluations. In addition, this report explores the role of teacher preparation programs and state licensure policies in ensuring that teaching candidates acquire knowledge and implement instructional practices called for by the Common Core Standards.

Observation protocols

Classroom observations have long been a staple of most teacher evaluation systems. In many cases, however, the observations and rubrics used in teacher evaluations are not grounded in clearly articulated models or frameworks of effective teaching.²⁵ In contrast the Framework for Teaching, the Classroom Assessment Scoring System, or CLASS, and other newer observation protocols are explicitly grounded in models of effective teaching and/or sets of teacher performance standards. These protocols are used with rubrics that differentiate among various levels of teacher performance. Evaluators, including principals, district administrators, and other teachers, are expected to provide detailed records of what they observe organized around the standards and rubrics. Thus, these observations are characterized as evidence-based and are less likely to rely on subjective judgments that are unsupported by evidence or concrete examples. Data from these observations can then be used to:

- Provide detailed feedback to teachers
- Plan tailored opportunities for professional learning for teachers
- Contribute to overall ratings of teacher performance when combined with other data

The Framework for Teaching—developed by Charlotte Danielson, an education consultant from Princeton, New Jersey—whose design for assessing instruction is one of the most-used in the United States, can be used with K-12 teachers in core content areas such as mathematics, English language arts, science, history/social studies, and world languages. Danielson’s framework features four domains: planning and preparation, classroom management, teaching for student learning, and professionalism.²⁶ Trained evaluators assess teachers in all four domains with classroom observations focusing on teacher performance in two of these domains: classroom management and teaching for student learning. In most districts that use the Framework for Teaching—or modified versions of it—teachers are observed four to six times per school year. Research from Cincinnati and Chicago schools indicates that teachers’ observation ratings based on the Framework for Teaching are related to their effects on student achievement gains.²⁷

On the other hand, the Classroom Assessment Scoring System teacher assessment model was first developed by Robert Pianta and his colleagues for use in early childhood settings. In a recent Center for American Progress report Pianta outlined the ways that CLASS can be employed across subject areas and grade levels to assess teachers’ interactions with children in three broad domains: classroom organization, instructional support, and emotional support. These domains are common from preschool through the 12th grade. CLASS identifies particular types of interaction within each domain that vary by grade. It assesses “effective teacher-student interactions across pre-K–12 in a way that is sensitive to important developmental and context shifts that occur as students mature.”²⁸ In research on CLASS and a precursor to CLASS—the Classroom Observation System—Pianta and colleagues have reported relationships between teachers’ ratings based on these observation protocols and achievement gains at the preschool and elementary school levels.²⁹

The Framework for Teaching, CLASS, and other newer observation protocols have the potential to support implementation of the Common Core standards and assessments in a number of ways. First, these observation protocols provide teachers and administrators with a common language for analyzing and documenting teaching practices. Second, these protocols enable teachers and evalua-

tors to use evidence to describe teachers' strengths and diagnose areas in need of improvement. Third, the observation results can be used to make decisions about professional development for individual teachers. At the same time, districts face challenges in enacting these new observation protocols. Districts must be able to ensure the validity and reliability of the protocols by implementing a standardized approach to training evaluators and monitoring evaluators' ratings. In addition, using new protocols requires principals to demonstrate instructional leadership, provide timely and meaningful feedback to teachers, and connect teachers to relevant opportunities for professional learning and development.

Despite their strong potential to contribute to changes in instruction, it is important to note that the Framework for Teaching and the Classroom Assessment Scoring System are both generic with regard to subject matter and, as noted, can be used across grade levels and content areas. On one hand, this leads to efficiencies for districts in that they only need to train evaluators to use one observation protocol. At the same time, it also means that such observation protocols are not able to measure or directly promote teachers' pedagogical content knowledge. If, as many believe, pedagogical content knowledge is necessary in order to implement the Common Core in mathematics and English language arts, such observation protocols would need to be supplemented by other forms of teacher evaluation.

Student surveys

A second approach to teacher evaluation, student surveys, has been gaining currency over the past few years. A prominent example, the Tripod survey developed by Ronald Ferguson, has recently been implemented in several school districts.³⁰ The Tripod survey asks students about their perceptions regarding seven constructs pertaining to their teachers' instructional practices: care, control, clarify, challenge, captivate, confer, and consolidate. The elementary and secondary versions of the Tripod surveys address all seven constructs although the survey items for each construct vary to some degree from elementary school to secondary school. This report focuses only on two of the seven constructs—clarify and challenge—because student responses to survey items related to these constructs provide insight into aspects of teachers' pedagogical content knowledge.

At both the elementary and the secondary levels, the Tripod surveys include the following items for the “clarify” construct: “My teacher explains difficult things clearly; My teacher has several good ways to explain each topic that we cover in

Despite their strong potential to contribute to changes in instruction, Framework for Teaching and the Classroom Assessment Scoring System are both generic with regard to subject matter and can be used across grade levels and content areas.

this class; My teacher knows when the class understands and when we do not” and “If you don’t understand something, my teacher explains it another way.” And at both schooling levels, the Tripod surveys include the following item for the “challenge” construct: “In this class, my teacher accepts nothing less than our full effort.” The other three items for the challenge construct in the elementary survey address students’ perceptions of their teacher’s expectations for student effort, for example: “My teacher pushes everybody to work hard.” The other seven items for this construct in the secondary school survey address students’ perceptions of whether their teacher asks them to explain their answers and of their teachers’ expectations for how hard they will work and how much they will learn, for example, “My teacher asks students to explain more about answers they give.”

The Tripod survey items that address the “clarify” and “challenge” constructs reveal information about teachers’ pedagogical content knowledge. That is to say, these items provide relevant data about teachers’ knowledge of examples, representations, and explanations that are appropriate for a given concept and specific students; their ability to make productive use of student thinking; and their ability to respond when students have misunderstandings. Thus, the Tripod surveys have the potential to measure and promote a key aspect of teacher knowledge, pedagogical content knowledge, which seems associated with effective enactment of the Common Core standards and assessments. In addition, the Measures of Effective Teaching project recently reported that when observation ratings from protocols such as the Framework for Teaching or the Classroom Assessment Scoring System were combined with Tripod survey data, the combined measures of teacher performance had higher correlations with teachers’ effects on achievement gains than the correlations between the observation ratings—without the Tripod survey data—and the measures of effects on achievement gains.³¹

While the Tripod surveys have strong potential to promote changes in instruction, districts confront a number of challenges in implementing them. In terms of survey administration, districts will need to ensure that the surveys are administered in a uniform way across schools and classrooms. They will need to make decisions about the sampling of classes at the secondary level and how to administer the surveys to young children or those with disabilities. With regard to teachers, some may initially be reluctant to have a significant part of their evaluation based on students’ perceptions of their practices. Finally, as with the observation protocols, principals will need to use student survey data to provide timely, meaningful feedback to teachers, and link the teachers to relevant professional learning opportunities.

Value-added models

A third approach to teacher evaluation that has become much more prevalent is the use of value-added models, or VAMs, to assess the effects of individual teachers on student learning. Value-added models are statistical models that attempt to explain the contribution of particular teachers to student achievement gains over time. In order to do this, such models require at least two years of student test scores and they typically control for other student and school characteristics such as the percentage of students from low-income families or the percentage of students who are English language learners. Value-added models are designed to take account of and control for differences among students with regard to their family situations and other outside-school factors that can influence their learning.³² Through *Race to the Top*, the Department of Education provided incentives for states to allow for the use of student achievement data in teacher evaluation. As a result, many states and districts have begun to enact value-added models for this purpose.

Some educational leaders and scholars have expressed strong support for the use of value-added models in summative teacher evaluation.³³ At the same time, other educational leaders and scholars have raised significant concerns about the use of value-added models for this purpose.³⁴ These include concerns about the stability of individual teacher value-added scores over time, nonrandom assignment to students, vertical scaling of constructs across grades, and the evaluation of teachers in nontested grades and subjects. Let's first look at these four concerns as they relate to the use of value-added models, each of which poses a challenge to districts that implement them. That will be followed with a discussion of how value-added models can be used, along with other measures, to evaluate teachers in a given district; and how their use could support the enactment of the Common Core standards and assessments.

Several studies have found that an individual teacher's value-added score is not very stable from one year to the next. In particular, research indicates that 40 percent or more of the teachers who perform in the top quartile (top 25 percent) in their district in one year will no longer be in the top quartile the following year.³⁵ At the same time, the correlation between an individual teacher's "underlying value-added score," based on three or more years of data, and their score for any given year is often much higher than the correlation between their value-added scores from one year to the next.³⁶ This means that (a) school districts should consider only using multiple years of value-added data in evaluating individual teachers; and (b) high-stakes decisions such as dismissal, career ladder promo-

tion, or merit pay should focus only on those teachers who consistently receive bottom-quartile or top-quartile value-added scores over multiple years.

Observable student characteristics that can affect achievement include prior achievement, socioeconomic status, and English language proficiency.

A second concern involves the assignment of students to teachers. Students are rarely assigned to teachers at random, instead, the assignment of students to teachers is influenced by various factors, some of which can be observed and some that are unobservable. Observable student characteristics that can affect achievement include prior achievement, socioeconomic status, and English language proficiency. Unobserved student characteristics that can affect achievement include the influence of parents and the influence of peers. When value-added models fail to take sufficient account of the possible effect of unobserved student characteristics, this can lead to biased estimates of teacher effectiveness. Recently, Jesse Rothstein of the University of California, Berkeley, and Derek Briggs of the University of Colorado, conducted separate studies that showed bias in value-added estimates in prominent jurisdictions in the United States due to this very reason.³⁷ Rothstein notes that this bias is lessened in a value-added model that controls for one year of prior achievement and is further reduced when additional measures of prior achievement are included as controls.³⁸ This provides further evidence that districts should consider only using multiple years of value-added data in evaluating individual teachers.

A third issue arising from the use of value-added models has to do with vertical scaling. Some researchers have defined vertical scaling as occurring when student assessments are intentionally built around the same constructs but differ in content across grades.³⁹ Meanwhile, other scholars have identified reasons why vertical scaling can be difficult to achieve.⁴⁰ Bill Schmidt and colleagues, for example, used data from the Third International Mathematics and Science Study, or TIMSS, to demonstrate that the mathematical “knowledge and skills needed by students to respond correctly to achievement test items are qualitatively different across grades.”⁴¹ They also show that the same phenomenon—construct-shift across grades—is evident when analyzing state of Michigan mathematics content standards.⁴² With regard to the possible use of the Common Core assessments in teacher value-added models, states and districts will need to show that these assessments measure the same constructs across grades.

A fourth concern about value-added models pertains to teachers in nontested subjects or grades. In many states, this group includes history/social studies teachers, world language teachers, and specialist teachers—for example, art, music, physical education. This group also includes high school mathematics and English language arts teachers in nontested grades, as well as pre-K, kindergarten, and,

in many states, first-grade and second-grade teachers. Some districts are using schoolwide value-added data to assign scores to such teachers; other districts rely only on data from classroom observations to generate summative evaluations for teachers. Districts' decisions regarding the use of schoolwide value-added data in evaluating teachers in nontested subjects and grades can influence:

- Expectations that are placed on such teachers—for example, to contribute to student learning in tested subjects
- Collaboration between tested and nontested teachers
- Teacher morale

In light of these four concerns, this report recommends that districts only use value-added models for summative teacher evaluation in combination with other sources of data such as data from classroom observations and/or student surveys. In addition, high-stakes decisions involving value-added scores—such as dismissal, career-ladder promotion, and merit pay—should focus only on those teachers who regularly receive bottom-quartile or top-quartile value-added scores over multiple years. The use of value-added models under these conditions can support implementation of the Common Core Standards and assessments in two main ways. First, value-added models are typically used to evaluate elementary and middle school teachers in mathematics and English language arts. As a result, the tested grades and subjects are likely to be very similar to those addressed by the Common Core. Second, when used appropriately—for example, to promote or dismiss teachers who are consistently in the top or bottom quartiles—value-added models can potentially serve as an effective form of accountability for all teachers in tested subjects and grades. In addition, the emphasis in the Common Core on measuring certain constructs in mathematics and English language arts across grades increases the likelihood that the Common Core assessments will be vertically scaled.

Teacher performance assessments

The fourth approach to teacher evaluation discussed in this report is teacher performance assessment, exemplified by the National Board for Professional Teaching Standards assessments, designed for experienced teachers, and the newer edTPA, designed for use with teaching candidates.

The National Board for Professional Teaching Standards has developed a series of content-specific assessments for experienced teachers at different schooling

levels and in various content areas—for example, elementary (generalist), middle school English language arts, high school mathematics.⁴³ The purpose of the National Board for Professional Teaching Standards assessments is to identify accomplished or expert teachers, and some studies have reported associations between earning national board certification—passing the National Board for Professional Teaching Standards assessment—and teacher effectiveness as measured by effects on student achievement.⁴⁴

Each of the National Board for Professional Teaching Standards assessments consists of several portfolio and assessment center exercises.⁴⁵ In the middle school English language arts assessment, for example, candidates compile portfolios that include students’ responses to a work of literature, writing samples from students, and video clips of the teacher exploring important topics in language arts with both a small group of students and then with the entire class. For each portfolio entry, the teacher must also submit a written commentary in which they contextualize the entry and analyze their teaching. For the assessment center exercises, candidates for the middle-school English language arts assessment plan a unit based on multiple works of literature and construct an argument about a teaching scenario based on professional articles about writing instruction. In addition, they use professional articles on English language learning and bilingualism to describe how they would plan instruction for a particular English language learner and what reading selections they would recommend to parents of middle school students.

Similarly, edTPA is a series of content-specific assessments designed for teaching candidates at different schooling levels and content areas. In contrast, though, the edTPA is designed for use in making initial licensure decisions. The edTPA requires teaching candidates to complete and submit several materials that are integrated around three to five lessons from the same instructional unit. These include a description of their teaching context, a set of lesson plans, video clips of instruction during the unit, student work samples, and written reflections on planning, instruction, and assessment of student progress.⁴⁶ The edTPA requirements are highly structured, and teaching candidates are given detailed instructions on how to meet them. Each completed assessment provides information about the organization of the teacher’s curriculum, the appropriateness of their instructional decisions, the range of pedagogical strategies they use, the quality of their assignments, and their ability to assess student learning and make changes based on evidence of such learning.

In a recent report for the Center for American Progress, called “Evaluating Teacher Effectiveness: How Teacher Performance Assessments Can Measure and

Improve Teaching,” Linda Darling-Hammond argues that completing a teacher performance assessment can help teachers improve their instructional practices. She notes that those teachers who complete the National Board for Professional Teaching Standards assessments report that analyzing student work and their own work in relation to professional standards helps them to better assess student learning and evaluate the effects of their own actions. Teachers also have to adopt new practices that are called for in the standards and assessments.⁴⁷ Darling-Hammond adds that districts can use teacher performance assessments “to provide an evidence-based methodology for making systematic decisions about recruitment, employment, professional development, and career development.”⁴⁸

In sum, the National Board for Professional Teaching Standards and edTPA assessments are different from the other approaches to teacher evaluation discussed above in that they are content-specific assessments designed to measure and promote teachers’ use of pedagogical content knowledge in instruction. The National Board for Professional Teaching Standards and edTPA assessments are based on subject-specific teaching standards, and they evaluate a teacher’s ability to plan and engage in the types of rigorous, content-rich instructional practices called for by the Common Core Standards and assessments. Despite their potential, these assessments—at least as presently designed—can’t be easily implemented by districts for use in annual teacher evaluations. This is because they are much more time- and work-intensive for practicing teachers than classroom observations, student surveys, and value-added models. In addition, while the national board standards and edTPA assessments provide very detailed information about single instructional units, they provide less information than the other three approaches about a teacher’s performance over the course of a school year. As a result, this report recommends that districts explore ways to combine aspects—or scaled-down versions—of the National Board for Professional Teaching Standards and the edTPA assessments with one or more of the other approaches discussed here for use in annual teacher evaluation.

Teacher preparation and licensing

As discussed above, in order for practicing teachers to acquire new knowledge and implement instructional practices called for by the Common Core Standards, they need regular opportunities to participate in high-quality comprehensive professional development. In addition, in order to ensure that teaching candidates are prepared to implement such instructional practices as they enter the profession, it is important for teacher education programs and state licensure policies to address the Common Core standards and assessments.

University-based preparation programs and alternative certification programs can address the Common Core Standards and assessments in mathematics and English language arts content courses, mathematics and English language arts methods (pedagogy) courses, and courses on student assessment. And they can arrange student teaching placements for candidates in schools that have successfully implemented the Common Core.

State licensure policies can address the Common Core Standards and assessments by requiring elementary and secondary teaching candidates to complete tests of subject-matter knowledge, along with teacher performance assessments such as the edTPA that address pedagogical-content knowledge. In many states, tests of knowledge of mathematics and English language arts will likely need to be revised to incorporate items that are aligned with the Common Core Standards in these areas. As discussed above, the edTPA is a content-specific assessment that measures teachers' pedagogical content knowledge in particular areas and evaluates teachers' ability to plan and engage in many of the instructional practices called for by the Common Core.

Implementing the Common Core Standards and state teacher evaluation reform

The 45 states and the District of Columbia that are implementing the Common Core have enacted different strategies, processes, and structures to support their efforts. Of particular note, two states—Delaware and Kentucky—have begun working on improving professional learning for all educators, tying educator evaluation to the Common Core, and aligning materials and resources to support high-quality instruction so students can master the rigor embedded in the Common Core. We are highlighting Delaware because it was one of the first two states, along with Tennessee, to receive Race to the Top funds. Likewise, Kentucky is highlighted because it was the first state to adopt the Common Core Standards.

Delaware

In 2010 Delaware became one of two Phase 1 states to receive federal Race to the Top funds.⁴⁹ Prior to winning Race to the Top, the state had already moved toward adoption of the Common Core Standards and enacted changes in educator evaluation policy that linked student growth to overall educator performance ratings. The Delaware Department of Education has supported school districts in a number of ways as they have moved to implement the Common Core and the new approaches to educator evaluation, using Race to the Top funding as a catalyst to do so since fall 2010.

In support of one effort, known as Common Ground for the Common Core, the Delaware Department of Education is convening representatives from each of the more than 200 schools in the state for a two-day meeting related to the Common Core Standards in March 2013.⁵⁰ Each school has been invited to send a team of three to six individuals, including principals, English language arts lead teachers, mathematics lead teachers, and lead teachers from other content areas to the March meeting. At this gathering, participants will discuss strategies for supporting teachers' efforts to incorporate the Common Core Standards into their instruction. Following the meeting each school will be expected to create a professional

development plan that features specific teacher learning activities associated with the Common Core Standards. The Delaware Department of Education will support schools' efforts to enact their plans and implement the standards by convening state and regional meetings and offering webinars related to the Common Core.

Shelley Rouser, special assistant to the Delaware secretary of education, notes that, "Most schools throughout Delaware have done a lot of work with professional learning communities. This has provided a strong foundation for implementing the Common Core and making significant changes in instruction."⁵¹

"The student growth component is the determining factor of a teacher's effectiveness rating—it acts as a gatekeeper."

Christopher Ruszkowski, chief officer of the Teacher and Leader Effectiveness Unit in the Delaware Department of Education

In terms of linkages between the implementation of the Common Core Standards and teacher evaluation reform, Delaware has adopted a modified version of the Charlotte Danielson's Framework for Teaching as the state's teaching standards and classroom observation instrument. Each teacher is evaluated in five domains—planning and preparation, the classroom environment, instruction, professional responsibilities, and student growth. The state fully implemented its new teacher evaluation system in the 2012–13 school year.

Christopher Ruszkowski, chief officer of the Teacher and Leader Effectiveness Unit in the Delaware Department of Education, explains that, "The student growth component is the determining factor of a teacher's effectiveness rating—it acts as a gatekeeper. If an educator is rated "satisfactory" in all aspects of practice related to classroom observation, but is rated "unsatisfactory" with regard to student growth, then his/her overall rating will be "needs improvement," for example."

The Delaware Department of Education provides extensive professional development for principals related to the new teacher evaluation system. New principals participate in four to five days of state-level training related to teacher evaluation in their first year, while returning principals participate in two to three days of state training each year, as well as district-provided training.⁵² The state education department also provides development coaches who work full time providing job-embedded support to principals related to teacher evaluation and the Common Core Standards.⁵³

A not-insignificant factor that has contributed to Delaware's progress with regard to the Common Core and teacher evaluation reform is the state's small size. As a consequence, Delaware has the ability to convene administrators and teachers from all of its 19 school districts and, when needed, all 200-plus schools in regular meetings. Another contributing factor is current Gov. Jack Markell's (D) unwavering focus on making Delaware's schools among the best in the nation.

Kentucky

In 2010 Kentucky became one of the first states to adopt the Common Core Standards. Since then, the Kentucky Department of Education has employed several strategies to integrate high-quality professional learning for teachers related to these standards with teacher evaluation reform. In the 2010–11, 2011–12, and 2012–13 school years, the state’s department of education held monthly network meetings at eight locations around the state.⁵⁴ Each of Kentucky’s 174 school districts was invited to send three to four mathematics specialists, three to four English language arts specialists, three to four school administrators, and three to four district administrators to each of these meetings.

At these meetings, participants analyzed the Common Core Standards, developed formative and summative assessments based on the standards, and discussed strategies for supporting teachers’ efforts to incorporate the standards into their instruction. In each district, the math specialists, English language arts specialists, school administrators, and district administrators who attend the network meetings form a leadership team. According to Felicia Smith, associate commissioner at the Kentucky Department of Education Office of Next-Generation Learners, “Each district leadership team is responsible for determining how to structure and provide local professional development for teachers and principals related to the Common Core Standards.”⁵⁵

The Kentucky Department of Education supplements the work of the district leadership teams in a number of ways. First, 16 department of education content specialists in mathematics and English language arts provide support to districts across the state related to the Common Core Standards and teacher professional development. Second, state department of education staff worked with staff from Learning Forward—an international nonprofit organization that provides technical assistance to states, districts, and other jurisdictions related to professional learning—to set up a Continuous Improvement Instructional Technology System.

The Continuous Improvement Instructional Technology System features several resources associated with the Common Core Standards. According to Karen Kidwell, director of the division of program standards in the Department of Education Office of Next-Generation Learners, “These resources include lesson plans, videotapes of instruction, videotapes of teacher conversations about the standards, protocols for curriculum development and alignment, and formative assessments. The CIITS resources are available to teachers and administrators throughout Kentucky.”⁵⁶

In terms of linkages between the implementation of the Common Core Standards and teacher evaluation reform, Kentucky has adopted Danielson’s Framework for Teaching as the state’s teaching standards and classroom observation instrument.⁵⁷ During the 2012–13 school year, the network meetings for mathematics and English language arts specialists and school and district administrators are focusing on the use of the Danielson Framework in observing and evaluating teachers and on making explicit connections between the framework and the Common Core Standards. The state plans to pilot its new teacher evaluation system in the 2013–14 school year and to implement it statewide in the 2014–15 school year.⁵⁸ The system will differentiate teachers across at least four levels of effectiveness. In addition to classroom observations, the state’s new teacher evaluation system is likely to include student growth measures such as teacher value-added measures.

Kentucky has received funding from the Gates Foundation to support integration of the Common Core Standards and assessments, and teacher evaluation measures. During the 2012–13 school year, several Kentucky districts are working on integration issues as part of this grant. The state education department’s Felicia Smith notes that the state’s efforts to integrate the Common Core State Standards Initiative with teacher evaluation reform is due in part to the fact that “in Kentucky, this work is housed in one office, the Office of Next-Generation Learners. This promotes coordination across the areas of Teaching and Learning, and Professional Growth and Effectiveness.”⁵⁹

Policy recommendations and conclusion

The Common Core Standards and assessments call for teachers to engage all students in high-level thinking, mathematical problem-solving, and reading and writing complex text. While past teacher evaluation practices did not assess or encourage teachers' ability to teach in these ways, new approaches to teacher evaluation have the potential to support changes in instruction. These include classroom observation protocols, student surveys, value-added models, and teacher performance assessments. In order for the potential of these new approaches to be fully realized, K–12 education would need to adopt a standard of care that calls for all teachers to continually acquire new knowledge about academic content, pedagogy, learning theory, and technology by participating in high-quality professional learning opportunities, along with revising their instructional practices based on their newly acquired knowledge. In addition, in enacting these new approaches, districts should consider taking the following steps:

- Districts can utilize instructional coaches in English language arts and mathematics to provide ongoing professional development to teachers related to the Common Core Standards and assessments. In addition, districts can support efforts by principals to connect teachers to relevant external professional development based on classroom observation and student survey ratings. In order to effectively implement the Common Core Standards and assessments, teachers must have opportunities to participate in high-quality, comprehensive professional development and to read and discuss research related to the Common Core on a regular basis. In addition, they need access to school-based instructional coaches, principals, and teacher colleagues as they attempt to integrate new knowledge and instructional strategies into their classroom practices.
- Districts can ensure the validity and reliability of observation protocols by implementing a standardized approach to training evaluators and monitoring evaluators' ratings. In addition, districts can support principals' efforts to provide timely, meaningful feedback to teachers based on observation data and connect them to tailored opportunities to improve their professional practice. Given that observation protocols such as the Framework for Teaching and CLASS are generic—not subject-specific—and, thus, do not measure or directly promote teachers' pedagogical content knowledge, districts should

supplement them with other forms of teacher evaluation such as student surveys and teacher performance assessments that measure and promote teachers' pedagogical content knowledge.

- Districts can ensure that student surveys are administered in a uniform way across schools and classrooms, and work out specific procedures for administering them to young children or those with disabilities. To address possible teacher resistance to the use of student surveys, districts can educate them about the value of student survey data. Further, districts can support the efforts of principals to provide timely and meaningful feedback to teachers, based on student survey data and connect them to relevant professional development. Since the items in the Tripod student surveys addressing the “clarify” and “challenge” constructs reveal information about teachers' pedagogical content knowledge, districts should consider using such surveys to supplement observation protocols and value-added measures.
- Districts should consider only using multiple years of value-added model data in evaluating individual teachers to offset variation from one year to the next. Researchers have raised concerns about the use of value-added measures in summative teacher evaluation. While studies have found that a teacher's value-added score is not stable from one year to the next, the correlation between an individual teacher's “underlying value-added score,” based on three or more years of data, and their value-added score for any given year is often much higher than the correlation between their value-added scores from one year to the next. High-stakes decisions such as dismissal, career-ladder promotion, or merit pay should focus only on those teachers who consistently receive bottom-quartile or top-quartile value-added scores over multiple years.
- Districts should consider combining aspects—or scaled down versions—of the National Board for Professional Teaching Standards and edTPA assessments with one or more of the other approaches discussed in this report for use in annual teacher evaluation. The National Board for Professional Teaching Standards and edTPA assessments differ from other teacher evaluation approaches in that they are content-specific and they measure and promote teachers' use of pedagogical-content knowledge in instruction. The national board standards and edTPA assessments are based on subject-specific teaching standards, and they evaluate teachers' ability to plan and engage in the types of rigorous, content-rich instructional practices called for by the Common Core Standards and assessments. Despite their potential, though, they are much more

time- and work-intensive for teachers than classroom observations, student surveys, or value-added models. In addition, while the National Board for Professional Teaching Standards and edTPA assessments provide very detailed information about single instructional units, they provide less information than other approaches about a teacher's performance over the course of a school year.

- States should review and revise licensure policies to require teaching candidates to successfully complete tests of subject matter knowledge and assessments of knowledge of teacher performance such as edTPA. Teacher-preparation organizations can address the Common Core by ensuring that coursework for teaching candidates aligns with the Common Core, and by arranging clinical experiences for teaching candidates in schools that have successfully implemented the Common Core.

The approaches to teacher evaluation described in this report, when combined with comprehensive professional development for teachers and principals, as well as changes in the organization and capacities of school districts, have the potential to support the types of teacher knowledge acquisition and changes in instructional practices called for by the Common Core Standards and assessments.

About the author

Peter Youngs is an associate professor of educational policy at Michigan State University. His research interests focus on education policy effects on teaching and learning in the core academic subjects. In particular, his work concentrates on state and district policy related to teacher induction, evaluation, and professional development in the United States and their effects on teachers' instructional practices, commitment to teaching, and retention in the teaching profession. Recent publications have appeared in *Teachers College Record*, *Elementary School Journal*, and *Teacher Assessment and Teacher Quality: A Handbook*.

Acknowledgements

The author wishes to express thanks to several people who offered very helpful ideas and assistance with the project: Glenda Partee, Jenny DeMonte, and Kate Pennington, Center for American Progress; Raegen Miller, Teach for America; Bruce Haslam, Policy Studies Associates; and Spyros Konstantopoulos, Michigan State University.

Endnotes

- 1 Andrew C. Porter, Peter Youngs, and Allan Odden, "Advances in Teacher Assessments and Their Uses." In V. Richardson, ed., *Handbook of Research on Teaching* (4th edition) (New York: Macmillan, 2001).
- 2 Robert C. Pianta, "Implementing Observation Protocols: Lessons for K–12 Education from the Field of Early Childhood Education" (Washington: Center for American Progress, 2012), available at http://www.americanprogress.org/wp-content/uploads/issues/2012/05/pdf/observation_protocols.pdf.
- 3 Robert L. Linn, "Linking Results of Direct Assessments," *Applied Measurement in Education* 6 (1) (1993): 83–102; Robert Mislevy, *Linking Educational Assessments: Concepts, Issues, Methods, and Practice* (Princeton, New Jersey: Educational Testing Service, 1992).
- 4 "About the Standards," available at <http://www.corestandards.org/about-the-standards> (last accessed January 2013).
- 5 Common Core State Standards Initiative: In the States," available at <http://www.corestandards.org/in-the-states> (last accessed January 2013).
- 6 Laura M. Desimone, "Improving Impact Studies of Teachers' Professional Development: Toward Better Conceptualizations and Measures," *Educational Researcher* 38 (3) (2009): 181–199.
- 7 Heather C. Hill, "The Nature and Predictors of Elementary Teachers' Mathematical Knowledge for Teaching," *Journal for Research in Mathematics Education* 41 (5) (2010): 513–545.
- 8 Patricia Burch and James P. Spillane, "Elementary School Leadership Strategies and Subject Matter: Reforming Mathematics and Literacy Instruction," *The Elementary School Journal* 103 (5) (2003): 519–535; Helen M. Marks and Susan M. Printy, "Principal Leadership and School Performance: An Integration of Transformational and Instructional Leadership," *Educational Administration Quarterly* 39 (3) (2003): 370–397.
- 9 Randi Weingarten, "Statement by Randi Weingarten, president, American Federation of Teachers, on Common Core Standards," Press release, March 10, 2010, available at http://www.corestandards.org/assets/k12_statements/StatementK12AFT.pdf.
- 10 Chester E. Finn, "Back to Basics: Are 'Common Standards' Good for American Education?" National Review Online, March 16, 2010, available at <http://www.nationalreview.com/articles/229317/back-basics/chester-e-finn-jr#>.
- 11 William H. Schmidt and Curtis C. McKnight, "Inequality for All: The Challenge of Unequal Opportunity in American Schools" (New York: Teachers College Press, 2012); Linda Darling-Hammond, "The Flat World and Education: How America's Commitment to Equity Will Determine Our Future" (New York: Teachers College Press, 2010).
- 12 Ibid.
- 13 David K. Cohen, "Teacher Quality: An American Educational Dilemma." In Mary Kennedy, ed., *Teacher Assessment and Teacher Quality: A Handbook* (San Francisco: Jossey-Bass, 2010).
- 14 Ibid.
- 15 Partnership for Assessment of Readiness for College and Careers, "PARCC Model Content Frameworks, Mathematics, Grades 3–11, Version 2.0" (2012); Smarter Balanced Assessment Consortium, "Content Specifications for the Summative Assessment of the Common Core State Standards for Mathematics" (2012).
- 16 Ibid.
- 17 "Item and Task Prototypes," available at <http://www.parcconline.org/samples/item-task-prototypes> (last accessed January 2013).
- 18 "Sample Items and Performance Tasks," available at <http://www.smarterbalanced.org/sample-items-and-performance-tasks> (last accessed January 2013).
- 19 Porter, Youngs, and Odden, "Advances in Teacher Assessments and Their Uses."
- 20 Tom Loveless, "The 2012 Brown Center Report on American Education: How Well Are American Students Learning?" (Washington: Brookings Institution, 2012).
- 21 William H. Schmidt and Richard Houang, "Curricular Coherence and the Common Core State Standards for Mathematics" *Educational Researcher* 41 (8) (2012): 294–308.
- 22 Daniel Weisberg and others, "The Widget Effect: Our National Failure to Acknowledge and Act On Differences in Teacher Effectiveness" (New York: The New Teacher Project, 2009); Porter, Youngs, and Odden, "Advances in Teacher Assessments and Their Uses."
- 23 Dirk C. Strauss and J. Meirion Thomas, "What Does the Medical Profession Mean by 'Standard of Care?'" *Journal of Clinical Oncology* 27 (32) (2009): e912–e913.
- 24 Peter Moffett and Gregory Moore, "The Standard of Care: Legal History and Definitions: the Bad and Good News," *Western Journal of Emergency Medicine* 12 (1) (2011): 109–112.
- 25 Porter, Youngs, and Odden, "Advances in Teacher Assessments and Their Uses."
- 26 Charlotte Danielson, "Enhancing Professional Practice: A Framework for Teaching" (Alexandria, Virginia: Association for Supervision and Development, 1996); Charlotte Danielson and Thomas McGreal, "Teacher Evaluation to Enhance Professional Practice" (Alexandria, Virginia: Association for Supervision and Development, 2000).
- 27 Thomas J. Kane and others, "Identifying Effective Classroom Practices Using Student Achievement Data" NBER Working Paper 15803 (Washington: National Bureau of Economic Research, 2010); Lauren Sartain, Sara Ray Stoelinga, and Eric R. Brown, "Rethinking Teacher Evaluation in Chicago: Lessons Learned from Classroom Observations, Principal-Teacher Conferences, and District Implementation" (Chicago: Consortium on Chicago School Research, 2011).
- 28 Pianta, "Implementing Observation Protocols: Lessons for K–12 Education from the Field of Early Childhood Education."

- 29 Robert C. Pianta and others, "Classroom Effects on Children's Achievement Trajectories in Elementary School," *American Educational Research Journal* 45 (2) (2008): 365–398; Jason T. Downer and others, "Observations of Teacher-Child Interactions in Classrooms Serving Latinos and Dual Language Learners: Applicability of the Classroom Assessment Scoring System in Diverse Settings," *Early Childhood Research Quarterly* 27 (1) (2012): 21–32.
- 30 Measures of Effective Teaching Project, "Gathering Feedback from Teaching: Combining High-Quality Observations with Student Surveys and Achievement Gains" (2012).
- 31 Ibid.
- 32 Douglas N. Harris and Daniel McCaffrey, "Value-Added: Assessing Teachers' Contributions to Student Achievement." In M. Kennedy, ed., *Teacher Assessment and Teacher Quality: A Handbook* (San Francisco: Jossey-Bass, 2010).
- 33 Brookings Brown Center Task Group on Teacher Quality, "Evaluating Teachers: The Important Role of Value-Added" (2010); Measures of Effective Teaching Project, "Gathering Feedback from Teaching."
- 34 Henry Braun, Naomi Chudowsky, and Judith Koenig, "Getting Value out of Value-Added: Report of a Workshop" (Washington: National Research Council Committee on Value-Added Methodology for Instructional Improvement, Program Evaluation, and Accountability, 2010); Eva L. Baker and others, "Problems With the Use of Student Test Scores to Evaluate Teachers" (Washington: Economic Policy Institute, 2010).
- 35 Daniel Aaronson, Lisa Barrow, and William Sander, "Teacher and Student Achievement in the Chicago Public High Schools," *Journal of Labor Economics* 25 (1) (2007): 95–135; Dale Ballou, "Value-Added Assessment: Lessons from Tennessee." In Robert W. Lissitz, ed., *Value-Added Models in Education: Theory and Practice* (Maple Grove, Minnesota: JAM Press, 2005); Cary Koedel and Julian R. Betts, "Re-Examining the Role of Teacher Quality in the Education Production Function" Working Paper No. 200703 (Nashville, Tennessee: National Center on Performance Incentives, 2007).
- 36 This is similar to the way in which a major league baseball player's batting average or earned run average may fluctuate significantly from one year to the next, but their career batting average or career earned run average is often highly correlated with their batting average or earned run average in a given year.
- 37 Jesse Rothstein, "Student Sorting and Bias in Value-Added Estimates: Selection on Observables and Unobservables," *Education Finance and Policy* 4 (4) (2009): 537–571; Jesse Rothstein, "Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement," *Quarterly Journal of Economics* 125 (1) (2010): 175–214; Derek Briggs and Ben Domingue, "A Review of the Value-Added Analysis Underlying the Effectiveness Rankings of Los Angeles Unified School District Teachers by the *Los Angeles Times*" (Boulder, Colorado: National Education Policy Center, University of Colorado, 2011).
- 38 Rothstein, "Student Sorting and Bias in Value-Added Estimates."
- 39 Linn, "Linking Results of Distinct Assessments."
- 40 Joseph A. Martineau, "Distorting Value-Added: The Use of Longitudinal, Vertically Scaled Student Achievement Data for Growth-Based, Value-Added Accountability," *Journal of Educational and Behavioral Statistics* 31 (1) (2006): 35–62; Derek Briggs and Jonathan P. Weeks, "The Sensitivity of Value-Added Modeling to the Creation of a Vertical Scale Score," *Education Finance and Policy* 4 (4) (2009): 384–414.
- 41 William H. Schmidt, Richard T. Houang, and Curtis C. McKnight, "Value-Added Research: Right Idea But Wrong Solution?" In Robert W. Lissitz, ed., *Value-Added Models in Education: Theory and Practice* (Maple Grove, Minnesota: JAM Press, 2005).
- 42 Schmidt, Houang, and McKnight, "Value-Added Research."
- 43 Anthony T. Milanowski, Herbert G. Heneman, and Steven M. Kimball, "Review of Teaching Performance Assessments for Use in Human Capital Management" (Madison, Wisconsin: Wisconsin Center for Education Research, Strategic Management of Human Capital, 2009).
- 44 Dan Goldhaber and Emily Anthony, "Can Teacher Quality Be Effectively Assessed? National Board Certification as a Signal of Effective Teaching," *Review of Economics and Statistics* 89 (1) (2007): 134–150; Leslie G. Vandervoort, Audrey Amrein-Beardsley, and David C. Berliner, "National Board Certified Teachers and Their Students' Achievement," *Education Policy Analysis Archives* 12 (46) (2004).
- 45 Porter, Youngs, and Odden, "Advances in Teacher Assessments and Their Uses."
- 46 For more details about the nature of the edTPA assessments, see: "edTPA," available at <http://edtpa.aacte.org/> (last accessed January 2013).
- 47 Linda Darling-Hammond, "Evaluating Teacher Effectiveness: How Teacher Performance Assessments Can Measure and Improve Teaching" (Washington: Center for American Progress, 2012).
- 48 Ibid.
- 49 U.S. Department of Education, "Delaware and Tennessee Win First Race to The Top Grants," Press release, March 29, 2010, available at <http://www.ed.gov/news/press-releases/delaware-and-tennessee-win-first-race-top-grants>.
- 50 Personal communication from Shelley Rouser, special assistant, Delaware Secretary of Education, January 11, 2013.
- 51 Ibid.
- 52 Personal communication from Christopher Ruskowski, chief officer of the Teacher and Leader Effectiveness Unit, Delaware Department of Education, January 11, 2013.
- 53 Ibid.
- 54 Personal communication from Felicia Smith, associate commissioner, Kentucky Department of Education Office of Next-Generation Learners, December 20, 2012.
- 55 Ibid.
- 56 Personal communication from Karen Kidwell, director of the division of program standards, Department of Education Office of Next-Generation Learners, January 4, 2013.
- 57 Ibid.
- 58 Ibid.
- 59 Personal communication from Felicia Smith.

The Center for American Progress is a nonpartisan research and educational institute dedicated to promoting a strong, just, and free America that ensures opportunity for all. We believe that Americans are bound together by a common commitment to these values and we aspire to ensure that our national policies reflect these values. We work to find progressive and pragmatic solutions to significant domestic and international problems and develop policy proposals that foster a government that is “of the people, by the people, and for the people.”

Center for American Progress

