Professional Development on Formative Assessment

Insights from Research and Practice ◣

by Elise Trumbull and Nancy Gerzon

State and district leaders, policymakers, and researchers are increasingly taking an interest in formative assessment as it becomes an important component in next-generation assessment systems, such as those called for by the federal Race to the Top legislation (Herman, 2010).

At first blush, formative assessment often seems a straightforward concept: It involves teachers' ongoing use of evidence of learning to inform instruction and to guide feedback for students. In developing this paper, we considered formative assessment to be "a process used by teachers and students during instruction that provides feedback to adjust ongoing teaching and learning to improve students' achievement of intended instructional outcomes" (FAST SCASS, 2008, p. 1). Key elements of formative assessment are explicit learning goals and success criteria, descriptive feedback to students about their performance, self and peer assessment, collaboration among teachers and students, interpretation of assessment evidence with reference to expected student learning progressions within a domain, and use of assessment information to improve instruction.

In our experience running professional development sessions, we have worked with many teachers who initially believed they knew and were familiar with both the concepts and the practices of formative assessment, but over time, as they adopted new practices of evidence collection, descriptive feedback, differentiation in instruction, and student involvement in learning, many of these teachers have reported dramatic changes in their daily work. These shifts relate to lesson planning, teachers' content knowledge, communication with students, the relationship between instruction and assessment, the processes of grading and reporting, and even how they think of their roles as teachers.

Considering both the complex task of learning formative assessment and the limited time and resources available for professional development in most schools and districts, we began a dialogue with WestEd

colleagues about our approach to this work. We asked: What do we know about how to best support teacher learning in the area of formative assessment? Is our theory of change research-based, and will it lead to the deep transformation that is necessary to implement formative assessment fully? What needs to be in place for professional learning among teachers to positively impact student learning?

This paper arose from our dialogue on these questions. It outlines how different programs and studies have addressed professional learning in formative assessment. We intend for the paper to inform readers on how to structure teacher learning for full adoption of classroom formative assessment practices. In this paper we examine the need for professional development on formative assessment, the literature

>> This paper is one in a series produced by WestEd on the topic of formative assessment.

WestEd
WestEd.org

on effective professional development in general, and examples of professional development related to formative assessment that are documented in the literature. We also discuss major issues in the design and implementation of professional development on formative assessment and make recommendations for future efforts.

Professional development, teacher quality, and student learning

Teacher quality is one of the strongest predictors of student learning (Chetty, Friedman, & Rockoff, 2011; Darling-Hammond, 2000; Rowan, Correnti, & Miller, 2002). Teacher professional development is one factor affecting teacher quality. However, documenting the relationship between professional development and student learning is challenging, "despite an intuitive and logical connection" between them (Yoon, Duncan, Lee, Scarloss, & Shapley, 2007, p. 3). Nonetheless, several high-quality, quasi-experimental studies do support the contention that professional development leads to increased teacher knowledge and skills, improved teaching, and enhanced student achievement (Yoon et al., 2007). Because research has shown that formative assessment contributes to student learning (Black & Wiliam, 1998; Wiliam, Lee, Harrison, & Black, 2004), there is reason to believe that professional development on formative assessment has the potential to enhance student achievement. Professional development works, according to an accepted theory of action, by (1) improving teachers'

knowledge, skills, attitudes, and beliefs; (2) improving instruction; and, hence, (3) improving student learning (Desimone, 2009, p. 185). At the school and district levels, the goal of professional development is to build the capacity of teachers (Mandinach & Jackson, 2011; Killion, 2012), and, in order to be effective, most innovations require "getting to scale"—that is, attaining a systemic reach (Elmore, 1996).

The scope of the need for professional development on formative assessment

Schools and districts throughout the country have been looking for professional development on formative assessment because of their belief that formative assessment is a key component of teaching that promotes students' higher-order thinking and helps students meet standards at high levels of proficiency (Schneider & Randel, 2010; Wylie, 2008). Stiggins (2010) notes that both teachers and administrators tend to have a very thin grounding in the kinds of assessment knowledge and skills they need because preservice programs offer so little "relevant assessment training" (p. 233). His observation applies to formative assessment in particular and to student assessment in general.

The need for professional development on formative assessment cannot be fully met with a few targeted workshops on formative assessment topics; teachers cannot be expected to incorporate new formative assessment practices into their teaching without the support of extensive in-service professional development (Heritage,

Kim, Vendlinski, & Herman, 2009; Shepard, 2000). Effective formative assessment in any academic domain requires teachers to integrate and apply knowledge of the academic content area, expected student learning progressions within that content area, content-area pedagogy, and how to elicit and analyze evidence to give feedback to students and adjust instruction. Wylie and Heritage (2010) assert that embracing and implementing formative assessment means huge changes for most teachers—changes in their views of themselves as teachers and in their understanding of the relationship between instruction and assessment. Effective professional development on such a complex endeavor as formative assessment needs to go deep in terms of meaningful content; and teachers need opportunities to try out, collaboratively reflect on, and revise their practices (Wylie, Lyon, & Mavronikolas, 2008; Darling-Hammond, Wei, Andree, Richardson, & Orphanos, 2009; Garet et al., 2011; Borko, 2004).

Research studies on formative assessment (see Trumbull & Lash, 2013) indicate that extensive teacher preparation and professional development are necessary if formative assessment is to realize its promise as a tool for promoting student learning. A model of continuous capacity building is needed (Mandinach & Jackson, 2011). To carry out formative assessment, teachers require support over time, since changes in instruction happen slowly and iteratively, and teachers need time to adapt and revise their instructional routines based on new learnings (Leahy & Wiliam, 2009).

Teachers also need time to reflect on their current practices, addressing issues of deepening content knowledge, pedagogy, and how to establish different methods for eliciting and using evidence during instruction (Herman, Osmundson, Ayala, Schneider, & Timms, 2006; Ruiz-Primo & Furtak, 2007).

A 2009 report on professional learning in education concludes that teachers typically need at least 50 hours of professional development related to any educational innovation in order to begin to make the innovation part of their repertoire (Darling-Hammond et al., 2009). To judge from the research reviewed, more than that may be needed to make headway toward successful use of formative assessment—depending upon the knowledge and experience base from which teachers and districts are starting.

Districts and teachers need to know that when they begin professional development on formative assessment, they will be in it "for the long haul" because "formative assessment practice requires teachers to think differently about the relationship between instruction and assessment, to see feedback as a central mechanism in promoting learning, and to come to regard students as partners in the learning process. Such fundamental shifts in practice take time, commitment, and patience on the part of both teachers and administrators" (Wylie & Heritage, 2010, p. 118). The breadth of content of professional development on formative assessment can be anticipated, but the specific design should depend to some degree on the nature of the teaching staff and

on many contextual factors within a school or district. Among these factors are what kinds of professional development have already been done on subject-area content, pedagogy, and assessment; the nature of the student population; the resources available to keep the formative assessment initiative going; and state or local policy mandates. Adapting to the scheduling demands of local districts is necessary as well; however, when such adaptations mean reducing the recommended professional development time, they may compromise outcomes (Wylie, Thompson, Lyon, & Snodgrass, 2008).

In order to be successful with formative assessment, teachers need an understanding of how students learn, a strong foundation in whatever content domains they are teaching, and knowledge of how students develop in those domains (learning trajectories), as well as familiarity with the appropriate academic standards and how to map them to learning goals and performance criteria (Heritage, 2010). Professional development on formative assessment is situated in this bigger context, and to be effective it needs to raise educators' awareness of the importance of this context—and sometimes, perhaps often, build needed expertise in all of those areas. For this reason, those conducting the professional development need considerable expertise in all the areas in which teachers need the relevant professional knowledge.

Most districts are not starting from scratch in the process of developing teachers' capacity to use formative assessment. Many

practices that have not been viewed from the perspective of formative assessment include some of its key features. Among these are process writing instruction, with its peer assessment and multiple rounds of feedback to the writer (Cowie, 1995); reciprocal teaching, in which the teacher models for the student how to question, clarify, and predict (Palincsar & Brown, 1984); and project-based learning, which emphasizes student autonomy and collaborative inquiry (Thomas, Mergendoller, & Michaelson, 1999). Teachers who have moved away from teacher-centered learning to student-centered teaching—reflecting an emphasis on learning rather than teaching—will have a leg up (Bransford, Brown, & Cocking, 2000). Teachers familiar with these kinds of practices will have already incorporated elements associated with formative assessment in their teaching, such as using formative feedback and promoting students' taking active roles in their own learning.

Even where teachers are far along in their thinking and practice related to formative assessment, district leaders need to resist the easy fix when making choices for professional development on formative assessment. Research has identified certain features of professional development as contributing to making the professional development the most effective it can be. In the following section, we review those features of effective professional development in general before discussing the research on professional development specifically focused on formative assessment. ⬊⬊

Features of Effective Professional Development and Implications for Professional Development on Formative Assessment

Effective professional development is "intensive, ongoing, and connected to practice" and, ideally, "embedded in the work of professional learning communities that support ongoing improvements in teachers' practice" (Darling-Hammond et al., 2009, p. 9). It is content-focused (Yoon et al., 2007), takes into account the local context of teaching with respect to the "affordances and constraints of the schools and districts in which [teachers] work" (Cobb, McClain, Lamberg, & Dean, 2003, p. 13), engages teachers in active learning and collective participation (Desimone, 2009; Blank, de las Aals, & Smith, 2008), and garners systemic support (Wylie, Lyon, & Goe, 2009). In addition, Schneider and Randel (2010) cite research indicating that professional development activities should also be part of a coherent program in which all parts support common educational goals. In the following sections, we further describe each of these features.

Intensive and ongoing

Rigorous research reviewed by Darling-Hammond et al. (2009) showed that professional development programs offering 30 to 100 hours (an average of 49 hours) over a period of 6 to 12 months had a "positive and significant effect" on student achievement (p. 9). Professional development offering a limited number of hours (5–14) produced no significant impact on achievement. If teachers' long-standing practices are not in line with the major tenets of formative assessment, they are likely to require lengthy and intensive professional development (Elmore, 2002), with repeated opportunities to explore the new territory.

Connected to practice

When teachers have opportunities to apply what they are learning through professional development to their own classroom instruction and reflect on what they have done (and its possible impact), the professional development is more likely to be effective (Killion, 2012; Curry & Killion, 2009; Garet et al., 2011). Connecting professional development to practice depends on adequate time for such trials and reflections.

Collaborative, embedded in a professional learning community

There is a national consensus that effective teacher professional development depends on strong, collaborative working relationships among teachers (Borko, 2004; Desimone, Porter, Garet, Yoon, & Birman, 2002). Though collaboration has been widely touted as an important element of an effective professional teaching community for some time, it continues to be hard to achieve—not least because it demands time and, quite likely, restructuring of school schedules (Darling-Hammond et al., 2009). One approach to structuring collaborative working relationships among teachers is known as professional learning communities (PLCs), sometimes called teacher learning communities (TLCs), which are groups of educators (ranging from discipline-specific and grade-level groups to the entire faculty of a school) who work collaboratively to share their expertise, promote their own professional growth, and focus on school improvement (Saunders, Goldenberg & Gallimore, 2009; McLaughlin & Talbert, 2006). PLCs may be compared to "communities of learners" (Rogoff, Matusov, & White, 1996) or "communities of practice" (Wenger, 1998) in which people have a shared interest in learning or accomplishing something and to which each person brings some knowledge or expertise. PLCs are intermediate structures that lie between the school or district level and the individual teacher level.

Content-focused

Numerous studies have suggested that professional development emphasizing subject-matter content and how students learn that content has the most impact on teacher learning and student learning, as compared to professional development on general principles of instruction or on the method of delivery (reported in Garet, Porter, Desimone, Birman, & Yoon, 2001).

Adapted to local context

"[S]ustained change in day-to-day practice is inherently local" (McLaughlin & Talbert, 2006, p. 4), and if professional development is to be effective in altering teachers' practices, it must be adapted to local conditions (Cobb et al., 2003; McLaughlin & Talbert, 1993). For

example, a plan of professional development needs to take into consideration local leadership structures. There may be content-area department coordinators, committees designed to develop curriculum or examine assessment data, or grade-level groups that have worked successfully to implement reform. A professional development effort related to formative assessment should be organized to make best use of those existing structures. However, a school or district may alter an innovation in such a way as to maintain its existing practices and avoid any change that disturbs organizational conditions (Berman & McLaughlin, 1979). For instance, in a district where instruction is driven by pacing guides, teachers may not have the latitude or time to integrate particular forms of formative assessment and make meaningful use of the data they gather from such assessments. In such a district, teachers will feel pressure to "square prescribed procedures and activities within the exigencies of their circumstances" (Cobb et al., 2003, p. 18). In that way, the integrity and meaning of an innovation are lost, and meaningful change is foreclosed (McLaughlin, 1976).

Active

Like their students, teachers need to engage actively with new concepts in order to understand and apply them effectively (Guskey, 2000; Loucks-Horsley, Love, Stiles, Mundry, & Hewson, 2003). Active learning may be promoted through discussions with colleagues, hands-on practice with particular techniques, and reviewing student work with other teachers (Schneider & Randel, 2010).

In particular, opportunities for teachers to collectively analyze their thinking about teaching and learning seem to be important to their professional development (Garet et al., 2001).

Systemically supported

School leadership is the second most important factor influencing student learning (Louis, Leithwood, Wahlstrom, & Anderson, 2010). Leadership entails "providing direction" and "exercising influence" (Louis et al., 2010, p. 9) and can be distributed across teachers, administrators, parents, and others. But it is the leadership of administrators that has been most studied and that has been shown to affect student learning. For this reason, professional development tends to be most effective when administrators understand the need for the professional development and actively participate in it (Brookhart, Moss, & Long, 2010).

Coherent

To be effective, professional development also needs to be part of a coherent program of teacher learning and district or school reform (Garet et al., 2001; Schneider & Randel, 2010). Professional development should build on broader efforts to improve teaching and learning and should be aligned with state and local standards. Each activity in a plan of formative assessment professional development should be designed to follow an expected teacher learning progression (Garet et al., 2001). Thus, professional development should be targeted to both

teacher enhancement and the district improvement process. Another key element in professional development coherence is ensuring that teachers within a system communicate with each other and share knowledge so as to support the same goals (Garet et al., 2001).

Professional Development on Formative Assessment: Additional Considerations

Effective professional development has the general features enumerated in the preceding sections. For professional development to be effective when focused on formative assessment, it may need to address additional challenges because formative assessment is built on such a broad foundation and depends on so many forms of teacher knowledge.

A professional culture for change

For most teachers, implementing formative assessment involves significant changes in practice—both in regard to the technical aspects of teaching and in their views of themselves as teachers. If these changes are to take place and take root, a school needs what Wylie and Heritage (2010) call "a professional culture for change" (p. 118). Collaborating in a PLC can promote that kind of culture. According to Loucks-Horsley, Stiles, Mundry, Love, and Hewson (2010), "Professional learning communities are associated with both changed teacher practices and changed professional culture by embedding continuous learning ⇗

into the culture" (p. 141). In a professional culture for change, teachers have opportunities to work together in an ongoing way, room to experiment, freedom to make mistakes, and the encouragement to open their practice to shared critique (DiRanna, Topps, Cerwin, & Gomez-Zwiep, 2009; Wylie & Heritage, 2010). Hargreaves and Fullan (2012) use the term "professional capital" to represent the assets that a school or district needs in order to function at the highest level to meet the needs of students and engage in continuous improvement. Professional capital is not gauged in terms of the talent of individuals within the school community but, rather, in collective terms. It includes the professional knowledge, judgment, and commitment represented by an entire school community and shared via such vehicles as PLCs.

Learning trajectories of teachers

Teachers, like all learners, follow a developmental trajectory vis-à-vis new ideas and practices related to what they already know and understand and to the nature of the new learning. Based on our experience doing extensive professional development on formative assessment with numerous groups, we believe that teachers need extended opportunities to evaluate proposed innovations by asking, "What's new, and what is already known?" We have found that teachers have a tendency to look at formative assessment practices and think, "I already do that." When given the chance to look more carefully, reflect, and discuss, they may realize that there is a lot that they don't do or don't do in

depth. And, as Wylie and Heritage (2010) note, "knowing what good practice is and doing it on a consistent basis are two different things" (p. 120). Teachers need to be recognized for their capacity to make professional judgments and adopt innovations in terms that make sense for their students and their instructional environments, even as they are in the process of learning about something new, such as formative assessment (Heritage & Heritage, 2011).

Formats and strategies parallel to those in the classroom

In considering options for professional development on formative assessment, districts would do well to choose those that model some of the broad strategies teachers will be using to carry out formative assessment in their classrooms. For example, professional developers' use of coaching, modeling, and feedback when working with teachers can have parallels with teachers' formative assessment practices in the classroom. Hence, the format of the professional development itself may exemplify important skills that teachers can use in their own formative assessment practices (Hirsh, 2011). Indeed, that is the case with several professional development efforts described in this paper; they use a PLC model, which mirrors in many ways the classroom community of learners. A PLC is a useful forum for teachers to discuss what they are learning (e.g., in workshops or readings), share what they have tried in their classrooms, elicit feedback, and make plans for revisions in practice. Experienced professional developers suggest

that a PLC include no fewer than five and probably no more than nine teachers (Wylie et al., 2009), to ensure that there will be opportunities for everyone to participate actively and get feedback.

A number of professional development researchers have used coaching and facilitating discussions as part of the professional development they have provided (see, e.g., Moss, Brookhart, & Long, 2012; Wylie et al., 2009). There are many forms of coaching, but they have in common several features: a respect for teachers as professionals, mutual dialogue between teacher(s) and coach, and an emphasis on student learning (Cornett & Knight, 2008, p. 193). Teacher inquiry is also often a part of professional development focused on formative assessment (see, e.g., Brookhart et al., 2010; DiRanna et al., 2008). Professional development providers need to be careful not to compromise genuine teacher inquiry by prescribing activities on the basis that the activities were successful elsewhere (Anderson & Herr, 2011). Teachers can use the examples of good professional developers as models for their own ways of engaging students in inquiry in the classroom.

Examples of Professional Development Related to Formative Assessment

In education research, most accounts of professional development on formative assessment are based on studies carried out by researchers in collaboration with schools or districts. In many cases,

Table 1. Professional Development Projects Reviewed and Their Features								
Project	Intensive, ongoing	Connected to practice	Collaborative, embedded in a PLC/TLC	Content-focused	Adapted to local context	Active	Systemically supported	Coherent
King's Medway-Oxfordshire	X	X	--	--	X	X	X	X
Educational Testing Service TLCs	X	X	X	--	X	X	X/NS	X
Syracuse City School District	X	X	X	X	X	X	X	X
Keeping Learning on Track	X	X	X	--	NM	X	X*	X*
Assessment for learning project	X	X	--	X	X	X	NM	X
Professional development for administrators	X	X	E	--	X	X	X	X
Rhode Island	X	X	X*	X*	X*	X	X	X
Science Assessment Leadership Academy	X	X	X	X	X	X	X	X
NBPTS certification process	X	X	NA	X	NA	NA	NA	NA
Professional development for common formative assessment—California	X	X	X	X	X	X	NM	NM
Professional development for common formative assessment—Delaware	X	X	X	MC	X	X	X	X

X = present; X* = dependent on school to implement; X/NS = present but not specific to formative assessment;
E = effectively present but not described as such; NM = not mentioned; NA = not applicable; -- = not present;
MC = Linked to multiple content areas

researchers have provided highly focused professional development sessions on specific techniques that can be used as formative assessment. For example, Ruiz-Primo and her colleagues have conducted a series of experimental research projects in which science teachers are trained to use particular formative assessment techniques. Among these techniques are concept maps (Furtak et al., 2008), curriculum-embedded written and oral prompts (Furtak & Ruiz-Primo, 2008; Herman et al., 2006), and assessment conversations (Ruiz-Primo & Furtak, 2007).

In this section, we review 11 cases of formative assessment implementation and related professional development efforts to promote changes in practice that were intended to enhance student learning. Our goal was to identify a mix of cases that examined how professional development on formative assessment affected teachers and/or students. We searched the following databases: EBSCO's Education Research Complete, Google Scholar, and the National Library of Education. To find evaluations of professional development on formative assessment, we used the search terms *professional development*, *classroom assessment*, and *formative assessment*. In an effort to cast a wide net, we also asked authors ◥◥

in the field of formative assessment to recommend sources, and we scanned key articles' reference lists for additional sources. Included in our review are summaries of reports that examined professional development for K–12 teachers or administrators, studied effects of the professional development on teachers/administrators and/or students, and described the professional development program in sufficient detail that we could identify the dimensions of effective professional development it entailed. Partly because of the paucity of studies, we did not apply any other criteria to exclude studies. Several of the cases are reviewed at somewhat greater length by Schneider and Randel (2010).

The cases we describe in this paper embody features of effective professional development to varying degrees. Table 1 characterizes each project according to its features (based on information available in published reports). It is important to note that programs designed to be used across many districts are intended to be adaptable for particular contexts and may make strong suggestions for methods to be used (e.g., PLCs), but they necessarily have less control over ensuring that such recommendations are acted on.

The King's Medway-Oxfordshire project

Formative assessment researchers collaborated with two school districts in England that had strong administrative support for formative assessment and where schools had already begun to think about implementing formative assessment practices (Wiliam et al., 2004). For a year and a half, the researchers worked with two mathematics teachers and two science teachers from each of six schools, for a total of 24 teachers. The researchers held six and a half days of in-service sessions, during which teachers were introduced to formative assessment principles and formulated their own action plans for implementing increased formative assessment in their classrooms. The first six months (January–June) were allocated to experimentation with various formative assessment techniques, and "real" implementation was set for the school year beginning in the following September. Researchers went to the schools periodically to observe teachers and meet with them to give guidance regarding the effective implementation of their plans. Comparison groups were either (a) a parallel class taught by the teacher in a previous year, (b) a parallel class taught by a different teacher the same year, or—if necessary—(c) a non-parallel class taught by the same teacher or another teacher (Wiliam et al., 2004, p. 57). Scores on assessments administered by the school were used to determine whether teachers' formative assessment practices had any effect on student learning.

Researchers did not direct the teachers to select any particular formative assessment practices to use, but they did encourage teachers to base their choices on research. The techniques that teachers elected to use fell into the broad categories of questioning, feedback, sharing criteria with learners, and self and peer assessment. Nearly all of the teachers mentioned something about "questioning" in their plans. Within that category, strategies chosen were teacher questioning (e.g., using a focal question or giving more wait time), students asking questions, and using existing pre-assessments. For feedback strategies, teachers mentioned using "comment-only marking" (providing descriptive feedback but no grades) and test review, among other strategies. Teachers gave considerable attention to strategies that would help students understand instructional objectives and criteria for grading, including setting learning targets at the beginning of a lesson, using exemplars of student work, and having students grade work examples using established criteria.

The students of teachers who used formative assessment did score higher, on average, on the school assessments than those in comparison groups. The mean effect size was 0.32 (Wiliam et al., 2004, p. 60). This statistic may be an underestimation of the effects of formative assessment because teachers' formative assessment practices changed slowly, increasing toward the end of the school year. The authors acknowledge that design constraints and some changes in the teacher group limited their ability to draw strong conclusions about the impact of the professional development on teachers' practice and, in turn, on student learning.

Educational Testing Service teacher learning communities

Wylie et al. (2009) document two cases of small teacher learning communities (TLCs) that met throughout a school year (one group monthly, the other group

six times), following several half-days of professional development. One of these groups was composed of teachers in a K–8 district and the other of mathematics teachers in a high school district. Each follow-up meeting was facilitated by an Educational Testing Service (ETS) expert in formative assessment who introduced techniques that had been deemed successful in previous efforts. Another ETS staff person took notes at each meeting, both for research purposes and a record of topics discussed by the group. At the end of the first workshop, teachers were urged to select one or two formative assessment practices to use in their classrooms. Teachers developed action plans, which they continuously modified, and kept reflective journals documenting what they were doing in their classrooms. This format parallels the format for teacher planning and documentation described by Wiliam et al. (2004).

Teachers exhibited a variety of levels of readiness, willingness, and ability to translate what they were learning into effective formative assessment practices. Some were hesitant to try new activities because they feared falling behind on prescribed curriculum. In the K–8 district, there was administrative support for expanding the use of formative assessment, and in that district more teachers beyond the initial group became involved in the effort. In the high school district, teachers were on their own, and the effort was largely contained to their initial group. Wylie et al. (2009) conclude that both district-level and school-level advocates are needed and that teachers need support for at

least two years, as they explore the concepts and implementation of formative assessment.

An interesting feature of this professional development is its focus on formative assessment strategies. Wylie et al. (2009) mention that some 80 strategies were documented and made available to teachers. Several of these strategies are described in their paper. For instance, "white boards" are plastic sleeves with a piece of white paper inserted. Students respond to teachers' questions on the white board, using an erasable marker directly on the plastic. They then hold up their boards so the teacher can scan their responses. Another strategy, "two stars and a wish," is a peer assessment. The idea is that, in response to another student's performance on a task, a peer makes two positive comments and one suggestion or wish for improvement. This strategy reportedly worked well with the students of a grade 3 teacher in the study, motivating their engagement in learning. A third strategy, "traffic light," is a way for students to indicate to the teacher their comfort with topics or concepts entailed in an assignment. Students use a red dot to show that they do not understand the concept, a yellow dot to show that they need the teacher to go more slowly or explain, and a green dot to indicate that they understand the concept. Wylie et al. (2009) suggest that, over time, students become more honest in their self-assessments. It is also likely that the students become more skilled, given time and experience with self-assessment.

Wylie et al. (2009) do not fully describe the content of the workshops provided, but the emphasis of this professional development program seems to have been primarily on formative assessment strategies. Detailed accounts of teacher meetings after the workshops do not mention the topics of assessing the cognitive demand of assessment tasks or using learning trajectories to plan instruction or interpret student performance, nor is there evidence that teachers explored the issue of how to give feedback to students that they could use to improve students' learning. While these formative assessment strategies are an important element of formative assessment practice, this case suggests that teachers are not yet using collected evidence to inform instruction or to engage students, both essential to impact student learning.

Syracuse City School District mathematics initiative

Formative assessment was integrated as a major pedagogical topic in a large, three-year professional development initiative designed to improve mathematics teaching and learning in the Syracuse City School District. By year 3 of the project, 163 teachers, administrators, and mathematics specialists for grades K–12 (82% of the originally targeted group of 200) had met the criterion of 60 or more hours of professional development. The professional development modalities included direct instruction provided by experts, turnkey training (a train-the-trainer model), mentoring of novice teachers by experienced teachers (who would coach, ⬦⬦

model, and co-teach), self-study (focused on teacher-set goals), and professional learning communities (regular meetings focused on a pedagogical topic, with feedback from peers and experts) (Newman & Gullie, 2010). The professional development mix was 11 percent direct instruction and 87 percent classroom mentoring, modeling, coaching, and other activities, including embedded formative assessment assistance (Newman & Gullie, 2010, p. 8). Prior to beginning their work with teachers, professional developers met with six top-level administrators multiple times to ensure that leadership had a common understanding of formative assessment and its alignment with district goals and other professional development efforts.

Those teachers who participated in 60 or more hours of professional development significantly increased their use of formative assessment practices, extended-response questions, systematic observations, and differentiated instruction. They were also more likely than teachers with less professional development to encourage their students to self-evaluate and engage in reflection (Newman & Gullie, 2010, p. 21). Teachers who participated in 60 or more hours of professional development showed gains in knowledge of mathematics content and mathematics pedagogy as well. The changes in teachers' pedagogy were not correlated with any changes in student performance on district or state mathematics tests, although cross-year comparisons were hampered by changes in the tests.

The evaluators found that teachers who participated in fewer than 60 hours of professional development decreased their use of extended-response questions and were more likely than those with 60 or more hours of professional development to use multiple-choice and/or true/false questions on mathematics assessments. In classroom observations, evaluators confirmed survey and interview data from teachers and noted particular feedback tools that teachers used, such as white boards, along with student oral and written reflections.

What rises to the surface in this project is the degree to which professional development was not primarily in the form of workshops but was much more in the form of sharing, mentoring, coaching, and building upon district expertise through faculty. Effective collaboration was facilitated by administrator support, which in turn was based on the administrators' foundation of understanding, engendered by early sessions with the professional development providers.

Keeping Learning on Track

Keeping Learning on Track (KLT) was developed by ETS staff under the leadership of formative assessment researcher Dylan Wiliam and is now owned by Northwest Evaluation Association (NWEA.org). KLT is a two-year program designed to train teachers in formative assessment theory and practice and to build capacity in schools and districts to scale up implementation of formative assessment. KLT implementation is intended to be "tight but loose" (Wylie, 2008). It is tight in the sense of adhering to fundamental formative assessment principles and loose in that it is necessarily adapted to local contexts. KLT emphasizes five "assessment for learning" strategies that have been shown to improve learning (Wiliam, 2007b). These are:

clarifying learning intentions and sharing criteria for success; engineering effective classroom discussion questions and learning tasks that elicit evidence of learning; providing feedback that moves learners forward; activating students as the owners of their own learning; and activating students as resources for one another. (Thompson & Wiliam, 2008, pp. 7–9 [reformatted and repunctuated from original])

KLT begins with professional development workshops, is sustained through ongoing guided learning and practice, and is supported by teacher learning communities that meet monthly.

Wylie (2008) describes implementation of KLT in four school districts (in New Jersey, Maryland, Pennsylvania, and Ohio) and one state (Vermont). Research on KLT has largely addressed its impact on teachers' and students' practices (for example, students doing self or peer assessment), and less so its impact on student learning outcomes. One large-scale study investigated the impact of KLT on student scores in reading and mathematics on statewide tests (Bell, Steinberg, Wiliam, & Wylie, 2008, discussed in Schneider & Randel, 2010). Fourteen KLT schools were compared to 73 non-KLT schools; nearly 12,000 students in grades

4 through 8 were included in this study. There were no significant differences between KLT and non-KLT schools, either overall or in post hoc analyses using a subset of schools that were matched to KLT schools. The resulting sample (a total of 28 schools) was too small for adequate statistical power, though there were "observable differences, and some significance levels at or below p = .10 in reading" (Schneider & Randel, 2010, p. 255). This study does not seem conclusive, in that changes in test scores would be unlikely to improve after a single year of intervention, given that it takes teachers time to learn and implement effective formative assessment. (However, see the outcomes of the Science Assessment Leadership Academy discussed later in this report and in DiRanna et al. [2008].)

Assessment for learning project

Brookhart et al. (2010) conducted professional development with six Title I remedial reading teachers of K–1 students in a large, rural school district, using an approach called "Teaching as Intentional Learning." The professional development consisted of both direct instruction and teachers' inquiry into their own formative assessment practices in their classrooms. Two university-based researchers and the district's Coordinator of Federal Programs provided the professional development and documented changes in teachers' thinking and practice over the course of a year. Teachers were asked to examine their beliefs about formative assessment and to experiment with formative assessment. At three points during the

year, they posted reports to an online bulletin board in response to questions about their inquiry into their assumptions about formative assessment, areas of concern, and efforts to improve their formative assessment practice. Researchers gave online feedback to the teachers. All posts were viewable by all participants. In addition, the teachers and researchers convened eight times during the year for "guided discussions" (Brookhart et al., 2010, p. 44), which were documented for later analysis.

Researchers investigated the impact of the professional development process on teacher learning as well as on student learning. Professional growth of teachers from the beginning to the end of the school year was documented. The performance of their students on standardized measures of reading (letter naming for the kindergarteners and phoneme segmentation fluency for the first-graders) was compared to the performance of other grade-matched Title I students.

Professional growth of teachers. Although these teachers were not in a formal professional learning community, they interacted as such by sharing professional concerns, inquiry, and learning. All showed similar patterns of learning. Common across teachers were:

» Increased mindfulness about formative assessment (i.e., greater awareness of and more intentional use of formative assessment);

» Greater specificity in feedback to students;

» More systematic note-taking and record-keeping about students, so as to give more useful feedback;

» Increased student involvement in using assessment information;

» A shift from a focus on achievement to a focus on motivation, as they saw students become excited about having control over their own learning;

» Increased instructional language to talk about formative assessment, linking formative assessment to differentiating instruction; and

» More creative use of a scripted reading program based on professional judgment, making adjustments to instruction on the basis of formative assessment information.

Teachers' apparent ability to use formative assessment to effectively differentiate instruction with such a program is evidence that formative assessment need not be tied to a particular approach to teaching and learning. Since teachers engaged in formative assessment must understand and respond to students' current learning needs, differentiation will necessarily follow—no two students are likely to be at the exact same point in their learning, so teacher responses must be differentiated.

Impact on student performance. For kindergarteners, researchers found no difference between the two groups of students: By the end of kindergarten, virtually all had learned letter names. For the first-graders, there was a small but significant difference in favor of ◥

the students whose teachers participated in the study. The more important student outcome, from the teachers' perspective, was the increase that teachers reported in students' motivation and sense of control over their own learning.

Limitations of the study. One limitation of this study arises from the instructional program choices made by the district. As with many other Title I programs, this district had chosen to use a highly scripted literacy curriculum that offered little instructional latitude to teachers. In keeping with the learning goals (letter naming and phoneme segmentation), assessments of student learning were focused on basic skills. Consequently, formative assessment in this case is tied to what seems to be a nonconstructivist approach to learning and an emphasis on low-level skills, with limited opportunities for student engagement in higher-level thinking or inquiry—a condition remote from what is usually associated with formative assessment. Test scores may not have been reflective of the potential of the professional development because the skills taught and assessed were routinely mastered by the majority of students in the grades studied.

Professional development for administrators

Moss et al. (2012) investigated the impact of professional development on 24 rural administrators in a Western Pennsylvania district, with regard to how they applied what they learned about formative assessment to the process of observing teachers and helping them implement

formative assessment. Prior to Moss et al.'s study, the district had developed a formative assessment model and provided professional development for all of its teachers, and administrators had led sessions for teachers. During the study, the administrators began observing teachers' implementation of formative assessment and continued to build their own professional knowledge through monthly coaching sessions led by university researchers.

The administrators used an observational protocol for classroom visits, during which they focused on whether and how teachers were establishing and sharing learning targets with students and clear criteria for judging whether targets had been met. They noted teachers' use of formative information and feedback to guide student learning. At the monthly coaching sessions, administrators learned how to sharpen their skills in classroom observation, give feedback to teachers, and identify evidence of student learning. Researchers gave feedback on the completed protocols and provided specific professional development as it seemed to be needed; they also took extensive notes during discussions at meetings.

Administrators discovered that, contrary to their expectations, not all teachers set clear learning targets. When teachers did set learning targets, they did not always express criteria for success. These findings applied, in some cases, to teachers who had been identified by administrators as "high fliers," or those most likely to succeed with formative assessment. The administrators discovered

that "the most telling evidence of the effects of formative assessment came from observing what the students were actually doing and whether the students' words and actions focused on learning" (Moss et al., 2012, p. 11). Administrators learned to be more explicit in their own feedback to teachers, so that teachers could act on it. Although they observed a lower rate of implementation of good formative assessment practices among "high fliers" than predicted, they were accurate in identifying how teachers would rank on implementation ("high fliers," "middle-of-the-road," and "struggling" teachers).

The researchers proposed three themes to capture the development of administrators during the two years of the study. First, in order to lead formative assessment efforts in their schools, administrators needed to see themselves in the role of "leading learner." Second, if they were to develop their own knowledge of formative assessment, they needed to "look for and analyze" what students were doing and learning during their observations (Moss et al., 2012, p. 14). Third, in order to know what to look for and analyze, they needed to have a deep understanding of formative assessment. If they were to be helpful to teachers (i.e., formative and timely in their own feedback), administrators had to develop the same kind of knowledge and understanding that teachers would need. The authors noted, "Until a principal or supervisor deeply understands formative assessment, classroom observations remain at the level of the principal telling the teacher what she did right instead of the

principal partnering with the teacher to learn something about student achievement" (Moss et al., 2012, p. 19).

Rhode Island online professional development project

The Rhode Island Department of Education contracted with WestEd to develop a series of online professional development modules on formative assessment, Linking Learning and Assessment in Rhode Island Schools. The second author of this paper has been the primary developer of the program. Piloted in the spring of 2012 with groups of teachers from six Rhode Island schools, the five modules are *Exploring the Foundations of Classroom Formative Assessment*, *Planning for Classroom Formative Assessment*, *Eliciting and Using Evidence of Learning*, *Engaging Students to Take Next Steps in Learning*, and *Developing and Sustaining Formative Assessment Practice*. These modules were first made available to Rhode Island teachers in the fall of 2012. Teachers can use the modules individually or in pairs or small groups. As part of the Rhode Island Race to the Top initiative, more than 60 percent of teachers in the state will participate in the online training. The program provides advance readings, interactive online content, video clips from a range of classrooms, and guided follow-up activities that can be done in collaboration with other teachers.

To support the online implementation, the program offers additional resources to be used by teacher teams at the school level, including a facilitator guide and a suggested agenda for school-site dialogues that follow each of the five modules. These resources were added to the program on the basis of participant feedback from the face-to-face piloting of the modules, as teacher feedback clearly showed that one of the most important elements for teachers' learning was the opportunity for them to talk with other teachers in their schools about what they were learning. Thus, the design team from the Rhode Island Department of Education and WestEd came up with a "blended" learning design: online instruction combined with activities within communities of practice (Rogoff et al., 1996) at schools. WestEd has written facilitation guides to help teachers make links to their own contexts, including what is going on in their schools and classrooms—whether related to, for instance, a particular textbook, an existing benchmark assessment, or a curricular unit.

During the summer of 2012, WestEd ran orientation sessions for 230 facilitators from around the state, representing teams from schools that had already signed up to use the modules. Each facilitator led a school-based team of 8–12 teachers. Part of the facilitator's job was to facilitate the community of practice sessions. The program developers emphasize that the program is not a prescription. It is a flexible resource designed so that teachers from any grade level or discipline can use it successfully. The program is not aligned with any content domain but uses examples from 30 different grade/content contexts to illustrate important principles and key strategies, such as identifying performance criteria against which to judge assessment evidence.

A distinguishing feature of the program is its focus on planning for how to *use* evidence of student learning *before* identifying the types of evidence to be gathered (e.g., answers to oral or written questions, student work). This is intended to better prepare teachers to clarify what to do with evidence for adjusting instruction. The program also takes participants through all the processes required for identifying learning goals (related to standards) and success criteria, and communicating them to students; mapping out learning progressions related to learning goals, and noting potential student misconceptions; identifying what counts as evidence of learning and how to elicit it through many different formative assessment strategies; giving feedback; adjusting instruction in light of assessment evidence; and promoting student engagement, self-assessment, and students' abilities to act as instructional resources for each other. Each process is dependent upon foundational knowledge. For example, if teachers have not been exposed to the notion of learning progressions, they will need preparatory education. Even with foundational knowledge, teachers may need considerable support to structure learning goals based on standards or use questioning strategies to elicit adequate evidence of student learning.

Readings and activities illustrate how to use five instructional routines familiar to teachers (pre-assessment, classroom discussions and academic ⮤

dialogue, questioning, analysis of student work, and classroom observations) to elicit evidence of student learning. Linking to well-known routines is intended to make the new concepts accessible and also to illustrate how those routines can be used more powerfully for the ultimate purpose of improving instruction. Teachers are urged to work together to develop proficiency and comfort with formative assessment processes and routines, compare experiences in implementation, and consequently amend their practices. Decision-making about when to gather what kinds of learning evidence from which students and how to revise instruction on the basis of that evidence is expected to be improved by such collaboration (Darling-Hammond et al., 2009).

The program addresses the need for both administrator and teacher leadership. The fifth module addresses how to develop a leadership team at the building level to move the formative assessment work forward. The program recommends that a facilitator be responsible for only one school-based team. The idea is to distribute the leadership across at least several people (Ritchie & Woods, 2007).

Science Assessment Leadership Academy

A broad-scale, in-depth professional development project was led by the Center for Assessment and Evaluation of Student Learning (CAESL), funded by the National Science Foundation, to "build science teachers' capacities to engage in formative assessment" (DiRanna et al., 2008, p. viii).

CAESL staff called their approach "assessment-centered teaching." Participants in the three-year project were California district teams, each composed of K–12 teachers and an administrator, who agreed to be part of a Science Assessment Leadership Academy for three years.

In designing its professional development, CAESL drew upon the latest theories of cognitive development (e.g., Pellegrino, Chudowsky, & Glaser, 2001); assessment, including integrating instruction and assessment (e.g., Black, Harrison, Lee, Marshall, & Wiliam, 2003; Bransford et al., 2000; National Research Council, 1996; Shavelson, Ruiz-Primo, & Wiley, 2005; Wilson, 2004); and instructional design (e.g., Bybee, 1997; Wiggins & McTighe, 2005). They situated the professional development in a standards-based, inquiry-based approach to science instruction (California Department of Education, 2000; National Research Council, 1996), guided by a conceptualization of professional development that incorporated research-based features of effective professional development (Loucks-Horsley et al., 2003).

At the outset, CAESL devoted attention to understanding district policies, histories of professional development and teacher learning, and organizational development. Four critical issues emerged: time for professional development, ensuring equity, building professional culture, and developing leadership. The Science Assessment Leadership Academy was constructed by professional developers and participants as part of building a professional

learning community. Professional developers nurtured teacher leadership through "a joint reflective process that required [teachers] to share responsibilities with [outside professional developers] as they worked toward implementation of quality assessment practices in their classrooms" (DiRanna et al., 2008, p. 161). Structured Academy activities took approximately three weeks a year, but participants spent much additional time on the project. Participants attended statewide meetings; they also met in district teams, whole-staff meetings, and grade-level groups, and as buddies. Each district team had the responsibility of disseminating assessment practices to other teachers in its district in ways that team members thought appropriate for the district contexts.

One prominent strategy used by the project was the assessment-centered portfolio, based on the one used by the National Board of Professional Teaching Standards, which "guided teachers through a process of planning, analysis, implementation, and evaluation of assessments for specific science units" (DiRanna et al., 2008, p. 168). Teachers documented their implementation efforts and used the portfolios to facilitate discussions during Academy meetings.

Evaluation data on 53 grade 3–8 project teachers show that these teachers increased their subject-matter knowledge significantly more than a control group of teachers who were comparable in grade levels taught, years of teaching experience, and student populations they taught. Likewise, their students scored significantly higher than control-group students

on standardized science tests after the intervention, beginning in year 1 of the project and continuing through year 3. Observations and analyses of lessons showed that teachers increased their use of recommended strategies for unit and lesson planning, high-level questions, and monitoring of student misconceptions in order to adjust instruction (Young, 2008–2011).

NBPTS certification process as professional development

The National Board for Professional Teaching Standards (NBPTS) certification process is not a recognized form of professional development but has been shown to have a strong positive effect on teachers' formative assessment skills (Sato, Wei, & Darling-Hammond, 2008). The NBPTS operates a voluntary system for certifying teachers who meet rigorous teaching standards. Sato et al. (2008) found that the three-year NBPTS process, using the NBPTS standards and assessment tasks, apparently served as professional development in helping nine science and mathematics teachers who were NBPTS candidates improve their formative assessment skills substantially. These teachers were matched for comparison to seven teachers who were non-NBPTS candidates (but who had considered being NBPTS candidates). The NBPTS and comparison teachers had similar average levels of experience (9.9 and 11.7 years, respectively); both groups had a preponderance of science teachers (eight for the NBPTS candidates and five for the comparison group). The NBPTS group came from three middle schools and six high schools, whereas the

comparison group came from four middle schools and three high schools. The NBPTS candidates' schools served somewhat academically needier populations (as measured by state Academic Performance Index) and had larger class sizes (averaging 28.9 students per classroom, compared to 26.5 students per classroom in the comparison group's schools). There were fewer credentialed teachers in the NBPTS candidates' schools.

The NBPTS candidates and the comparison teachers followed by the researchers were assessed according to six dimensions of formative assessment: views and uses of assessment; range, quality, and coherence of assessment methods; clarity and appropriateness of goals and expectations for learning; opportunities for self-assessment; modifications to teaching based on assessment information; and quality and appropriateness of feedback to students (Sato et al., 2008, p. 673). The researchers also collected videotapes of classroom lessons, written responses to questions about those lessons, student work samples, and teacher lesson plans. Researchers also conducted teacher interviews, student and teacher surveys, and final reflective interviews with teachers.

At the outset, the NBPTS candidates scored lower than the comparison group on four out of six dimensions, but their scores increased by the second year and continued to surpass those of the comparison group into the third year. "Pronounced changes were in the variety of assessments used and the way assessment information was used to support student learning" (Sato et al., 2008, p. 669).

Teachers' understanding of their own practices increased in the process of being asked what they understood, how they came to understand it, and what they still needed to learn. Although teachers were asked some questions pertaining to assessment (e.g., "What counts as an assessment?"), many questions were not about assessment. However, the opportunity to think about relationships among standards, goals, instruction, and assessment was rich ground for professional development in general and affected the teachers' understanding of the appropriate and effective uses of varied types of classroom assessment practices—including formative assessment. The comparison-group teachers whose scores did improve described professional development experiences that were similar to the NBPTS process.

The study authors concluded that the NBPTS process is useful not only for identifying effective teachers but also for developing improved teaching, based on the professional learning that takes place (Sato et al., 2008, p. 670). A previous study conducted with science teachers found that assessment was an area in which teachers evidenced some of the most significant learning through the NBPTS process (Lustik & Sykes, 2006, reported in Sato et al., 2008).

Professional development for common formative assessment—California urban elementary school

While formative assessment entails ongoing and daily evidence collection to guide instruction, the use of common formative ◣◣

assessments typically involves teachers' co-developing weekly assessments and then discussing student results in collaborative teacher teams. In this way, the definition of formative assessment used elsewhere in this paper does not necessarily apply to the case described in this section and the following case. However, these cases are included because the professional development is similar to that required for formative assessment. In these cases, teachers are described as learning how to collect evidence, quickly review data, and explore different ways of supporting students' next steps in learning.

Frey and Fisher (2009) document a case in which the joint development of common formative assessments served as a powerful professional development opportunity for teachers in a large, urban K–5 school in California. A group of classroom teachers and a reading specialist, without benefit of formal professional development on formative assessment, worked together to develop a system for "writing, scoring, evaluating, and using common assessments to inform their instruction" in the area of literacy (Frey & Fisher, 2009, p. 675). The system entailed three phases. First, teachers met to do backward curriculum planning, through which student outcomes were identified first (as they are in the Rhode Island program), then curriculum and instruction were mapped to those outcomes (Wiggins & McTighe, 2005). State standards and a curriculum pacing guide were used as reference points. Second, teachers met to design prompts for a common assessment (typically

10–12 questions, plus a writing task). Third, teachers met to discuss student performance and hypotheses for students' errors and to plan future instruction. They also shared instructional strategies and materials during the meetings. The study's authors interviewed teachers, documented teacher meetings and all professional development activities, and observed in teachers' classrooms.

Over the course of four years, through observations and interviews, the researchers found the following: Teachers became more familiar with standards, increased their knowledge of grade-level content and how it related to standards, improved their ability to develop assessments, and found ways to link their instruction to assessment and to consider the instructional implications of the assessment data they gathered. They also found the process helpful for identifying students who needed particular interventions.

At the end of the four-year period, student scores on state standardized reading tests had significantly improved at all tested grade levels (2–5). Grade 5 scores were most improved, at 26 percent higher, with grade 2 up 23 percent, grade 3 up 6 percent, and grade 4 up 21 percent. The school's Academic Performance Index, an accountability score calculated by the state on the basis of students' test performance, had increased from 573 out of a possible 1,000 points in 2001 to 746 in 2005. All subgroups of students met the state's targets for annual improvement, including English learners, students living in poverty, Latino students, African American students,

and Asian American students. Teachers regularly commented that students' improvement in reading allowed them to teach at a higher level and devote more time to science and other subjects. This self-professional development must have spread beyond the initial group of teachers, but Frey and Fisher (2009) do not make clear by what mechanisms that happened.

Another important condition is that the teachers in this study had access to peer coaching and other professional development that likely supported their ability to capitalize on yet another way to deepen their knowledge and hone their skills. The case also illustrates potential problems when teachers who are not trained in item design create formal assessment items. For example, one question on a common formative assessment read: "Which one is not true for the index of a book?" (Frey & Fisher, 2009). A fundamental guideline of item development is to avoid framing questions in the negative (Haladyna, Downing, & Rodriguez, 2002). Such questions increase the need for accurate reading, and they are linguistically more complex than questions framed in the positive (e.g., "Which item below is found in the index of a book?"). A teacher expressed frustration that so many of her students got the item wrong. She and her colleagues decided that students had overlooked the word "not," but they apparently did not conclude that negatives should be avoided in future assessment questions.

Professional development for common formative assessment—Delaware special educators

Palucci (2010) examined the success of a Delaware PLC constituted for the purpose of increasing K–8 special-education teachers' use of data on students' mathematics and science achievement to make instructional decisions. Main concerns of the group were improving special-education students' performance on state tests and meeting targets for annual improvement. The PLC was composed of eight special-education teachers and a principal; the researcher (Palucci) worked directly with the PLC.

The group met one to two times a month throughout the school year. It is not clear whether professional development in the form of presentations or activities provided by an expert was part of the PLC meetings, though Palucci (2005) mentions that instruction was provided by "various sources," including herself, at times (p. 52). Extensive documentation of teachers' thinking and practice was gathered through discussion, surveys, observations, and teacher reflective journals.

Developing common formative assessments, setting student goals, and sharing lesson plans and instructional strategies were tasks of the PLC. In addition, teachers learned and reviewed their use of three classroom formative assessment techniques: asking higher-level questions, following up with probe questions, and helping students correct mistakes. Teachers consciously linked assessment to specific academic standards and, on observation, all were seen to be

using standards-guided instruction. When the evidence from common formative assessments showed that a learning goal had not been met, teachers reviewed lesson plans, discussed how to improve them, and provided students additional instruction to meet learning goals. When students were not successful with additional instruction, teachers focused on identifying additional information teachers might need in order to understand the specific student-learning problem as well as instructional approaches that would best support the student. On occasion, they identified outside readings that they thought would help with this process.

According to several measures, all teachers markedly increased their use of data to plan and revise instruction. They rated themselves higher on ability to use data, standards, and grade-level expectations to plan instruction. Teachers asserted a need for more time for collaboration but believed that they did not need more professional development on writing lesson objectives and assessments. Students in grades 2–6 and 8 improved significantly on standardized mathematics tests. Students in grades 2–5 improved significantly on standardized reading tests (Palucci, 2010).

Recommendations to Guide Effective Professional Development for Formative Assessment

In this section, we offer recommendations for professional development in formative assessment,

based on the analyses summarized in the preceding sections. These recommendations take into account such topics as developing leadership, attending to the long-term change process of formative assessment, and addressing local context in professional learning design. While some of these recommendations are similar to those recommended for effective professional development in general (see, for example, Learning Forward, 2011), this section discusses these recommendations, as much as possible, in the context of formative assessment implementation. There is, of course, much to be learned from future research; new findings may alter these recommendations and add new ones to the list.

Build on what teachers know

Professional development needs to be respectful of teachers' existing practices. It should be grounded in the understanding that many of the elements of formative assessment are not new to teachers. The components of formative assessment include many of the elements that reformers have been promoting for a long time—e.g., engagement of students in inquiry and higher-level thinking, promoting students' self-regulated learning and self-assessment, and gathering data to inform instruction so as to adjust it to learners' needs. Part of building on what teachers know is recognizing that they have important professional knowledge and experience. Teacher professionalism should be supported as teachers are encouraged to use their judgment and move away from test preparation and reliance on summative assessment (Black & Wiliam, 2003).

Do not underestimate the magnitude of change required

Formative assessment is likely to come up against long-standing practices that run counter to formative assessment assumptions and underpinnings—e.g., grading, approaches to student learning, beliefs about student abilities and where those abilities come from, and the idea that the teacher is the owner of knowledge. Formative assessment is not a *program*, and full implementation is thus quite different from other types of adoptions. As we have discussed, for many teachers, deep changes in knowledge and beliefs may need to take place on several fronts (assessment theory, cognitive development, content pedagogy, teacher-student roles, and more). In some cases, the whole enterprise of teaching may be tossed into the air for reorganization, including not only the components of instruction and assessment but also classroom organization, grading practices, approaches to student motivation, and relationships to other adults involved in students' education.

In particular, effective formative assessment practice involves a shift in students' role in the learning process (Heritage, 2010)—a shift away from what prevails in many classrooms. The definition of formative assessment from the Formative Assessment for Students and Teachers State Collaborative on Assessment and Student Standards (FAST SCASS, 2008) outlines "collaboration" as one of five attributes of formative assessment. This does not mean collaboration between teachers, but collaboration within the classroom, both between teachers and students and *among* students. Developing a culture of collaboration in the classroom requires far more than additional technical knowledge related to eliciting evidence of learning and using that evidence to inform instruction. It entails teachers and students working together actively to clarify learning goals, to make sense of emerging understandings, and to differentiate next steps in learning, based on evidence (Wiliam, 2007a). Teachers must be supported as they move through different stages of this change process (Curry & Killion, 2009; Reeves, 2010; Learning Forward, 2011).

Provide many opportunities for teachers to try out and discuss formative assessment

One strategy that seems to boost teachers' willingness to engage in formative assessment processes is to introduce tools or artifacts that move students toward thinking more deeply about their own learning or that help teachers gather evidence of learning (R. P. Durán, personal communication, December 7, 2011). One such tool is the "exit ticket" (Ishii, 2003; Wylie et al., 2009). This takes the form of a small index card given to a student at the end of an activity and used by the student to write two or three sentences reflecting about what he or she has just learned. This is the student's ticket to leave the class or go on to the next activity. Later, the teacher can use it as assessment information to help in planning future instruction.

Another formative assessment tool is a simple clipboard with a matrix of students' names and spaces for the teacher to write observations about students. Rubrics that help students assess their own level of proficiency with regard to a particular skill are an additional tool (see, e.g., Andrade, Du, & Wang, 2008). The "white boards" described earlier (in lieu of the electronic technology that allows students to respond to a question instantly) are also popular in some settings (Wylie et al., 2009). Using tools such as each of these "requires a relatively small change to teacher practice but may result in large changes in teacher pedagogy, the classroom culture, and student learning" (Wylie et al., 2009, p. 2).

As previously mentioned, teachers need opportunities to share their strategies with other teachers, in order to examine what works and what does not seem to work. It may take such a forum for a teacher to be able to analyze just how well he or she used assessment information to reformulate instruction. Practice opportunities should focus not only on particular techniques, such as the ones mentioned in this paper, but also on the broader processes of assessment and instructional planning.

Select packaged programs and outside providers critically

Schools and districts that do not have professional development specialists who are knowledgeable about formative assessment will need to identify outside resources that can build internal capacity. For example, they may find a program such as Keeping Learning on Track (KLT) attractive. However, initial reports of KLT's impact on student achievement are not

particularly encouraging, though many users seem to be committed to the program (Wylie, 2008).

Anderson and Herr (2011) argue that externally developed programs violate "principles of authentic inquiry" that underlie professional learning communities (p. 287). Such programs risk reproducing an "outside-in, top-down, teacher-deficit model" that does not work (p. 287). A reliance on so-called "evidence-based practices" can foreclose inquiry (p. 288) and places the emphasis on fidelity of implementation rather than adaptation to local context. Anderson and Herr (2011) also suggest that "even the notion, inherent to PLCs and action research, that data can be used to make better decisions through inquiry, has been appropriated by a testing and data management industry that has commodified and fetishized data" (p. 288 [citing Burch [2009]]). They caution that data should be used to support or open up inquiry rather than to narrow the decision-making process prematurely, and that any ready-made products should be viewed with a critical eye. According to Anderson and Herr (2011), if teachers' professional judgment is the basis of formative assessment, processes that circumvent or diminish development and use of that professional judgment are detrimental.

Select participants and outside collaborators judiciously

Self-selection of teachers or nomination of teachers thought by administrators to be most likely to succeed increases the likelihood of successful implementation of formative assessment but reduces the ability to draw generalizable inferences. However, randomized assignment of teachers to experimental or control groups within a school or district for a research study forecloses the opportunity to use existing professional learning communities and runs counter to what is known about how to provide the most effective professional development. If the entire professional staff cannot participate in a professional development program, a school or district will be faced with determining how to scale up eventually to include all staff, if an innovation is to succeed. That issue should be addressed from the outset of any professional development effort, and resources and incentives should be identified to ensure complete participation and buy-in (Elmore, 1996).

When considering a collaboration with outside researchers, district leaders need to ask questions such as the following: If one of the features of a PLC is "shared values" (DiRanna et al., 2009), are PLCs that are externally created for the purpose of a research project the same as internally created ones or naturally arising ones? Are there professional development research designs that capitalize on or even enhance the strengths of a teacher professional community? Are there mechanisms within the school district that can support continued teacher learning on formative assessment? Although much can be gained from collaboration with outside researchers, these questions should be weighed in advance of any agreement.

Collaborating with researchers may offer the chance for teachers to develop skills aligned with district goals for formative assessment use. For instance, recent research showed that short-term professional development workshops focused on identifying student errors and misconceptions resulted in enhanced assessment skills in high school biology teachers, compared to a control group who attended workshops focused on content-area knowledge (Buschang, 2012).

Balance adherence to a professional development design with adaptation to local context

Professional development programs on formative assessment that are developed by outside providers need to be adapted to meet local conditions, but if the basic design of a program is heavily altered, there is no way to know if it might have been more effective in its original form. Thompson and Wiliam (2008) address this tension between fidelity and adaption in relation to the KLT program. They say that any formative assessment professional development and associated implementation must be "tight but loose" (p. 1). It needs to be tight in adhering to fundamental principles but loose when alterations in plans support teacher participation and implementation of a desired change without compromising those principles.

Districts, schools, and teachers are likely to try to adapt external professional development programs to fit their own contexts. This is necessary and important to the success of any program ⬚

(McLaughlin & Talbert, 2006). However, it can be challenging to determine what kinds and degrees of adaptation are reasonable. When district personnel try to modify a project to suit their local conditions, if outside researchers or professional developers are in charge, the outsiders may try to stand firm to maintain what they see as fundamental design features or content elements. But in the long run, the district or school is likely to prevail. Wylie, Thompson, et al. (2008) discuss a problem with the TLCs in a 10-school KLT project in the Cleveland Municipal School District:

> Some TLCs had as many as 18 members (far more than was advised, but they did not want to break into smaller groups, this being their "only chance to get to know one another!") and some had as few as three (fewer than was advised, and these groups included only those teachers who were originally trained as teacher learning community leaders). (p. 82)

It became apparent that some schools were recruiting teachers who had not participated in the training for their TLCs, and in some schools the TLCs had more teachers who had *not* attended training than teachers who had attended. That meant that many TLCs did not begin as communities of teachers who knew each other and that there was a considerable number of newly participating teachers who were not prepared for the next steps in the KLT process. If a district wants to contract with a professional development provider, it needs to get assurance that the program will be customized adequately to meet district needs and objectives (Mandinach & Jackson, 2011).

One consideration for districts is whether to link professional development on formative assessment to a particular academic domain. Because of the connection between domain knowledge and productive use of formative assessment, experts have recommended that professional development on formative assessment practice be nested within subject-matter professional development (Shepard, 2005). Formative assessment researchers Black and Wiliam (2003) say that a subject-specific focus offers teachers "concrete ideas about the directions in which they can productively take their practice" (p. 236). Collaboration among teachers and between teachers and researchers can also build key domain knowledge as well as pedagogical content knowledge and assessment skill linked to modifying instruction (Darling-Hammond et al., 2009; Shepard, 2005). However, in some cases, it may be more practical to provide opportunities to connect professional development on formative assessment to multiple domains, so as to make it meaningful for a mixed population of teachers. In small rural middle and high school districts, where teachers are often responsible for more than one subject, a focus on a single academic domain may be unrealistic and unproductive.

Furthermore, domain knowledge is useful only to the degree that teachers understand cognitive development in general, as well as how it may take place within a particular domain (Heritage, 2010). Research on professional development has shown that a focus on content and *how students learn that content* is associated with better student achievement outcomes (Doppelt et al., 2009). Hence, content-focused professional development should also take a developmental view of learning within a domain. This conclusion is consonant with recent thinking on how to use the notion of learning trajectories to link teaching and learning conceptually and in research (Sztajn, Confrey, Wilson, & Edgington, 2012).

Ensure enough time and availability of teachers

Most of the professional development programs described in this paper have offered dozens of hours or more of education and support and have taken place over a period of at least a year. These programs contrast with the kind of professional development or training that is generally provided in the course of many research projects on formative assessment, which is not ongoing and does not allow teachers multiple opportunities to try a practice, get feedback, and retry it (see Trumbull & Lash, 2013). On the whole, these research studies showed that even experienced teachers with deep knowledge of subject-matter domains had problems completing some key elements of formative assessment: planning, collecting evidence, using evidence to inform instruction, and involving students in the learning process. The biggest challenge usually lay with the step of revising instruction on the basis of formative assessment information. That finding may be largely attributable to the limited

nature of the professional development provided.

Teachers who have not already been consciously using formative assessment to improve instruction and learning will most certainly need to try to apply what they may be learning in workshops or study groups. Without such connection to practice, teachers are not likely to anticipate how to successfully carry out all the elements of formative assessment.

Simply getting enough time allocated to professional development in a school or district is one of the biggest challenges to successful professional development (Lyon, Cleland, & Gannon, 2008; Heritage & Popham, 2008). The often relentless chipping away at time allotted for workshops, meetings, coaching, or other components of a professional development program is a major threat to effective implementation. If a school or district does not protect teachers' time, the fidelity of implementation of any professional development program will be compromised.

Well-designed professional development can be foiled by attendance issues, discontinuities in membership of PLCs or TLCs, shortages of substitutes, failure of principals to understand the value of the meetings (because they missed the initial professional development session for administrators) and consequent double bookings of teachers, and administrator turnover (Wylie, Thompson, et al., 2008). In one instance, described by Wylie, Thompson, et al. (2008), teachers were required to vacate school buildings 90 minutes after the end of the school day; as a result, the two-hour minimum meeting requirement for professional development could not be met.

Recognize that PLCs can be powerful but will not address all needs

A PLC is a frequently selected professional development structure because it is situated in a school and its members are able to act on the basis of an understanding of the school culture and context. Thus, the PLC is more likely to have the capacity to promote teacher change that alters entrenched practice (Thompson & Wiliam, 2008, p. 14). The PLC can also continue in a sustained manner over long periods of time, something that outside professional development providers often cannot do. In addition, PLCs can carry out inquiry and action research focused on the particular needs of a school. Subgroups within a school may function well for ongoing professional development related to formative assessment without external support. Grade-level collaborative teams, for instance, have shown promise for improving classroom learning (Saunders, Goldenberg, & Gallimore, 2009). Even when there is no single person in the group who is well grounded in formative assessment, the members are likely to bring varying levels of expertise in areas of education related to formative assessment and can use outside resources to build their formative assessment knowledge (Wylie et al., 2009).

Teachers' own conscious inquiry within the context of their schools and classrooms will be an important part of their developing knowledge and understanding related to formative assessment (Putnam & Borko, 2000). As Borko (2004) observes, teacher learning is "situative," taking place in multiple contexts (e.g., at a workshop, in a school hallway, in one's own classroom, within a professional learning community), each of which has its particular characteristics that interact with or affect one's learning.

But as powerful as they can be, PLCs are not the only important element of a professional development program on formative assessment. Formative assessment needs to be addressed as part of a system in which standards guide the development of learning goals, which are aligned with immediate performance objectives (Stiggins, 2005). Clear criteria for what counts as successful student learning and performance are also important; hence, many professional development programs incorporate rubric development in order to help teachers think about student outcomes in a nuanced way (Wylie et al., 2009). Teachers may also think ahead to how to engage students in evaluating their own learning using similar tools (Andrade et al., 2008).

A PLC operating without strong administration support and understanding is likely to be powerless to implement sustainable change. Other constituencies besides teachers must be persuaded that formative assessment is desirable. Some formative assessment practices, such as peer assessment, may not sit well with parents who are used to teachers' having ⮞

total control over assessment and judgments of student learning (Carless, 2005). Administrators may need to pave the way with parents before any major change effort begins. Some professional development would be advisable for parents, administrators, and any other stakeholders who need to be on board with an innovation.

Moderate expectations for speedy impact

Because effective professional development takes many hours and is conducted over an extended period of time, formative assessment professional development can be expected to take time to implement and more time to pay off. It is not reasonable to expect big changes in student learning at the end of a single year of professional development. Teachers may not be fully implementing or effectively implementing desired practices until toward the end of the year, a change that may coincide with the end of research documentation, unfortunately, as in the study by Wiliam et al. (2004).

Plan for administrator learning prior to implementation with teachers

The support of administrators is meaningful only to the degree that their support is coupled with their elaborated and sophisticated understanding of formative assessment. Administrators need to have a full grasp of the scope of a formative assessment reform and what can be expected to come out of professional development that is effective. On a practical level, planning for long-term professional development and related districtwide

support (e.g., staffing, release time for PLC meetings) should be done at the outset of any effort.

Administrators themselves need professional development on formative assessment if they are to give proper support to teachers who are trying to strengthen their formative assessment practices (Moss et al., 2012). Administrative support—and, perhaps more important, administrators' deep knowledge of formative assessment—is a central factor influencing the success of formative assessment professional development with teachers (Moss et al., 2012; Noyce & Hickey, 2011). Professional development for administrators should focus not just on content but also on how to support teachers to improve their practices so that the administrators can "lead a school culture that is focused on learning rather than evaluation" (Moss et al., 2012, p. 3).

Nothing, not even a lengthy memorandum of understanding signed by all responsible parties, can ensure that administrators actually participate in any reform effort—in this case, professional development on formative assessment. For example, in the initiation phase of the KLT professional development in Cleveland, all district administrators were expected to attend a one-day kickoff workshop. Every single principal failed to attend the workshop after the district scheduled a districtwide head count that required them to stay at their schools on that day (Wylie, Thompson, et al., 2008, pp. 80–81).

Reflecting on several district implementations of formative assessment, we have noticed that

mandating formative assessment as a required practice had some unintended negative consequences, in particular when there was not yet shared understanding throughout the district of what formative assessment entails. In some cases, administrators construed formative assessment as a set of strategies that could easily be added to teachers' daily instructional practices. The thoughtful process of determining what kind of evidence about learning should be gathered and how that evidence should then be used was undercut by a focus on the technical aspects of strategy use. When administrators evaluated teachers based on only one aspect of formative assessment use, without a full understanding of the complexity and range of practice, it appeared to undercut full implementation.

Areas for Future Work

Addressing special needs

Professional development in general does not often address the needs of students with learning differences or those who are still learning English (Darling-Hammond et al., 2009; Darling-Hammond & Wood, 2008). In the literature we have surveyed, no accounts of preparing teachers to conduct formative assessment with different populations have emerged, although there are articles on how to adapt and conduct formative assessment with special populations (e.g., Abedi, 2010; Elliott, Kettler, Beddow, & Kurz, 2010). A study by Amaral, Garrison, and Klentschy (2002) in a large district in the Imperial Valley of California (an area with a high

population of English language learners) documented teachers' learning about and use of kit- and inquiry-based science instruction over a period of four years. Professional development for the teachers reportedly included attention to formative assessment, whose purpose was described as "to gather information for program improvement" (p. 223). No further details were offered.

Carless (2005) describes educational reform efforts in Hong Kong (a bilingual educational setting), including a shift toward assessment for learning and away from a great emphasis on norm-referenced tests. He refers to a model of professional growth (Clarke & Hollingsworth, 2002) but does not describe how professional development on formative assessment was carried out.

The previously cited study by Palucci (2010) showed that formative assessment implemented by special-education teachers can have an impact on student test performance. Special-education teachers are likely to be prime candidates for effectively using formative assessment. In fact, long-standing special-education practices of curriculum-based assessment and dynamic assessment are comparable in many ways to formative assessment (Fuchs, 2004; Fuchs et al., 2007).

Measuring the outcome of professional development

Should the efficacy of professional development be judged solely on its eventual impact on student performance? If so, what measures of student learning are most appropriate,

and under what circumstances? How long should one expect to wait to see the impact of professional development on student learning? There are no simple answers to these questions. To begin with the first question, summative district or state tests may not be tied closely enough to the taught curriculum to serve as measures of the success of professional development. In addition, these tests rarely address the kinds of 21st-century skills that formative assessment can promote, including self-assessment, self-regulation, collaboration, and how to give and receive constructive feedback. Locally developed assessments are not quality controlled, and asking teachers to develop new assessments has equal drawbacks (Wiliam et al., 2004). This issue will need to be addressed by schools and districts as they attempt to gauge the effectiveness of professional development over a number of years—a necessary step in the process of determining what further support teachers need. With regard to the expected timing of impact, it seems reasonable to conclude that no judgment should be made in less than a year's time and that it may take two years or more to see results. Perhaps a comprehensive program such as the Science Assessment Leadership Academy could be expected to produce faster results, but external factors could certainly dampen the outcomes of even the most well-planned and well-executed program.

Ensuring that changes "take" and endure

As suggested previously in this paper, conditions for maintaining and continuing to develop effective formative assessment practices

must go beyond the level of individual teacher change. It cannot be assumed that even the best ideas for school improvement will become institutionalized because of their merit. Organizations are notorious for maintaining the status quo, so additional elements must be present as well (Elmore, 1996). Elmore argues that most meaningful educational reforms are directed at what he calls "the core of educational practice" (p. 2), or "how teachers understand the nature of knowledge and the student's role in learning, and how these ideas about knowledge and learning are manifested in teaching and classwork" (p. 2). This core of educational practice extends to the ways that schools are organized physically and socially: e.g., classroom layout, student grouping, teacher relationships, and assessment and grading practices (Cobb et al., 2003).

The core is not readily changed, as evidenced by the decades of thought, research, and effort devoted to moving from a teacher-centered approach to a student-centered approach to instruction—without widespread impact (Cuban, 1999; Cuban, Kirkpatrick, & Peck, 2001). Elmore (1996) suggests that districts will have to provide incentives to teachers if any major innovation is to take root and endure: "Encouragement and support, access to special knowledge, time to focus on the requirements of the new task, time to observe others doing it—all suggest ways in which the environment of incentives in the organization comes to reflect the requirements of learning" (p. 25). To this we would add that another important element ⬎

of the environment could be revisions to the teacher evaluation process to reflect the shift in values associated with the innovation.

Conclusion

Formative assessment, once in place, has been shown to increase student learning, but the mechanisms to fully implement formative assessment, both at the level of individual classrooms and system-wide, are not yet *fully* understood. This paper highlights the need for ongoing, site-based professional learning that invests in teachers so that they can develop the skills required for formative assessment implementation. More importantly, this paper presents evidence that supports providing professional learning in ways that will help teachers over time to examine their beliefs about teaching and learning, the role of teacher and student, and how teachers support and encourage student involvement in the learning process.

While the mechanisms necessary for full implementation of formative assessment are not yet altogether known, there is consistency across the cases in this paper, which can inform school and district leaders, professional developers, and researchers. The cases document an emerging consensus about the necessity of ensuring that school leaders understand and support this work over time, about the value of teacher dialogue to build capacity to use evidence of learning during instruction, and about the importance of attending to teachers' content knowledge as they develop

formative assessment practices. In addition, these cases suggest that professional learning in formative assessment benefits from modeling the kinds of changes one hopes to see in a formative assessment classroom—professional development that is grounded in the daily work of teaching and learning, takes place within teachers' classrooms as much as possible, and provides multiple opportunities to explore, practice, and reflect on new content.

The cases in this paper also serve as a reminder that the field has more to learn to continue to clarify the varied mechanisms that will result in systemwide implementation of formative assessment. Professional development providers working in the realm of formative assessment should be urged to find ways to document their activities, the impact on teachers, and eventual impact on students. The literature will continue to be impoverished if such documentation does not make its way into publication. Experimental studies conducted by university-based or education agency–based personnel can contribute some insights into how teachers learn aspects of formative assessment, but such studies do not tend to illuminate relationships among all the parts and across a whole school or district. It would be particularly useful to the field to have more explicit accounts of how professional development programs on formative assessment have been specifically adapted (successfully or unsuccessfully) to local contexts.

It may be daunting to districts to try to build a professional development program that pulls

together all the related pieces of an effective approach to formative assessment. Not only are significant resources likely to be required, but the local conditions that readily combine to prevent easy implementation of professional development and attenuate its effects are seemingly innumerable. The programs of professional learning described in this paper illustrate how a complex array of features and conditions must come together if teachers and districts are to benefit fully from professional development on formative assessment.

Above all, education leaders must strive to ensure that formative assessment is not reduced to a set of technical practices abstracted from the instructional context. Administrators need to anticipate and fend off pressures to homogenize formative assessment, as they seek to promote its use throughout a district. The very power of formative assessment lies in its contextualization, its being tailored to specific students' learning within specific domains at particular points in development.

References

Abedi, J. (2010). Research and recommendations for formative assessment with English language learners. In H. Andrade & G. Cizek (Eds.), *Handbook of formative assessment* (pp. 181–197). New York: Routledge.

Amaral, O. M., Garrison, S., & Klentschy, M. (2002). Helping English learners increase achievement through inquiry-based science instruction. *Bilingual Research Journal*, *26*(2), 213–239.

Anderson, G. L., & Herr, K. (2011). Scaling up "evidence-based" practices

for teachers is a profitable but discredited paradigm. *Educational Researcher, 40*(6), 287–289.

Andrade, H. L., Du, Y., & Wang, X. (2008). Putting rubrics to the test: The effect of a model, criteria generation, and rubric-referenced self-assessment on elementary school students' writing. *Educational Measurement: Issues and Practice, 27*(2), 3–13.

Berman, P., & McLaughlin, M. (1979). *An exploratory study of school district adaptation*. R-2010-NIE. Santa Monica, CA: Rand.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Berkshire, England: Open University Press.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy, & Practice, 5*(1), 7–74.

Black, P., & Wiliam, D. (2003). In praise of educational research: Formative assessment. *British Educational Research Journal, 5*, 623–637.

Blank, R., de las Aals, N., & Smith, C. (2008). *Does teacher professional development have effects on teaching and learning? Analysis of evaluation findings from programs for mathematics and science teachers in 14 states*. Washington, DC: Council of Chief State School Officers.

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher, 33*(8), 3–15.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (2000). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academies Press.

Brookhart, S. M., Moss, C. M., & Long, B. A. (2010). Teacher inquiry into formative assessment practices in remedial reading classrooms. *Assessment in Education: Principles, Policy, & Practice, 17*(1), 41–58.

Buschang, R. E. (2012). *The impact of short-term science teacher professional development on the evaluation of student understanding and errors related to natural selection*. CRESST Report 822. Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing.

Bybee, R. W. (1997). *Achieving scientific literacy: From purposes to practices*. Portsmouth, NH: Heinemann.

California Department of Education. (2000). *Science content standards for California public schools: Kindergarten through grade twelve*. Sacramento, CA: Author.

Carless, D. (2005). Prospects for the implementation of assessment for learning. *Assessment in Education, 12*(1), 39–54.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2011). *The long-term impacts of teachers: Teacher value-added and student outcomes in adulthood. Working Paper 17699*. Cambridge, MA: National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w17699

Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education, 18*(8), 947–967.

Cobb, P., McClain, K., Lamberg, T. D., & Dean, C. (2003). Situating teachers' instructional practices in the institutional setting of the school and district. *Educational Researcher, 32*(6), 13–24.

Cornett, J., & Knight, J. (2008). Research on coaching. In J. Knight (Ed.), *Coaching: Approaches and perspectives* (pp. 192–216). Thousand Oaks, CA: Corwin Press.

Cowie, N. (1995). Students of process writing need appropriate and timely feedback on their work, and in addition, training in dealing with that feedback. *Saitama University Review, 31*(1), 181–194.

Cuban, L. (1999). *How scholars trumped teachers: Change without reform in university curriculum, research, and teaching, 1890–1990*. New York: Teachers College Press.

Cuban, L., Kirkpatrick, H., & Peck, C. (2001). High access and low use of technologies in high school classrooms: Explaining an apparent paradox. *American Educational Research Journal, 38*, 813–834.

Curry, M., & Killion, J. (2009). Slicing the layers of learning: Professional learning communities fill the gaps as educators put new knowledge into practice. *JSD, 30*(1), 56–62.

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives, 8*. Retrieved from http://epaa.asu.edu/ojs/article/view/392/515

Darling-Hammond, L., Wei, R. C., Andree, A., Richardson, A., & Orphanos, S. (2009). *Professional learning in the learning profession: A status report on teacher development in the United States and abroad*. Oxford, OH: National Staff Development Council.

Darling-Hammond, L., & Wood, G. (2008). *Democracy at risk: The need for a new federal policy in education*. Washington, DC: The Forum for Education and Democracy.

Desimone, L. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher, 38*(3), 181–199.

Desimone, L., Porter, A., Garet, M., Yoon, K., & Birman, B. (2002). Effects of professional development on teacher instruction: Results from a three-year longitudinal study. *Educational Evaluation and Policy Analysis, 24*(2), 81–112.

DiRanna, K., Osmundson, E., Topps, J., Barakos, L., Gearhart, M., Cerwin, K., Carnahan, D., & Strang, C. (2008). *Assessment-centered teaching: A reflective practice*. Thousand Oaks, CA: Corwin Press.

DiRanna, K., Topps, J., Cerwin, K., & Gomez-Zwiep, S. (2009). Teacher learning collaborative: A process for supporting professional learning ⭷

communities. In S. Mundry & K. Stiles (Eds.), *Professional learning communities for science teaching* (pp. 35–53). Arlington, VA: National Science Teachers Association Press.

Doppelt, Y., Schunn, C., Silk, E., Mehalik, M., Reynolds, B., & Ward, E. (2009). Evaluating the impact of a facilitated learning community approach to professional development on teacher practice and student achievement. *Educational Administration Quarterly, 27*(3), 339–354.

Elliott, S. N., Kettler, R. J., Beddow, P. A., & Kurz, A. (2010). Research and strategies for adapting formative assessments for students with special needs. In P. E. Noyce & D. T. Hickey (Eds.), *New frontiers in formative assessment* (pp. 159–180). Cambridge, MA: Harvard Education Press.

Elmore. R. (1996). Getting to scale with good educational practice. *Harvard Educational Review, 66*(1), 1–26.

Elmore, R. (2002). *Bridging the gap between standards and achievement: Report on the imperative for professional development in education*. Washington, DC: Albert Shanker Institute.

FAST SCASS. (2008). *Attributes of effective formative assessment: A work product coordinated by Sarah McManus, NC Department of Public Instruction, for the Formative Assessment for Students and Teachers (FAST) Collaborative*. Washington, DC: Council of Chief State School Officers.

Frey, N., & Fisher, D. (2009). Using common formative assessments as a source of professional development in an urban American elementary school. *Teaching and Teacher Education, 25*, 674–680.

Fuchs, D., Fuchs, L. S., Compton, D. L., Bouton, B., Caffrey, E., & Hill, L. (2007). Dynamic assessment as responsiveness to intervention. *Teaching Exceptional Children, 39*(5), 58–63.

Fuchs, L. (2004). The past, present, and future of curriculum-based measurement research. *School Psychology Review, 33*(2), 188–192.

Furtak, E. M., & Ruiz-Primo, M. A. (2008). Making students' thinking explicit in writing and discussions: An analysis of formative assessment prompts. *Science Education, 92*(5), 799–824. Retrieved from http://www.interscience.wiley.com

Furtak, E. M., Ruiz-Primo, M. A., Shemwell, J. T., Ayala, C. C., Brandon, P. R., Shavelson, R. J., & Yin, Y. (2008). On the fidelity of implementing embedded formative assessments and its relation to student learning. *Applied Measurement in Education, 21*, 360–389.

Garet, M. S., Ludwig, M., Yoon, K., Wayne, A., Birman, B., & Milanowski, A. (2011, April). *Making professional development more strategic: A conceptual model for district decision-makers*. Paper presented at the American Educational Research Association conference, New Orleans, LA.

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal, 38*(4), 915–945.

Guskey, T. R. (2000). *Evaluating professional development*. Thousand Oaks, CA: Corwin Press.

Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A review of multiple-choice item–writing guidelines for classroom assessment. *Applied Measurement in Education, 15*(3), 309–334.

Hargreaves, A., & Fullan, M. (2012). *Professional capital: Transforming teaching in every school*. New York: Teachers College Press.

Heritage, M. (2010). *Formative assessment: Making it happen in the classroom*. Thousand Oaks, CA: Corwin Press.

Heritage, M., & Heritage, J. (2011, April). *Teacher questioning: The epicenter of formative instruction and assessment*. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA.

Heritage, M., Kim, J., Vendlinski, T., & Herman, J. L. (2009). From evidence to action: A seamless process of formative assessment? *Educational Measurement: Issues and Practice, 28*(3), 24–31.

Heritage, M., & Popham, W. J. (2008). Tight but loose: Through the looking glass. In E. C. Wylie (Ed.), *Tight but loose: Scaling up teacher professional development in diverse contexts* (ETS Research Report RR-08-29) (pp. 125–133). Princeton, NJ: Educational Testing Service.

Herman, J. L. (2010). *Coherence, key to next generation assessments*. (AACC Report.) Los Angeles: University of California.

Herman, J. L., Osmundson, E., Ayala, C., Schneider, S., & Timms, M. (2006). *The nature and impact of teachers' formative assessment practices*. CRESST Report 703. Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing.

Hirsh, S. (2011). *Building professional development to support new student assessment systems*. Oxford, OH: Learning Forward/National Staff Development Council.

Ishii, D. K. (2003). First-time teacher-researchers use writing in middle school mathematics instruction. *The Mathematics Educator, 13*(2), 38–46.

Killion, J. (2012). *Meet the promise of content standards: Professional learning required*. Oxford, OH: Learning Forward.

Leahy, S., & Wiliam, D., (2009, April). *From teachers to schools: Scaling up professional development for formative assessment*. Paper presented at the American Educational Research Association conference, San Diego, CA. Retrieved from http://dylanwiliam.net

Learning Forward. (2011). *Standards for professional learning*. Oxford, OH: Author.

Loucks-Horsley, S., Love, N., Stiles, K. E., Mundry, S., & Hewson, P. W. (2003). *Designing professional development for teachers of science and mathematics* (2nd ed.). Thousand Oaks, CA: Corwin.

Loucks-Horsley, S., Stiles, K. E., Mundry, S., Love, N., & Hewson, P. W. (2010). *Designing professional development for teachers of science and mathematics* (3rd ed.). Thousand Oaks, CA: Corwin.

Louis, K. S., Leithwood, K., Wahlstrom, K., & Anderson, S. (2010). *Learning from leadership: Investigating the links to improved student learning.* Minneapolis, MN: Center for Applied Research and Educational Improvement, University of Minnesota.

Lyon, C. J., Cleland, D., & Gannon, M. (2008). Letting go of the reins: Learning about scalability in the context of one district-wide implementation. In E. C. Wylie (Ed.), *Tight but loose: Scaling up teacher professional development in diverse contexts* (ETS Research Report RR-08-29) (pp. 67–78). Princeton, NJ: Educational Testing Service.

Mandinach, E. B., & Jackson, S. S. (2011). *Transforming teaching and learning through data-driven decision making.* Thousand Oaks, CA: Corwin.

McLaughlin, M. (1976). Implementation as mutual adaptation: Change in classroom organization. *Teachers College Record, 77*(3), 339–351.

McLaughlin, M., & Talbert, J. (1993). *Contexts that matter for teaching and learning: Strategic opportunities for meeting the nation's educational goals.* Palo Alto, CA: Stanford University Center for Research on the Context of Secondary School Teaching.

McLaughlin, M., & Talbert, J. (2006). *Building school-based teacher learning communities.* New York: Teachers College Press.

Moss, C. M., Brookhart, S. M., & Long, B. A. (2012). *Administrators' roles in helping teachers use formative assessment information.* Manuscript submitted for publication.

National Research Council. (1996). *Classroom science education standards.* Washington, DC: National Academies Press.

Newman, D. L., & Gullie, K. (2010). *Syracuse City School District Title II B Mathematics and Science Partnership: Mathematics initiative. Year three report 2009–2010.* Albany: University at Albany/State University of New York, The Evaluation Consortium.

Noyce, P. E., & Hickey, D. T. (Eds.). (2011). *New frontiers in formative assessment.* Cambridge, MA: Harvard Education Press.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension monitoring activities. *Cognition and Instruction, 1*(2), 117–175.

Palucci, D. M. (2010). *Fostering data-driven decision-making in a professional learning community of special-education teachers* (Unpublished doctoral dissertation). Wilmington University, Wilmington, Delaware.

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.). (2001). *Knowing what students know: The science and design of educational assessment.* Washington, DC: National Academies Press.

Putnam, R., & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning? *Educational Researcher, 29*(1), 4–15.

Reeves, D. B. (2010). *Transforming professional development into student results.* Alexandria, VA: Association for Supervision and Curriculum Development.

Ritchie, R., & Woods, P. A. (2007). Degrees of distribution: Towards an understanding of variations in the nature of distributed leadership in schools. *School Leadership and Management, 27*(4), 363–381.

Rogoff, B., Matusov, E., & White, C. (1996). Models of teaching and learning: Participation in a community of learners. In D. R. Olsen & N. Torrance (Eds.), *The handbook of education and human development: New models of learning, teaching and schooling* (pp. 388–414). Malden, MA: Blackwell.

Rowan, B., Correnti, R., & Miller, R. J. (2002, November). *What large-scale,* survey research tells us about teacher effects on student achievement: Insights from the Prospects study of elementary schools. CPRE Research Report Series RR-051. Philadelphia: Consortium for Policy Research in Education, University of Pennsylvania Graduate School of Education. Retrieved from http://www.cpre.org/images/stories/cpre_pdfs/rr51.pdf.

Ruiz-Primo, M. A., & Furtak, E. M. (2007). Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry. *Journal of Research in Science Teaching, 44*(1), 57–84.

Sato, M., Wei, R. C., & Darling-Hammond, L. (2008). Improving teachers' assessment practices through professional development: The case of National Board certification. *American Educational Research Journal, 45*(3), 669–700.

Saunders, W., Goldenberg, C., & Gallimore, R. (2009). Increasing achievement by focusing grade-level teams on improving classroom learning: A prospective, quasi-experimental study of Title I schools. *American Educational Research Journal, 46*(4), 1006–1033.

Schneider, M. C., & Randel, B. (2010). Research on characteristics of effective professional development programs for enhancing educators' skills in formative assessment. In H. Andrade & G. Cizek (Eds.), *Handbook of formative assessment* (pp. 251–276). New York: Routledge.

Shavelson, R. J., Ruiz-Primo, M. A., & Wiley, E. (2005). Windows into the mind. *Higher Education, 49,* 413–440.

Shepard, L. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*(7), 4–14.

Shepard, L. (2005, October). *Formative assessment: Caveat emptor.* Paper presented at the ETS Invitational Conference 2005, The Future of Assessment: Shaping Teaching and Learning, New York, NY.

Stiggins, R. (2005). From formative assessment to assessment FOR learning: A path to success in standards-based schools. *Phi Delta Kappan, 87*(4), 324–328.

Stiggins, R. (2010). Essential formative assessment competencies for teachers and school leaders. In H. Andrade & G. Cizek (Eds.), *Handbook of formative assessment* (pp. 233–250). New York: Routledge.

Sztajn, P., Confrey, J., Wilson, P. H., & Edgington, C. (2012). Learning trajectory based instruction: Toward a theory of teaching. *Educational Researcher, 41*(4), 147–156.

Thomas, J. W., Mergendoller, J. R., & Michaelson, A. (1999). *Project-based learning: A handbook for middle and high school teachers*. Novato, CA: The Buck Institute for Education.

Thompson, M., & Wiliam, D. (2008). Tight but loose: A conceptual framework for scaling up school reforms. In E. C. Wylie (Ed.), *Tight but loose: Scaling up teacher professional development in diverse contexts* (ETS Research Report RR-08-29) (pp. 1–43). Princeton, NJ: Educational Testing Service.

Trumbull, E., & Lash, A. (2013). *Understanding formative assessment: Insights from learning theory and measurement theory*. San Francisco: WestEd.

Wenger, E. (1998). Communities of practice: Learning as a social system. *Systems Thinker*. Retrieved from http://www.co-i-l.com/coil/knowledge-garden/cop/lss.shtml

Wiggins, G., & McTighe, J. (2005). *Understanding by design* (expanded 2nd ed.). Alexandria, VA: ASCD.

Wiliam, D. (2007a). Content then process: Teacher learning communities in the service of formative assessment. In D. Reeves (Ed.), *Ahead of the curve: The power of assessment to transform teaching and learning* (pp. 182–204). Bloomington, IN: Solution Tree.

Wiliam, D. (2007b). Keeping Learning on Track: Classroom assessment and the regulation of learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 1051–1098). Greenwich. CT: Information Age Publishing.

Wiliam, D., Lee, C., Harrison, C., & Black, P. J. (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education: Principles, Policy, & Practice, 11*(1), 49–65.

Wilson, M. (Ed.). (2004). *Towards a coherence between classroom assessment and accountability. 103rd Yearbook of the National Society for the Study of Education, Part II.* Chicago: University of Chicago Press.

Wylie, E. C. (Ed.). (2008). *Tight but loose: Scaling up teacher professional development in diverse contexts* (ETS Research Report RR-08-29). Princeton, NJ: Educational Testing Service.

Wylie, E. C., & Heritage, M. (2010). Developing and deepening formative assessment practice. In M. Heritage, *Formative assessment: Making it happen in the classroom* (pp. 117–132). Thousand Oaks, CA: Corwin Press.

Wylie, E. C., Lyon, C. J., & Goe, L. (2009). *Teacher professional development focused on formative assessment: Changing teachers, changing schools* (ETS Research Report RR-08-29). Princeton, NJ: Educational Testing Service. Retrieved from http://www.ets.org/Media/Research/pdf/RR-09-10.pdf

Wylie, E. C., Lyon, C. J., & Mavronikolas, E. (2008). *Effective and scalable teacher professional development: A report of the formative research and development* (ETS Research Report RR-08-65). Princeton, NJ: Educational Testing Service.

Wylie, C., Thompson, M., Lyon, C., & Snodgrass, D. (2008). Keeping Learning on Track in an urban district's low performing schools. In E. C. Wylie (Ed.), *Tight but loose: Scaling up teacher professional development in diverse contexts* (ETS Research Report RR-08-29) (pp. 79–92). Princeton, NJ: Educational Testing Service.

Yoon, K. S., Duncan, T., Lee, S. W.-Y., Scarloss, B., & Shapley, K. (2007). *Reviewing the evidence on how teacher professional development affects student achievement* (Issues & Answers Report, REL 2007-No.033). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southwest. Retrieved from ies.ed.gov/ncee/edlabs/regions/southwest/pdf/rel_2007033_sum.pdf

Young, M. J. (2008–2011). *Evaluation report: Lake Elsinore CaMSP cohort 5*. Tucson, AZ: MJ Young Associates.