

# Understanding Formative Assessment

## *Insights from Learning Theory and Measurement Theory* ▲

by Elise Trumbull and Andrea Lash

*This paper explores formative assessment, a process intended to yield information about student learning—information that teachers can use to shape instruction to meet students' needs and that students can use to better understand and advance their learning. This purpose—promoting learning by informing instruction—distinguishes it from other kinds of student assessment, such as diagnostic, which is used to identify students who have special learning needs, or summative, which is used by teachers to form final judgments about what students have learned by the end of a course, or is used at the state level for the purpose of evaluating schools.*

This paper comes at a time in education when this last purpose, school accountability, has been dominating assessment use for more than a quarter of a century (Shepard, 2006). Since implementation of No Child Left Behind in 2001, state departments of education have assessed students annually in English language arts and mathematics with tests that survey a broad spectrum of content. Although each student is assessed, these tests are not intended to help identify an individual student's learning needs or to provide information that can be used to modify subsequent instruction. Instead, the tests serve an accounting or monitoring function, such as counting

the number of individuals who meet grade-level standards; test results of individual students are aggregated into reports of school and district progress, reports that are useful for district- and state-level decision-makers. But while such tests may identify students who lack the knowledge and skills expected for their grade level, these achievement tests do not identify why students are not proficient; the tests are not linked closely enough to classroom instruction and curriculum to identify what misconceptions students hold or what skills they are missing, information that could help guide instruction.

Increasingly, educators are calling for education assessment systems that are more balanced, and that yield useful information for a variety of education purposes, from how to shape ongoing instruction in the classroom to accountability decisions made at the state level (Darling-Hammond & Pecheone, 2010; Pellegrino, 2006; Wilson & Draney, 2004; Pellegrino, Chudowsky, & Glaser, 2001). They are also calling for coherent systems, in which assessments at all levels (from classroom to state) would be aligned with the same learning goals and views of what constitutes learning and would produce relevant information about student learning over time (Herman, 2010; Pellegrino, 2006).

The purpose of this paper is to help readers understand the importance and potential of formative assessment as a key component of

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

WestEd   
WestEd.org

**Formative assessment is defined by its purpose which is to help form, or shape, a student's learning during the learning process.**

a balanced and coherent assessment system—a component that has been somewhat eclipsed by the focus on assessment for accountability purposes. The paper first describes formative assessment and its key features. It then turns to learning theory and measurement theory and their implications for effective use of formative assessment. Subsequent to that, and prior to the conclusion, is a brief review of summaries of research on how formative assessment affects student learning.

**Features of Formative Assessment in Classroom Instruction**

Black and Wiliam (1998a) characterize formative assessment as “all those activities undertaken by teachers and/or by their students [that] provide information to be used as feedback to modify the teaching and learning activities in which they are engaged” (p. 7). The goal of any modifications to instruction is enhanced student learning. It is often claimed that the practice of formative assessment is rooted in Bloom’s concept of “mastery learning,”

an instructional approach that espouses the use of assessments to gauge students’ progress toward mastering a learning goal (Bloom, Hastings, & Madaus, 1971). Bloom suggested that, rather than waiting to assess students at the end of a unit (common practice at the time), teachers use assessments “as an integral part of the instructional process to identify individual learning difficulties and prescribe remediation procedures” (Guskey, 2010, p. 108). According to Guskey, Bloom borrowed the term “formative” from Scriven (1967), who used it to describe program evaluation activities conducted *during* the course of a program to give feedback on the program’s progress so that it could be improved if need be.

Formative assessment does not take the form of a particular instrument or task (Moss, 2008), but is defined by its purpose (Shepard, 2009), which is to help form, or shape, a student’s learning during the learning process. Some suggest that formative assessment is better described as a *process* (“using assessment formatively” [Frohbeiter, Greenwald, Stecher, & Schwartz, 2011, p. 3]) than as a type of assessment (see also McManus, 2008). Erickson (2007) has used the term “proximal formative assessment” to indicate that it is an activity close to instruction (Ruiz-Primo, Shavelson, Hamilton, & Klein, 2002). Erickson (2007) defines it as “the continual ‘taking stock’ that teachers do by paying firsthand observational attention to students during the ongoing course

of instruction—careful attention focused upon specific aspects of a student’s developing understanding” (p. 187) in order to make decisions about next steps in instruction (see also Heritage, Kim, Vendlinski, & Herman, 2009). To facilitate this process, the teacher needs to use practices that will reveal not only *whether* a student appears to have mastered a concept but also *how* he or she understands it (Pryor & Crossouard, 2005). The assessment practices need to be so well grounded in the instructional process that the information they reveal will identify whether and how instruction should be adapted to advance students’ understandings. Heritage, who has made significant contributions to the theory and practice of formative assessment, emphasizes the close linkage—if not the inseparability—of formative assessment, teaching, and learning (Heritage, 2010a).

In theory, any assessment—including a commercially developed test—could be used for formative purposes. However, as Pellegrino et al. (2001) caution, using the same assessments for different purposes tends to lessen their effectiveness for each purpose (see also Shavelson, Black, Wiliam, & Coffey, 2007). For example, it would be difficult to design an assessment for school accountability systems that elicits student performance at the level necessary for fine-grained understanding of individual learning needs without compromising the scope necessary for an accountability measure or without making excessive time

demands for administration and scoring. Such accountability assessments are generally not coupled closely enough to instruction to yield information that would help a teacher think about what a student might need in order to better learn what has been assessed.

To serve a formative purpose, assessment needs to provide actionable information for teachers and students (Heritage, 2010a; Shepard, 2005). Ideally, it reveals something about a student's progress toward certain learning goals, the student's thought processes, and any misconceptions the student may hold (Supovitz, 2012). Formative assessment is highly "contingent" on the instructional situation and the student(s) (Black & Wiliam, 2009, p. 12). Thus, it should be tailored to the particular students being assessed, the relevant learning targets, and a specified point in the instructional process; also, it should take a form most likely to elicit the desired learning evidence (Ruiz-Primo & Li, 2011). There can be no prescription for what a single instance of formative assessment should look like. Any instructional activity that allows teachers to uncover the way students think about what is being taught and that can be used to promote improvements in students' learning can serve a formative purpose.

Formative assessment is often highly integrated with instruction (Herman et al., 2006) and most commonly takes the form of classroom exchanges between teachers and students (or, less commonly, between students). These exchanges have the potential

to make students' thinking explicit and thus open to examination and revision. In this way, the exchanges serve as learning opportunities (Ruiz-Primo, 2011). Given insights into students' thinking, a teacher is in a position to counter misconceptions and steer learning back on track through feedback or instructional modifications (Black & Wiliam, 2004). Teachers can also mentor students to become proficient at asking their own questions of each other and responding with ideas, reasoning, and evidence, as well as providing feedback to each other (Black & Wiliam, 1998b). Some have called feedback the "linchpin" that links the components of the formative assessment process (Brookhart, Moss, & Long, 2010, p. 41).

Feedback is "information provided by an agent (e.g., teacher, peer, parent, the assessment itself) regarding aspects of one's performance or understanding" (Hattie & Timperley, 2007, p. 81). Feedback takes on a formative role when it provides information about the gap between a student's current understanding and the desired level of understanding, and it is most effective for the student when it is targeted at the right developmental level and helps the student identify ways to close the gap (Hattie & Timperley, 2007; Sadler, 1989). Feedback helps students clarify the goals of learning, their progress toward such goals, and what they need to do to reach the goals (Hattie & Timperley, 2007). The challenge for a teacher is to gain insight into students' way of thinking about the subject matter at hand and to frame feedback

**There can be no prescription for what a single instance of formative assessment should look like. Any instructional activity that allows teachers to uncover the way students think about what is being taught and that can be used to promote improvements in students' learning can serve a formative purpose.**

that helps them move toward specific learning goals (Black & Wiliam, 2009).

The topic of feedback is large and complex, with a lengthy research history; yet much remains to be done to clarify just how to meet the challenge that Black and Wiliam (2009) identify. Research in classrooms (not laboratory settings) documenting how feedback is used and with what impact over time is particularly needed (Ruiz-Primo & Li, 2013).

Heritage and Heritage (2011) refer to teacher questioning as "the epicenter of instruction and assessment" (title). Teachers' questioning during instruction may be informal and spontaneous or may be formal and planned prior to the lesson (Shavelson et al., 2008). A teacher's informal questions to students during class may be for

### Exhibit 1. Some Dimensions on Which Formative Assessment May Vary

<b>1. Informal vs. formal</b>
<b>2. Immediate feedback vs. delayed feedback</b>
<b>3. Curriculum embedded vs. stand-alone</b>
<b>4. Spontaneous vs. planned</b>
<b>5. Individual vs. group</b>
<b>6. Verbal vs. nonverbal</b>
<b>7. Oral vs. written</b>
<b>8. Graded/scored vs. ungraded/unscored</b>
<b>9. Open-ended response vs. closed/constrained response</b>
<b>10. Teacher initiated/controlled vs. student initiated/controlled</b>
<b>11. Teacher and student(s) vs. peers</b>
<b>12. Process oriented vs. task/product oriented</b>
<b>13. Brief vs. extended</b>
<b>14. Scaffolded (teacher supported) vs. independently performed</b>

the purpose of checking certain students' learning, or for probing more deeply to gather evidence that will yield better understanding of their thinking. At the other end of the spectrum of formative assessment are more formal procedures, such as specific prompts that require a written response and that are embedded in instruction at key points to help identify the next steps needed to advance student learning (Furtak et al., 2008). These embedded tasks may be so integrated with instruction as to seem natural and unobtrusive, or they may be given to students at the end of a lesson, as a separate activity, such as when

students make entries in their science notebooks for the teacher to examine later.

Formative assessments can be described along a number of different dimensions. Some of the most salient dimensions are listed in Exhibit 1 above. While formative assessments may vary on a number of dimensions, "the crucial feature is that evidence is evoked, interpreted in terms of learning needs, and used to make adjustments [to instruction] to better meet those learning needs" (William, 2006, p. 3).

As noted earlier, because formative assessment is so tightly linked

to instruction, there is a conceptual question as to whether formative assessment is more like instruction or more like assessment, as traditionally conceived. Some writers (e.g., Heritage 2010a) situate formative assessment within a paradigm of learning and instruction; others (e.g., Phelan et al., 2009) have placed it squarely within a measurement paradigm. The following sections examine formative assessment within each paradigm because both contain concepts that are helpful to understanding effective use of formative assessment.

### Formative Assessment Within a Theory of Learning and Instruction

Formative assessment is not necessarily associated with any particular theory of learning (William, 2010). However, current conceptualizations of formative assessment are typically rooted in a sociocultural constructivist view of learning (Heritage, 2010a; Pellegrino et al., 2001; Shepard, 2000). This theory of learning is supported by research (Pellegrino et al., 2001), is most compatible with current goals of education, and best explains the processes of effective formative assessment (Heritage, 2010b; Pellegrino et al., 2001; Shepard, 2000).

From a sociocultural constructivist perspective, learners are seen as actively constructing knowledge and understanding through cognitive processes (Piaget, 1954) within a social and cultural context (Greenfield, 2009; Rogoff,

1998, 2003; Vygotsky, 1978); as building new knowledge on what they already know (i.e., prior knowledge) (Bransford, Brown, & Cocking, 2000); and as developing the metacognitive skills necessary to regulate their own learning (Bransford et al., 2000; Bruner, 1985; Vygotsky, 1978). These understandings about learning and development have implications for the use of formative assessment in classroom instruction.

The work of Vygotsky (1962, 1978) forms much of the basis for current conceptualizations of the sociocultural aspects of constructivist learning theory and has been widely applied to models of formative assessment. Students are seen to develop knowledge and understanding in a domain over time, not only as individuals but in an interactive social context, guided by others with greater expertise (e.g., teacher, parent, peer) (Tharp & Gallimore, 1991; Vygotsky, 1978; Wenger, 1998). One assumption of sociocultural theory is that learning is enhanced by what Vygotsky referred to as “joint productive activity” within a social setting, such as in a classroom where students and teachers collaborate as a community of learners (Ash & Levitt, 2003; Koschmann, 1999).

The “zone of proximal development” (ZPD), a concept taken from Vygotsky (1978), has been invoked by formative assessment theorists as useful for understanding the gap between a student’s actual understanding and the student’s targeted or potential learning. The ZPD is the developmental space between the level at which

a student can handle a problem or complete a task independently and the level at which the student can handle or complete the same task with assistance from a more competent other, such as a teacher. Work within the ZPD is a particular example of joint productive activity, that is, teacher and student are working jointly to ensure that the student reaches a learning goal (Ash & Levitt, 2003). In teaching, the teacher serves as a mediator between the student and the learning goal, providing scaffolding (i.e., learning support) to aid attainment of the goal (Black & Wiliam, 2009; Walqui & van Lier, 2010). Formative assessment is part of this process—whether implicitly or explicitly—as the teacher uses information about how a student responds to instruction in order to give feedback to the student and/or adjust instruction so as to prompt learning or performance. In this case, formative assessment is almost indistinguishable from instruction, as the teacher introduces content; assesses how the student is responding; offers supports for understanding and modifies instruction as needed; re-assesses how the student’s learning is progressing; continues with new content or returns in a new way to the same content, and so forth.

### The Roles of Teachers and Students in Formative Assessment

The kind of classroom evoked by the sociocultural constructivist theory of learning is one in which teachers and students share responsibility for learning

(Heritage, 2010a; Tunstall & Gipps, 1996). In this classroom, one would see teacher and students working together as part of an interactive community of learners, in roles that may be new to some (Brown & Campione, 1994; Rogoff, 1994), including engaging in formative assessment. Formative assessment calls upon teachers not only to determine *whether* students have learned something, but also to probe students’ ways of thinking to get at *why* any learning gaps exist. In addition to using assessment evidence to plan future instruction, teachers are expected to use it to help students (1) judge the state of their own knowledge and understanding, (2) identify the demands of a learning task, (3) judge their own work against a standard, (4) grasp and set learning goals, and (5) select and engage in appropriate strategies to keep their learning moving forward (Andrade, 2010; Black & Wiliam, 1998b, 2009; Bransford et al., 2000; Heritage, 2010b; Stiggins, Arter, Chappuis, & Chappuis, 2009). These metacognitive skills are critical to the development of intentional learning and of independent, self-propelled learners who can regulate their own learning and self-correct as needed (Bransford et al., 2000).

Students are expected to be active agents in their own learning by engaging, in increasingly independent ways, in the previously enumerated skills (Clark, 2012). As Black and Wiliam (2009) observe, “[S]ince the responsibility for learning rests with *both* the teacher *and* the learner, it is incumbent on each to do all they can to ■■

**Formative assessment places demands on students to take a more serious approach to learning and to work harder.**

mitigate the impact of any failures of the other” (p. 7). International studies on the impact of formative assessment practices show that such practices can indeed support students’ ability to take responsibility for and regulate their own learning, but that this occurs only when students understand that assessment can serve purposes other than summative purposes (Organization for Economic Co-operation and Development, 2005). Perrenoud (1991) notes that formative assessment places demands on students to take a more serious approach to learning and to work harder—demands they may not happily embrace; however, when they do, they may be their own best sources of feedback about their own learning. Student self-assessment does contribute to higher student achievement, and it is most likely to do so when students are trained in using sets of performance criteria, such as rubrics, to evaluate their work or when they receive other direct instruction on self-assessment (Ross, 2006). While the self-assessments of students may not always be in sync with their teachers’ assessments of them, discrepancies can form the basis of

“productive conversations about student learning needs” (Ross, 2006, p. 9).

Some forms of formative assessment require students not only to be active agents in their own learning but also to be, at times, facilitators of each other’s learning through a process of peer assessment. Peer assessment has students serving as instructional resources to each other in much the way that collaborative learning does (Black & Wiliam, 2009). Students’ feedback to each other during peer assessment is another source of information about their level of understanding (Black & Wiliam, 2009). For students to adopt such roles requires that they have a clear understanding of learning goals and performance criteria. Some suggest having teachers and students jointly construct assessment criteria in order to increase the reliability of peer assessment (Topping, 2010) or having teachers model the process for students in order to facilitate their participation (Black & Wiliam, 2009, p. 25).

### **The Role of Learning Progressions in Formative Assessment**

A learning progression is a kind of developmental model (Harris, Bauer, & Redman, 2008) that describes “the trajectory of learning in a domain” over an extended period of time—months to years (Heritage, 2008, p. 3). Learning progressions, also known as “learning trajectories” (Sztajn, Confrey, Wilson, & Edgington, 2012) and “progress maps” (Hess, 2010), have been defined as “descriptions of

successively more sophisticated ways of thinking about an idea that follow one another as students learn: [The descriptions] lay out in words and examples what it means to move toward more expert understanding” (Wilson & Bertenthal, 2006, p. 3). Learning progressions help teachers decide where to take instruction next, based on what they have observed students being able to do independently and with support (i.e., within the ZPD).

Learning progressions are intended to help teachers organize the curricular topics associated with standards. In some cases, learning progressions can be constructed logically, with reference to what experts in a domain perceive as a necessary sequence. For example, in a mathematics curriculum, addition logically precedes multiplication because multiplication is repeated addition and because a child is unlikely to have avoided learning addition before being able to understand multiplication (Leahy & Wiliam, 2011). In other cases, the “logical” progression may not capture a learner’s likely developmental path. In addition, learning progressions may vary to some degree from student to student and from country to country. For these reasons, there is no substitute for empirical validation of a learning progression.

Learning progressions or trajectories can help teachers to anticipate and identify common misconceptions students may have and, thus, to shape feedback—which, in turn, reshapes learning (Sztajn et al., 2012). Sztajn et al. write of “learning trajectory based instruction” as a promising approach

that brings teaching and learning theory together in a way not previously done: “Overall, we contend that, despite disciplines, when teachers organize teaching around learning from [a learning trajectory] perspective, the trajectory serves as the unifying element for their instruction” (p. 152).

Very few learning progressions have been empirically validated, so almost any available one needs to be viewed as a tentative heuristic—a way of helping teachers think about learning development in a given domain—as opposed to a map that is faithful to the terrain. Heritage (2008) summarizes seven sample learning progressions in mathematics, history, science, and oral language, as well as a tool that teachers can use to develop their own progressions in science (“Conceptual Flows” [DiRanna et al., 2008]). It is not clear, from her review, whether some or all of the progressions have been empirically validated or are based on logical progressions as identified by domain experts. Sztajn et al. (2012) refer to several different efforts to develop dozens of learning progressions related to different subdomains of mathematics.

Learning progression research is complex and time-consuming, and generalizing on the basis of such research is somewhat risky because of differences in context. Recently, researchers validated a learning progression for linear measurement in grades 2 and 3 (Barrett et al., 2012). They caution, however, that this learning progression is preliminary, having been tested with only eight

students. There is currently no standard way of approaching the framing of learning progressions. Sztajn et al. (2012) note that the progressions resulting from these efforts “varied in span, grain size, use of misconceptions, and level of detail” (p. 148).

Research shows that cognitive development in a domain does not necessarily follow a linear path (Harris et al., 2008; Shavelson & Kurpius, 2012; Steedle & Shavelson, 2009). Moreover, “[p]rogressions are not developmentally inevitable but dependent on instruction interacting with students’ prior knowledge and new-knowledge construction” (Shavelson & Kurpius, 2012, p. 15). Whereas there is not likely to be a single progression for any complex learning goal, many believe that educators will be able to identify paths that are consistent with the ways that many students learn (Mosher, 2011). These common paths can be annotated by teachers as they observe differences in students; this is a necessary step to providing differentiated feedback and instruction for learners who veer from the common path.

Much research remains to be done on learning progressions. Researchers at the Center for Policy Research in Education conclude, “If this work is pursued vigorously and rigorously, the end result should be a solid body of evidence about what most students are capable of achieving in school and about the particular sequence(s) of learning experiences that would lead to proficiency on the part of most students” (Corcoran, Mosher, &

Rogat, 2009, p. 8). Shavelson and Kurpius (2012) believe that experimental research, as well as action research by teams of teachers and researchers, may yield knowledge of how to proceed with the development of defensible learning progressions.

Even though there are not empirically validated developmental sequences for the major concepts and skills in every academic domain, the concept of learning progressions is likely to be helpful to teachers in conducting and interpreting formative assessments. The hypothesized progressions may guide teachers’ explorations of student learning through formative assessment, their decisions about developmentally appropriate feedback to students, and their planning of next instructional steps.

### Formative Assessment Within Measurement Theory

The role of measurement theory with regard to formative assessment is somewhat contested; it is not altogether clear whether and, if so, how and to what degree accepted measurement principles should guide formative assessment (Bennett, 2011). This section discusses ways in which established thinking about measurement in general may contribute to conceptualizing and designing effective formative assessment, as well as ways in which traditional practices based on principles of measurement theory may *not* be applicable to formative assessment. The

**As in other forms of assessment, the primary activity for the person using formative assessment results is to reason from evidence—to make an inference about what a student knows and can do, based on assessment information that is not perfect and may be, to some degree, incomplete or imprecise.**

section concludes with a discussion of why language and culture should be considered when planning and implementing formative assessment, so as not to bias results.

As in other forms of assessment, the primary activity for the person using formative assessment results is to reason from evidence—to make an inference about what a student knows and can do, based on assessment information that is not perfect and may be, to some degree, incomplete or imprecise (Pellegrino et al., 2001). Measurement theory identifies desired qualities of the inferences made from assessments: reliability, validity, and fairness. *Reliability* has to do with the consistency of the assessment information—for example, whether replication of an assessment at different times or in different

settings would result in the same judgment about the student (Haertel, 2006). *Validity* has to do with the extent to which the interpretation of a student's performance and the actions based on it are appropriate and justified (Messick, 1989). Are the decisions made on the basis of students' performance suitable and accurate? *Fairness* requires that validity does not change from one student group to another (Pellegrino et al., 2001). For example, are the interpretations of student performance as appropriate for students who are English learners as they are for students who are native speakers of English?

Measurement theory also provides statistical methods to assess the qualities of inferences. In large-scale assessments of achievement—such as statewide testing for school accountability—reliability, validity, and fairness are examined in statistical studies that are based on measurement models about the factors that influence student performance on tests. These statistical methods would not be helpful in formative assessment conducted in classrooms, for a couple of reasons. First, they require performance information from a large number of students, on a large number of tasks, possibly from multiple occasions. In classrooms, a teacher might use a particular assessment technique simply to evaluate a few students in a brief segment of a class discussion. Second, the statistical analyses generally are built on theories of test interpretation that summarize the quantity, rather than the qualities, of student knowledge. Thus,

the interpretations would focus on whether a student's test performance indicated that the student had acquired an adequate level of knowledge, rather than focusing on the nature of the student's reasoning or the patterns of thinking displayed by the student. It is this last type of information that generally is useful to teachers in understanding what students know and what they still need to learn.

Exploration of the qualities of inferences derived from formative assessment is in its infancy. Still to be investigated are such issues as how the types of strategies that teachers use in formative assessment affect the quality of evidence elicited from students, whether the strategies are interchangeable with regard to the instructional decisions to which they lead, and whether the strategies differ in effectiveness for different students (Brookhart, 2003, 2005; Shepard, 2009).

There are good reasons to believe that concerns for reliability, validity, and fairness are mitigated by the nature of how formative assessment is carried out. With formative assessments, teachers can evaluate students frequently via different strategies that can be tailored to the particular students (Durán, 2011). In formative assessment, Shavelson et al. (2007) argue, issues of reliability and validity are addressed over time, as teachers collect ongoing data about student performance and, as appropriate, make corrections to their previous inferences. Teachers are in an ideal position to adjust their methods to probe information that will resolve any

discrepancies or to test competing hypotheses as to why students respond the ways they do.

### Useful Principles of Assessment Design

An understanding of fundamental principles of assessment design can be useful to teachers in their efforts to obtain high-quality information from students. One useful heuristic is an assessment triangle that shows the three elements present in any type of assessment: a *model of student cognition*, which describes how students develop competence in an academic domain and how they organize their knowledge at different levels of development; *observations*, which are the tasks or activities in which students' performance can be observed, scored, and evaluated for the purpose of gathering evidence of learning; and *interpretation*, which is the rationale for making sense of and deriving inferences from the evidence gathered (Pellegrino et al., 2001, pp. 44–51). Whether or not they are made explicit, these elements are equally present in any instance of assessment, including formative assessment, and the quality of inferences derived from the assessment will depend on how well these three elements have been linked (Pellegrino et al., 2001). In formative assessment, a fourth element needs to be present: effective translation of the interpretation of assessment performance to instructional decisions and actions.

One approach to assessment development that makes explicit, and links, the three elements of the

assessment triangle is Evidence Centered Design (ECD) (Mislevy, Steinberg, & Almond, 2003; Zhang et al., 2010). ECD provides a framework for building valid and fair assessments. In this process, assessment developers identify the nature of the evidence that is needed to make a judgment about specified aspects of student learning; then, they examine any proposed assessment task to ensure that it does not preclude or reduce the opportunity for any student to participate in the task and show certain knowledge, skills, and abilities (KSAs). Sometimes an assessment task, including a formative assessment, may call on additional, unwanted (non-target) KSAs, and the task may end up eliciting evidence not only of the target KSAs but also of language skill or some other skill not related to the concepts ostensibly being assessed. In such a case, lack of proficiency with non-target KSAs can prevent the student from demonstrating proficiency with the actual targets of the assessment, limiting the assessment's fairness and the validity of interpretations derived from it. An example would be a task designed to evaluate students' understanding of a mathematical concept by having students solve a problem that is couched in many pages of text. Performance on the task would be influenced not only by the target KSA—that is, knowledge of the mathematical concept—but also by the non-target KSA—that is, ability to read extended text. The following series of questions (based in part on Harris et al., 2008) are useful when developing formative assessment activities in

general, but can also help teachers develop assessment processes and tools that minimize the intrusion of unwanted KSAs and increase the likelihood of making valid inferences about student learning:

- » What KSA(s) do I wish to assess (e.g., knowledge, skills, processes, understanding toward competency in a particular part of a domain)?
- » What is the cognitive/developmental path (i.e., learning trajectory) I would expect to see with regard to these KSAs?
- » What evidence (i.e., observable features of students' performances and responses) would I need in order to determine the student's level of KSAs?
- » What are the characteristics of tasks that will elicit this evidence?
- » What KSAs that are *not* wanted (e.g., unnecessarily complex language, need for speed of response) might this type of formative assessment process introduce?
- » How can I modify my formative assessment process to make it inclusive for all students, to minimize the impact of non-target KSAs?

Non-target KSAs are most commonly introduced unwittingly by unnecessarily complex language and/or by content or processes unfamiliar to students from particular cultural backgrounds. The following section shows how this can happen and offer suggestions for avoiding such problems. ▀ ▀

## Language and Culture in Formative Assessment

Language is the principal medium for teaching and learning (Bailey, Burkett, & Freeman, 2008), for mentally representing and thinking through problems (Durán, 1985; Lager, 2006), and for gaining an understanding of how other people think (Bronkard, 1995). As Bailey et al. (2008) write, “Classrooms are first and foremost, language environments” (p. 608). However, teachers are generally not educated to think linguistically (Bailey et al., 2008), to see how language is an integral element in all teaching and learning. Hence, language is often a kind of silent partner in instruction and assessment. This is unfortunate because good teaching depends on a teacher’s considerable knowledge of language development and the use of language in learning—a grounding in the fundamentals of educational linguistics (Wong Fillmore & Snow, 2000).

Students’ responses to formative assessments, which teachers expect to interpret as evidence of students’ content knowledge or skill, may be affected by students’ relative familiarity with the forms and uses of language in the assessment tasks. For example, a student may not understand the grammar (form) of a question or may lack the skills to mount an evidence-based argument (one use of language) to respond to the question adequately. The language forms and uses found in students’ classrooms, in both instruction and assessment tasks, may be more familiar or less familiar to students, depending on the dialects

of English spoken at home, the purposes for which their parents use language, the influence of another language or culture, their prior knowledge and past experience (related to opportunities to learn), their socioeconomic status, and a host of other factors (Heath, 1983; Ochs & Schieffelin, 2011; Solano-Flores & Trumbull, 2003).

When educators consider the role of language in assessment, the needs of students who are English language learners may come quickly to mind. These students are learning a new language at the same time they are learning content in that language, learning the specialized discourse of the different subject-matter domains, and learning how to use language as a tool for learning and for demonstrating their learning (Abedi, 2010, 2011; Bailey et al., 2008; Lee, Santau, & Maerten-Rivera, 2011; Trumbull & Solano-Flores, 2011). With these issues in mind, teachers will want to evaluate their formative assessment practices with a view to reducing language demands, providing choices in the ways they expect students to demonstrate understanding of a concept, and rewording the language of an assessment when apparently needed. They can also ask students directly about how they, the students, interpret assessment questions or tasks (Basterra, 2011; Lee et al., 2011; Spinelli, 2008). A student’s difficulty in interpreting the meaning of an assessment question is itself a clue to the presence of one or more non-target KSAs.

Formative assessment activities can be designed to be credible sources of learning evidence with

students who are English language learners (Kopriva & Sexton, 2011). However, having good information about a student’s level of English proficiency is critical to planning appropriate formative assessment processes. Abedi (2010) and others recommend that teachers use the formative process itself to gather and make note of information about students’ levels of English knowledge on a continuous basis. A teacher is in a better position than an outside specialist to observe language use in a range of circumstances and to make judgments about a student’s ability to use language as a medium of learning and assessment.

In written assessments, it is advisable to avoid high-level vocabulary not related to the learning goal being assessed, eliminate complex syntax, and avoid the passive voice (Abedi, 2006)—for any student, not just for English language learners. (See Trumbull & Solano-Flores [2011] for a list of linguistic features to avoid, with explanatory examples.) Unnecessary language complexity is probably the greatest source of non-target KSAs for a great many students (Trumbull & Solano-Flores, 2011). In spoken language, there are opportunities for a teacher to clarify language, and he or she may want to model language that is slightly beyond a student’s level so as to encourage language growth (by working in the ZPD).

Students who are poor readers or have a developmental language problem, but who (given appropriate supports) have the intellectual capacity to learn the taught curriculum, are also penalized

by instructional and assessment uses of language that do not take their needs into account (Fuchs et al., 2007; Lerner, 2000; Troia, 2011). A more complex challenge in formative assessment emerges for students who have a language-based learning disability and are also English language learners (Figueroa & Newsome, 2006; Hoover & Klingner, 2011). Experts in the education of these students see value in formative assessment processes precisely because such assessment is tailored to the students' contexts and can be used on a continuing basis to monitor student progress (National Joint Committee on Learning Disabilities, 2010; Hoover & Klingner, 2011). Should a student need a higher level of intervention, the teacher will have important evidence about the student's needs, from formative assessment, which he or she can share with specialists evaluating the student and with parents.

Cultural differences in students' orientation to language use, as well as their orientation to assessment, must be considered if formative assessment is to be valid and fair. Any use of language in the education process has cultural underpinnings—that is, culture-based assumptions about what is appropriate or acceptable (Trumbull & Solano-Flores, 2011). All students face the task of learning how classroom discourse and the discourses of different subject-matter domains work. But, based on their home language and/or culture, some students may be less prepared for this task than others, because they may not be

oriented to the dominant culture's ways of using language, which are reflected in the classroom (Heath, 1983; Schleppegrell, 2004). These kinds of differences have implications for formative assessment. For example, in some families, parents may routinely talk with their children about how they are progressing in learning a particular skill or what they may be able to do next (Moss, 2008). In other families, such conversations are not common, and, when asked to engage in an evaluative conversation about learning, their children may need more teacher modeling and more time to develop comfort and skill with this process. Explaining the school's expectations to parents while, at the same time, respecting parents' differences in expectations for children may be an important step for teachers to take. If parents believe that children learn best by listening and showing respect for the teacher, they may discourage the very behaviors that teachers are trying to elicit (Greenfield, Quiroz, & Raeff, 2000). It is not surprising that cultural orientation has been shown to affect students' ways of approaching tests or test items (Demmert, 2005; Li, Solano-Flores, Kwon, & Tsai, 2008; Swisher & Deyhle, 1992). In fact, the expectations for how to communicate during assessment, whether formal or informal, constitute what is, in effect, a "cultural script" (Emihovich, 1994).

Students' cultural backgrounds influence their beliefs about social roles in the classroom, the purpose of schooling, how to use language in the learning process (Bransford et al., 2000;

**In formative assessment, issues of reliability and validity are addressed over time, as teachers collect ongoing data about student performance and, as appropriate, make corrections to their previous inferences.**

Goldenberg & Gallimore, 1995; Greenfield, Suzuki, & Rothstein-Fisch, 2006; Greenfield, Trumbull et al., 2006), and how and when to demonstrate one's knowledge and understanding—an area entirely germane to formative assessment (Nelson-Barber & Trumbull, 2007). For example, when asked to name birds that live in a particular habitat, children from some cultural backgrounds may respond with stories of family outings that include sightings of birds, rather than the expected scientific discourse that focuses on observations or facts about birds, abstracted from experience or instruction (Trumbull, Diaz-Meza, & Hasan, 1999). This is because, in these students' homes, social and cognitive learning are seen as integrated, not separable. In such a case, if the teacher responds by discouraging personal stories, he or she may suppress students' participation in discussion. Instead, the teacher can demonstrate to students how to extract from their experiences what they have **■**

**Teachers must make any act of formative assessment contingent on what has been taught and on how students have responded to the teaching, and they must shape modifications to instruction in ways that make sense for students at different developmental levels within particular domains of study.**

learned about birds and record it on the board or in their journals (Trumbull et al., 1999).

Informal questioning, the most common form of formative assessment, may not always be the best way to assess students who are still learning English or who are from communities where English is used in ways different from those expected in the classroom. Such students may misconstrue the intent of a teacher's question, may thus respond differently than expected, and may then be misjudged about their understanding (Trumbull & Solano-Flores, 2011). A young student still learning the protocol of the classroom may think, "If the teacher already knows the answer to the question, why is she asking it? Is this a trick?"

Students from many cultural backgrounds (e.g., Asian, Latino,

Native American cultures) may avoid answering direct questions in a group of peers because being singled out in front of others is not common in their cultures and may cause discomfort or confusion (Greenfield & Cocking, 1994). Among Native American groups, a "right/wrong" approach to knowledge is not culturally congruent: Many such students have been socialized to consider all sides to a question and to avoid dichotomous (e.g., right/wrong) thinking. Historically, in federal boarding schools and (more recently) in classrooms using scripted basic skills programs, Native American students have suffered through direct questioning approaches, often responding with silence rather than participating in the question/answer ritual (McCarty, 2002). A recent study showed that teachers' oral questioning during discussions was negatively associated with Native American and Alaska Native students' later mathematics performance (Huang, Nelson-Barber, Trumbull, & Sexton, 2011). Likewise, an inquiry approach that requires students to reason aloud, on demand, about a question may be particularly uncomfortable for Native American and Alaska Native students who have learned at home to observe and mentally rehearse any complex task before attempting public performance (Swisher & Deyhle, 1995). Teachers not privy to the communication norms in some communities may at times be introducing non-target KSAs into assessment by using the very formative assessment practices that are most accepted (e.g., questioning students during a whole group discussion).

Cultural differences may also be associated with differences in responses to various forms of feedback (Hattie & Timperley, 2007; Kaplan, Karabenick, & De Groot, 2009; Maehr & Yamaguchi, 2001; Otsuka & Smith, 2005; Trumbull & Rothstein-Fisch, 2011). For example, some students may be uncomfortable with praise, particularly if it is given publicly (Markus & Kitayama, 1991; Rothstein-Fisch & Trumbull, 2008); they may be more motivated by negative feedback and criticism, at least in part because of a cultural value of working to meet the expectations of teachers and family (Heine, Takata, & Lehman, 2000). Teachers need to observe how their particular students respond to various forms of feedback in order to tailor feedback to those students' needs.

Given a basic understanding of how linguistic and cultural factors may intersect with formative assessment processes and tasks, educators can be alert to sources of non-target KSAs in order to achieve what has been called "cultural validity" in formative assessment (Solano-Flores & Nelson-Barber, 2001). Cultural validity is achieved when an assessment takes into consideration students' sociocultural backgrounds, including their cultural worldviews, their life contexts and values, the kinds of home and school experiences they have had (i.e., the foundation of their prior knowledge), their language preferences and proficiency, and their ways of using language to communicate and learn. Because formative assessment has the flexibility

to incorporate attention to context, it can more easily address issues of cultural validity.

### Research into the Effectiveness of Formative Assessment

Formative assessment has been highly touted for its purported positive impact on student learning (Black & Wiliam, 1998a; Organization for Economic Co-operation and Development, 2005). Black and Wiliam (1998), reviewing some 681 publications on studies related to formative assessment, concluded that “attention to formative assessment can lead to significant learning gains” (p. 9) and asserted that there is no evidence to suggest that it may have negative effects. However, caution should be exercised in making an uncritical endorsement of formative assessment (Bennett, 2011; Dunn & Mulvenon, 2009; Kingston & Nash, 2012a; Shepard, 2005). One issue is that the term “formative assessment” itself has been interpreted to mean different things (Bennett, 2011). For example, it may be used to describe commercial assessments that are not truly capable of serving a formative purpose because they are not tied closely enough to the teaching and learning context (Perie, Marion, Gong, & Wurtzel, 2007; Popham, 2006; Shepard, 2010).

Another issue is that the body of research on which claims of the positive impact of formative assessment are based is relatively small, and many of the relevant studies do not have the methodological

rigor to support conclusions about the effectiveness of formative assessment. Most claims about the benefits of formative assessment begin with the Black and Wiliam (1998a) review of research on formative assessment. Their review is often referred to as a “meta-analysis,” but, as the authors themselves observe, a true meta-analysis was not feasible for them because the studies they used represented such a wide range of practices and research methods. What the studies they reviewed had in common was teachers’ use of some of the features of formative assessment (e.g., feedback, teacher questioning, student self-assessment); these features were associated with moderate-to-large effect sizes. Bennett (2011) suggests that the characterization of Black and Wiliam’s review as a meta-analysis is education’s equivalent of an urban legend.

A recent meta-analysis of studies on the impact of formative assessment on K–12 student achievement concludes that, if only the studies hewing to rigorous methods are examined, the effect sizes of formative assessment are quite modest (a mean of .20); however, the effects are usually positive (of the 42 effect sizes reported, only 7 were negative), and some positive effects are greater than others (Kingston & Nash, 2012a). This meta-analysis has been criticized for the methods it employed, leading to a debate as to whether the findings were limited by the methodology (See Briggs, Ruiz-Primo, Furtak, Shepard, & Yuen, 2012; Kingston & Nash, 2012b).

In other summaries, implementation of particular formative assessment strategies that teachers had learned in professional development sessions resulted in an average effect size of .30 (Wiliam, Lee, Harrison, & Black, 2004), and use of a computer-based formative assessment system of writing resulted in an effect size of .28 (Rich, Harrington, Kim, & West, 2008).

There is some suggestion in the research literature as to why the effects of formative assessment are not as large as one might expect: Teachers are unsure what to do in response to what they learn about their students from formative assessment. The evidence gathered through formative assessment should be used to determine whether instruction needs to be modified and, if so, how. However, this part of the formative assessment cycle often falters: Teachers may succeed in gathering evidence about student learning and may accurately interpret the evidence to identify what knowledge a student lacks, yet may not be able to identify, target, and carry out specific instructional steps to close the learning gaps (Heritage, et al., 2009; Herman et al., 2006).

### Conclusion

Formative assessment is not new. Though they may not have called it by that name, effective teachers have always probed, in the course of their instruction, to understand students’ thinking and learning. Through questioning and observation, among other activities, they have strived to ▀

see behind the curtain, to expose why and how their students might get stuck or go off track. These teachers have taken what they have learned about their students and used that knowledge, along with their knowledge of pedagogy and the subject of study, to provide actionable feedback to students and to tailor their teaching to meet students' learning needs.

Shavelson (1973) noted that "any teaching act is the result of a decision ... that the teacher makes after the complex cognitive processing of available information," and he argued that "what distinguishes the exceptional teacher from his or her colleagues is not the ability to ask, say, a higher-order question, but the ability to decide when to ask such a question. (p. 144)" That decision, according to Shavelson, would incorporate information about students' understanding of course material and how alternative teaching actions would affect students' understanding.

As educators and researchers have been examining how teachers use assessments to inform instruction, it has become clear that conducting formative assessment is not only a complex process but one that requires extensive knowledge, including knowledge about student learning, domains of study, assessment, and pedagogy. Teachers must make any act of formative assessment contingent on what has been taught and on how students have responded to the teaching, and they must shape modifications to instruction in ways that make sense for students at different developmental levels within particular domains of study. There

is no prescription for how to tailor formative assessment to meet the needs of a particular classroom or student, but this tailoring is what good teaching demands of teachers. Thus, the full burden of implementing formative assessment falls on the teacher.

While there are efforts to develop supports for teachers who want to use assessments formatively, there is much work to be done. Research into learning progressions—those cognitive models of knowledge development within specific domains—may eventually provide teachers with validated models that they can use to guide formative assessment. Professional development and coaching on formative assessment may advance teachers' skill in using assessment to provide feedback to students and to inform their own instruction; advances in technology may help teachers meet the challenges of tailoring assessment and instruction to individual students. And the growing demand for balanced assessment systems presents both a rationale and an opportunity for the field to refocus some of the attention that is currently given to assessment onto classrooms and the important activities of teachers and students working to promote learning.

## References

- Abedi, J. (2006). Language issues in item-development. In S. M. Downing & T. M. Haladyna (Eds.), *Handbook of test development* (pp. 377–398). Mahwah, NJ: Erlbaum.
- Abedi, J. (2010). Research and recommendations for formative assessment with English language

learners. In H. Andrade & C. Cizek (Eds.), *Handbook of formative assessment* (pp. 181–197). New York: Routledge.

Abedi, J. (2011). Assessing English language learners. In M. Basterra, E. Trumbull, & G. Solano-Flores (Eds.), *Cultural validity in assessment: Addressing linguistic and cultural diversity* (pp. 49–71). New York: Routledge.

Andrade, H. (2010). Students as the definitive source of formative assessment. In H. Andrade & C. Cizek (Eds.), *Handbook of formative assessment* (pp. 90–105). New York: Routledge.

Ash, D., & Levitt, K. (2003). Working within the zone of proximal development: Formative assessment as professional development. *Journal of Science Teacher Education*, *14*(1), 1–26.

Bailey, F., Burkett, B., & Freeman, D. (2008). The mediating role of language in teaching and learning: A classroom perspective. In B. Spolsky & F. M. Hult (Eds.), *Handbook of educational linguistics* (pp. 606–625). Malden, MA: Blackwell.

Barrett, J. E., Sarama, J., Clements, D. H., Cullen, C., McCool, J., Witkowski-Rumsey, C., & Klanderma, D. (2012). Evaluating and improving a learning trajectory for linear measurement in elementary grades 2 and 3: A longitudinal study. *Mathematical Thinking and Learning*, *14*(1), 28–54.

Basterra, M. (2011). Cognition, culture, language, and assessment: How to select culturally valid assessments in the classroom. In M. Basterra, E. Trumbull, & G. Solano-Flores (Eds.), *Cultural validity in assessment: Addressing linguistic and cultural diversity* (pp. 72–95). New York: Routledge.

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, 18(1), 5–25.

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

Black, P., & Wiliam, D. (1998b). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139–149.

Black, P., & Wiliam, D. (2004). The formative purpose: Assessment must first promote learning. In M. Wilson (Ed.), *Towards coherence between classroom assessment and accountability*. The 103rd yearbook of the National Society for the Study of Education, Part II (pp. 20–50). University of Chicago Press.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation, and Accountability*, 21(1), 5–31.

Bloom, B. S., Hastings, J. T., & Madaus, G. F. (1971). *Handbook on formative and summative evaluation of student learning*. New York: McGraw-Hill.

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

Briggs, D. C., Ruiz-Primo, M. A., Furtak, E., Shepard, L., & Yuen, Y. (2012). Meta-analytic methodology and inferences about the efficacy of formative assessment. *Educational Measurement: Issues and Practice*, 31(4), 13–17.

Bronkard, J.-P. (1995). Theories of action, speech, natural language,

and discourse. In J. V. Wertsch, P. del Rio, & A. Alvarez (Eds.), *Sociocultural studies of mind* (pp. 75–91). Cambridge, England: Cambridge University Press.

Brookhart, S. M. (2003). Developing measurement theory for classroom assessment purposes and uses. *Educational Measurement Issues and Practices*, 22(4), 5–12.

Brookhart, S. M., (2005, April). *Research on formative assessment classroom assessment*. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.

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.

Brown, A. L., & Campione, J. C. (1994). Guided discovery in a community of learners. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 229–270). Cambridge, MA: MIT Press.

Bruner, J. (1985). *Actual minds, possible worlds*. Cambridge, MA: Harvard University Press.

Clark, I. (2012). Formative assessment: A systematic and artistic process of instruction for supporting school and lifelong learning. *Canadian Journal of Education*, 35(2), 24–40.

Corcoran, T., Mosher, F. A., & Rogat, A. (2009). *Learning progressions in science: An evidence-based approach to reform* (CPRE Research Report #RR-63). New York: Center on Continuous Instructional Improvement, Teachers College—Columbia University.

Darling-Hammond, L., & Pecheone, R. (2010). *Developing an internationally comparable balanced assessment system that supports high-quality learning*. Princeton, NJ: Educational Testing Services.

Demmert, W. (2005). The influences of culture on learning and assessment among Native American students. *Learning Disabilities Research & Practice*, 20(1), 16–23.

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.

Dunn, K. E., & Mulvenon, S. (2009). A critical review of research on formative assessment: The limited scientific evidence of the impact of formative assessment in education. *Practical Assessment, Research and Evaluation*, 14(7). Retrieved from <http://www.pareonline.net/pdf/v14n7.pdf>

Durán, R. P. (1985). Influences of language skills on bilinguals' problem solving. In S. F. Chipman, J. W. Segal, & R. Glaser (Eds.), *Thinking and learning skills* (pp. 187–207). Hillsdale, NJ: Lawrence Erlbaum Associates.

Durán, R. P. (2011). Ensuring valid educational assessments for ELL students. Scores, score interpretation, and assessment uses. In M. Bastera, E. Trumbull, & G. Solano-Flores (Eds.), *Cultural validity in assessment: Addressing linguistic and cultural diversity* (pp. 115–142). New York: Routledge.

Emihovich, C. (1994). The language of testing: An ethnographic sociolinguistic perspective on standardized tests. In K. Holland, D. Bloome, & J. Solsken (Eds.), *Alternative perspectives in assessing children's* ▀

language and literacy (pp. 33–52). Norwood, NJ: Ablex.

Erickson, F. (2007). Some thoughts on “proximal” formative assessment of student learning. In P. A. Moss (Ed.), *Evidence and decision making*. The 106th Yearbook of the National Society for the Study of Education, 106(1), 186–216. Malden, MA: Blackwell Publishing.

Figueroa, R. A., & Newsome, P. (2006). The diagnosis of LD in English language learners: Is it non-discriminatory? *Journal of Learning Disabilities*, 39(3), 206–214.

Frohbeiter, G., Greenwald, E., Stecher, B., & Schwartz, H. (2011). *Knowing and doing: What teachers learn from formative assessment and how they use information*. CRESST Report 802. Los Angeles: UCLA National Center for Research on Evaluation, Standards, and Student Testing.

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.

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.

Goldenberg, C., & Gallimore, R. (1995). Immigrant Latino parents’ values and beliefs about their children’s education. Continuities and discontinuities across cultures and generations. In P. Pintrich & M. Maehr (Eds.), *Advances in achievement motivation* (pp. 183–228). Greenwich, CT: JAI Press.

Greenfield, P. M. (2009). Linking social change and developmental change: Shifting pathways of human development. *Developmental Psychology*, 45(2), 401–418.

Greenfield, P. M., & Cocking, R. R. (Eds.). (1994). *Cross-cultural roots of minority child development*. Mahwah, NJ: Lawrence Erlbaum Associates.

Greenfield, P. M., Quiroz, B., & Raeff, C. (2000). Cross-cultural conflict and harmony in the social construction of the child. In S. Harkness, C. Raeff, & C. M. Super (Eds.), *Variability in the social construction of the child: New directions for child and adolescent development*, 87, 93–108.

Greenfield, P. M., Suzuki, L., & Rothstein-Fisch, C. (2006). Cultural pathways in human development. In W. Damon (Series Ed.) & I. E. Sigel & K. Renninger (Vol. Eds.), *Handbook of child psychology* (6th ed.), Vol. 4: *Child psychology in practice* (pp. 655–699). New York: Wiley.

Greenfield, P. M., Trumbull, E., Keller, H., Rothstein-Fisch, C., Suzuki, L., Quiroz, B., & Maynard, A. (2006). Cultural conceptions of learning and development. In P. Winne (Ed.), *Handbook of educational psychology* (2nd ed.) (pp. 675–692). Washington, DC: American Psychological Association.

Guskey, T. R. (2010). Formative assessment: The contributions of Benjamin S. Bloom. In H. Andrade & C. Cizek (Eds.), *Handbook of formative assessment* (pp. 106–124). New York: Routledge.

Haertel, E. H. (2006) Reliability. In R. L. Brennan (Ed.), *Educational measurement* (4th ed.) (pp. 65–110). Westport, CT: Praeger Publishers.

Harris, K., Bauer, M., & Redman, M. (2008). *Cognitive based developmental*

*models used as a link between formative and summative assessment*. Princeton, NJ: Educational Testing Service. Retrieved from [http://www.iaea2008.cambridgeassessment.org.uk/ca/digitalAssets/164898\\_Harris.pdf](http://www.iaea2008.cambridgeassessment.org.uk/ca/digitalAssets/164898_Harris.pdf)

Hattie, U., & Timperley, H. (2007). The power of feedback, *Review of Educational Research*, 77(1), 81–112.

Heath, S. B. (1983). *Ways with words: Language, life, and work in communities and classrooms*. Cambridge, England: Cambridge University Press.

Heine, S. J., Takata, T., & Lehman, D. R. (2000). Beyond self-presentation: Evidence for self-criticism among Japanese. *Personality and Social Psychology Bulletin*, 26, 71–78.

Heritage, M. (2008). *Learning progressions: Supporting instruction and formative assessment*. Paper prepared for the Formative Assessment for Teachers and Students (FAST) State Collaborative on Assessment and Student Standards (SCASS) of the Council of Chief State School Officers (CCSSO). Washington, DC: CCSSO.

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

Heritage, M. (2010b). *Formative assessment and next-generation assessment systems: Are we losing an opportunity?* Paper prepared for the Council of Chief State School Officers. Los Angeles: UCLA National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

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.

Herman, J. L. (2010). *Coherence: Key to next generation assessment success (AACC report)*. Los Angeles, CA: 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: UCLA National Center for Research on Evaluation, Standards, and Student Testing.

Hess, K. (2010, April). *Strategies for helping teachers make better use of assessment results. Best practices for state assessment systems. Workshop II*. National Research Council, Washington, DC.

Hoover, J. J., & Klingner, J. (2011). Promoting cultural validity in the assessment of bilingual special education students. In M. Basterra, E. Trumbull, & G. Solano-Flores (Eds.), *Cultural validity in assessment: Addressing linguistic and cultural diversity* (pp. 143-167). New York: Routledge.

Huang, C.-W., Nelson-Barber, S., Trumbull, E., & Sexton, U. (2011, April). *Student, teacher, and school factors associated with NAEP mathematics test performance by American Indian and Alaska Native students*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Kaplan, A., Karabenick, S., & De Groot, E. (2009). Introduction: Culture, self, and motivation: The contribution of Martin L. Maehr to the fields of achievement motivation and educational psychology. In A. Kaplan, S. Karabenick, & E. De Groot (Eds.), *Culture, self, and motivation: Essays in honor of Martin L. Maehr* (pp. vii-xxi). Charlotte, NC: Information Age Publishing.

Kingston, N., & Nash, B. (2012a). Formative assessment: A meta-analysis and a call for research. *Educational Measurement: Issues and Practice*, 30(4), 28-37.

Kingston, N., & Nash, B. (2012b). How many formative angels can dance on the head of a meta-analytic pin: .2. *Educational Measurement: Issues and Practice*, 31(4), 18-19.

Kopriva, R., & Sexton, U. (2011). Using appropriate assessment processes: How to get accurate information about the academic knowledge and skills of English language learners. In M. Basterra, E. Trumbull, & G. Solano-Flores (Eds.), *Cultural validity in assessment: Addressing linguistic and cultural diversity* (pp. 96-114). New York: Routledge.

Koschmann, T. (1999, December). Toward a dialogic theory of learning: Bakhtin's contribution to understanding learning in settings of collaboration. In C. Hoadley & J. Roschelle (Eds.), *Proceedings of the Computer Support for Collaborative Learning (CSCL) 1999 conference* (pp. 308-313). Mahwah, NJ: Lawrence Erlbaum Associates. Retrieved from <http://www.gerrystahl.net/proceedings/cscl1999/A38/A38.HTM>

Lager, C. A. (2006). Types of mathematics-language reading interactions that unnecessarily hinder algebra learning and assessment. *Reading Psychology*, 27(2-3), 165-204.

Leahy, S., & Wiliam, D. (2011, April). *Devising learning progressions*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans: LA.

Lee, O., Santau, A., & Maerten-Rivera, J. (2011). Science and literacy assessments with English language learners. In M. Basterra, E. Trumbull, & G. Solano-Flores (Eds.), *Cultural validity in assessment: Addressing linguistic and cultural diversity* (pp. 254-274). New York: Routledge.

Lerner, J. W. (2000). *Learning disabilities: theories, diagnosis, and teaching strategies*. Boston: Houghton Mifflin.

Li, M., Solano-Flores-G., Kwon, M., & Tsai, S. P. (2008, April). "It's asking me as if I were the mother." *Examining how students from different groups interpret test questions*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Baltimore, MD.

Maehr, M. L., & Yamaguchi, R. (2001). Cultural diversity, student motivation, and achievement. In F. Salili, C. Chiu, & Y. Hong (Eds.), *Student motivation: The culture and context of learning* (pp. 121-148). New York: Kluwer Academic/Plenum.

Markus, H., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98, 224-253.

McCarty, T. L. (2002). *A place to be Navajo*. New York: Routledge.

McManus, S. (2008). *Attributes of effective formative assessment*. Washington, DC: Council of Chief State School Officers. Retrieved from <http://www.ccsso.org>

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational* ▀

measurement (3rd ed.) (pp. 13–103). New York: American Council on Education and Macmillan.

Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2003). On the structure of educational assessments. *CSE Technical Report 597*. Los Angeles: UCLA National Center for Research on Evaluation, Standards, and Student Testing.

Mosher, F. A. (2011). The role of learning progressions in standards-based education reform. *CPRE Policy Briefs RB-52*. Philadelphia: University of Pennsylvania, Consortium for Policy Research in Education. Available at <http://www.cpre.org>.

Moss, P. (2008). Sociocultural implications for the practice of assessment I: Classroom assessment. In P. A. Moss, D. C. Pullin, J. P. Gee, E. H. Haertel, & L. J. Young (Eds.), *Assessment, equity, and opportunity to learn* (pp. 222–258). New York: Cambridge University Press.

National Joint Committee on Learning Disabilities. (2010). *Comprehensive assessment and evaluation of students with learning disabilities*. Retrieved from <http://www.ldanatl.org>

Nelson-Barber, S., & Trumbull, E. (2007). Making assessment practices valid for Native American students. *Journal of Indian Education*, Special Issue.

Ochs, E., & Schieffelin, B. (2011). The theory of language socialization. In A. Duranti, E. Ochs, & B. Schieffelin (Eds.), *The handbook of language socialization* (pp. 1–21). Malden, MA: Wiley-Blackwell.

Organization for Economic Co-operation and Development (OECD). (2005). *Formative assessment: Improving learning in secondary classrooms*. [Policy Brief].

Retrieved from <http://www.oecd.org/edu/ceri/35661078.pdf>

Otsuka, S., & Smith, I. D. (2005). Educational applications of the expectancy-value model of achievement motivation in the diverse cultural contexts of the West and the East. *Change: Transformations in Education*, 8(1), 91–109.

Pellegrino, J. W. (2006, November). *Rethinking and redesigning curriculum, instruction, and assessment: What contemporary research and theory suggests*. Paper commissioned by the National Center on Education and the Economy for the New Commission on the Skills of the American Workforce. Washington, DC: National Center on Education and the Economy.

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.

Perie, M., Marion, S., Gong, B., & Wurtzel, J. (2007). *The role of interim assessment in a comprehensive system: A policy brief*. Aspen, CO: The Aspen Institute.

Perrenoud, P. (1991). Towards a pragmatic approach to formative evaluation. In P. Weston (Ed.), *Assessment of pupils' achievements: Motivation and school success* (pp. 79–101). Amsterdam: Swets and Zeitlinger.

Phelan, J., Kang, T., Niemi, D. N., Vendilinski, T., & Choi, K. (2009). *Some aspects of the technical quality of formative assessments in middle school mathematics*. CRESST Report 750. Los Angeles: UCLA National Center for Research on Evaluation, Standards, and Student Testing.

Piaget, J. (1954). *The construction of reality in the child*. New York: Basic Books.

Popham, W. J. (2006). All about accountability/phony formative assessments. Buyer beware! *Educational Leadership* 64(3), 86–87.

Pryor, J., & Crossouard, B. (2005, September). *A sociocultural theorization of formative assessment*. Paper presented at the Sociocultural Theory in Educational Research and Practice Conference, Manchester, England.

Rich, C. S., Harrington, H., Kim, J., & West, B. (2008, March). *Automated essay scoring in state formative and summative writing assessment*. Paper presented at the annual meeting of the American Educational Research Association, New York.

Rogoff, B. (1994). Developing understanding of the idea of communities of learners. *Mind, Culture, and Activity*, 1(4), 209–229.

Rogoff, B. (1998). Cognition as a collaborative process. In W. Damon, D. Kuhn, & R. S. Siegler (Eds.), *Handbook of child psychology: Cognition, perception, and language* (5th ed.) (pp. 679–744). New York: Wiley.

Rogoff, B. (2003). *The cultural nature of human development*. New York: Oxford University Press.

Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment, Research & Evaluation*, 11(10). Retrieved from <http://pareonline.net/pdf/v11n10.pdf>

Rothstein-Fisch, C., & Trumbull, E. (2008). *Managing diverse classrooms: How to build on students' cultural strengths*. Alexandria, VA: Association for Supervision and Curriculum Development.

Ruiz-Primo, M. A. (2011). Informal formative assessment: The role of instructional dialogues in assessing students' learning. *Studies in Educational Evaluation*, 37, 15–24.

Ruiz-Primo, M. A., & Li, M. (2011, April). *Looking into teachers' feedback practices: How teachers interpret students' work*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Ruiz-Primo, M. A., & Li, M. (2013). Examining formative feedback in the classroom context: New research perspectives. In J. H. McMillan (Ed.), *Handbook of research on classroom assessment* (pp. 215–234). Los Angeles: Sage.

Ruiz-Primo, M. A., Shavelson, R. J., Hamilton, L., & Klein, S. (2002). On the evaluation of systemic science education reform: Searching for instructional sensitivity. *Journal of Research in Science Teaching*, 39(5), 369–393.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Sciences*, 18, 119–144.

Schleppegrell, M. J. (2004). *The language of schooling: A functional linguistics perspective*. Mahwah, NJ: Erlbaum.

Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R. M. Gagne, & M. Scriven (Eds.), *Perspectives of curriculum evaluation* (AERA Monograph Series on Curriculum Evaluation, No. 1) (pp. 39–83). Chicago: Rand McNally.

Shavelson, R. J. (1973). What is the basic teaching skill? *Journal of Teacher Education*, 24(2), 144–151.

Shavelson, R. J., Black, P. J., William, D., & Coffey, J. (2007). *On linking*

*formative and summative functions in the design of large-scale assessment systems*. Report from the Stanford Education Assessment Lab. Stanford, CA: Stanford Graduate School of Education. Retrieved from <http://www.stanford.edu/dept/SUSE/SEAL/>

Shavelson, R. J., & Kurpius, A. (2012). Reflections on learning progressions. In A. C. Alonzo & A. W. Gotwals (Eds.), *Learning progressions in science* (pp. 13–26). Rotterdam, The Netherlands: Sense Publishers.

Shavelson, R. J., Yin, Y., Furtak, E. M., Ruiz-Primo, M. A., Ayala, C. C., Young, D. B., ... Pottenger, F. M., III. (2008). On the role and impact of formative assessment on science inquiry teaching and learning. In J. Coffey, R. Douglas, & C. Stearns (Eds.), *Assessing science learning* (pp. 21–36). Arlington, VA: NSTA Press.

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

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

Shepard, L. A. (2006). Classroom assessment. In R. L. Brennan (Ed.), *Educational Measurement (4th ed.)* (pp. 623–645). Westport, CT: Praeger Publishers.

Shepard, L. A. (2009). Commentary: Evaluating the validity of formative and interim assessment. *Educational Measurement: Issues and Practice*, 28(3), 32–37.

Shepard, L. A. (2010). What the marketplace has brought us:

Item-by-item teaching with little instructional insight. *Peabody Journal of Education*, 85(2), 246–257.

Solano-Flores, G., & Nelson-Barber, S. (2001). On the cultural validity of science assessments. *Journal of Research in Science Teaching*, 38(5), 553–573.

Solano-Flores, G., & Trumbull, E. (2003). Examining language in context: The need for new research and practice paradigms in the testing of English language learners. *Educational Researcher*, 32(2), 3–13.

Spinelli, C. G. (2008). Addressing the issue of cultural and linguistic diversity and assessment: Informal evaluation measures for English language learners. *Reading & Writing Quarterly*, 24, 101–118.

Steedle, J. T., & Shavelson, R. J. (2009). Supporting valid interpretations of learning progression level diagnoses. *Journal of Research in Science Teaching*, 46, 699–715.

Stiggins, R. J., Arter, J. A., Chappuis, J., & Chappuis, S. (2009). *Classroom assessment for student learning: Doing it right—Using it well*. Portland, OR: Assessment Training Institute.

Supovitz, J. (2012). Getting at student understanding—the key to teachers' use of test data. *Teachers College Record*, 114(11), 1–29.

Swisher, K., & Deyhle, D. (1992). Adapting instruction to culture. In J. Reyhner (Ed.), *Teaching American Indian students* (pp. 81–95). Norman, OK: University of Oklahoma.

Sztajn, P., Confrey, J., Wilson, P. H., & Edgington, C. (2012). Learning trajectory based instruction: Toward a theory of teaching. *Educational Researcher*, 41(5), 147–156. ♣

Tharp, R., & Gallimore, R. (1991). *The instructional conversation: Teaching and learning in social activity* (Research Report 2). Santa Cruz, CA: The National Center for Research on Cultural Diversity and Second Language Learning, University of California, Santa Cruz.

Topping, K. J. (2010). Peers as a source of formative assessment. In H. L. Andrade and G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 61–74). New York: Routledge.

Troia, G. A. (2011). How might pragmatic language skills affect the written expression of students with language learning disabilities? *Topics in Language Disorders, 31*(1), 40–53.

Trumbull, E., & Rothstein-Fisch, C. (2011). The intersection of culture and achievement motivation. *The School Community Journal, 21*(2), 25–53.

Trumbull, E., & Solano-Flores, G. (2011). The role of language in assessment. In M. Basterra, E. Trumbull, & G. Solano-Flores (Eds.), *Cultural validity in assessment: Addressing linguistic and cultural diversity* (pp. 22–45). New York: Routledge.

Tunstall, P., & Gipps, C. (1996). Teacher feedback to young children in formative assessment: A typology. *British Educational Research Journal, 22*, 389–404.

Vygotsky, L. S. (1962). *Thought and language*. Cambridge, MA: Massachusetts Institute of Technology.

Vygotsky, L. S. (1978). *Mind and society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Walqui, A., & van Lier, L. (2010). *Scaffolding the academic success of adolescent English language learners*. 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>

William, D. (2006). *Formative assessment: Getting the focus right*. Retrieved from [http://eprints.ioe.ac.uk/1128/1/EAJ\\_FA\\_special\\_issue\\_commentary\\_v2.pdf](http://eprints.ioe.ac.uk/1128/1/EAJ_FA_special_issue_commentary_v2.pdf)

William, D. (2010). An integrative summary of the research literature and implications for a new

theory of formative assessment. In H. L. Andrade and G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 18–40). New York: Routledge.

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

Wilson, M., & Draney, K. (2004). Some links between large-scale and classroom assessments: The case of the BEAR Assessment System. In M. Wilson (Ed.), *Toward coherence between classroom assessment and accountability*. The 103rd Yearbook of the National Society for the Study of Education, Part II (pp. 132–154). University of Chicago Press.

Wilson, M. R., & Bertenthal, M. W. (2006). Executive summary. In M. R. Wilson & M. W. Bertenthal (Eds.), *Systems for state science assessment* (pp. 1–10). Washington, DC: National Research Council, National Academies Press.

Wong Fillmore, L., & Snow, C. E. (2000). *What teachers need to know about language*. Paper prepared for the U. S. Department of Education, Office of Educational Research and Improvement. Washington, DC: Center for Applied Linguistics.

Zhang, T., Mislavy, M. J., Haertel, G., Javitz, H., Murray, E., Gravel, J., & Hansen, E. G. (2010). *A design pattern for a spelling assessment for students with disabilities*. Assessment for Students with Disabilities Technical Report 2. Menlo Park, CA: SRI International.

©2013 WestEd. All rights reserved.

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



WestEd — a national nonpartisan, nonprofit research, development, and service agency — works with education and other communities to promote excellence, achieve equity, and improve learning for children, youth, and adults. WestEd has 16 offices nationwide, from Washington and Boston to Arizona and California, with its headquarters in San Francisco. For more information about WestEd, visit [WestEd.org](http://WestEd.org); call 415.565.3000, or toll-free (877)4-WestEd; or write: WestEd | 730 Harrison Street | San Francisco, California 94107-1242.