Assessment *for* Learning

What Policymakers Should Know About Formative Assessment

by Martin Orland and Janice Anderson

Forty-five states across the United States, along with Washington, D.C., and a number of U.S. territories, have committed to revamping their education systems through reforms designed to better prepare American students for meaningful participation in postsecondary education and the world of work (Common Core State Standards Initiative). Serving as the keystone for their efforts are challenging new standards—starting with the Common Core State Standards (CCSS, 2012) and including the upcoming Next Generation Science Standards—that demand more-effective teaching and learning in U.S. classrooms, including interactions around subject matter that delve deeper, that are more focused, that encourage students to define and argue their own conclusions, and that support application of, and generalization beyond, concepts (Moschkovich, 2012; Bunch, Kibler, & Pimentel, 2012). As states prepare to implement the new standards and grapple with how to create new assessment systems aligned with them, a window of opportunity to revisit the role of assessment in student learning has opened.

For too long, too many have viewed assessment primarily as a means of measuring student learning for accountability purposes (Gordon Commission on the Future of Assessment in Education, 2013). But in recent years, policymakers, educators, researchers, and even parents have begun calling for "balanced" assessment systems in which assessments are used not just to measure student learning and inform accountability decisions but also to advance learning by guiding classroom instruction (Stiggins, 2008; Rabinowitz, 2010; Gordon Commission on the Future of Assessment in Education, 2013; Bogira, 2013).

Achieving this balance requires addressing two distinct but interrelated assessment system goals. The first is to produce valid and reliable assessments to measure what students *have* learned, for purposes of accountability at the state, district, school, teacher, and student levels. Assessment system components for this purpose are labeled *summative* because they provide summary measures of what students know and can do at particular points in their education careers, such as at the end of each course or school year.

The second goal is to improve the quality and use of assessments to elucidate what and how students *are* learning. Assessment system components for this purpose are labeled *formative*. They represent a type

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

of assessment that directly supports instruction by generating information, at multiple points during the course of instruction, about student thinking, about if and how students are learning, and about what misunderstandings or misconceptions might be getting in their way or sidetracking their learning.

Unlike summative assessments, formative assessments are intended to have direct and immediate applicability for classroom teachers and their students in improving teaching and learning.[1] Based on what teachers find out through formative assessments, they can reshape instruction and provide actionable feedback for students to use to enhance their learning (Trumbull & Lash, 2013).

Few dispute that there is an important role for summative assessment in American education. So given the significant technical, resource, and administrative challenges related to producing valid and reliable summative assessments for the CCSS, it is no surprise that individual states, as well as the two federally funded multistate consortia charged with developing next-generation assessments (i.e.,the Partnership for Assessment of Readiness for College and Careers and the Smarter Balanced Assessment Consortium), have been focusing much of their early attention on the accountability side of assessment systems.[2] However, there is a danger that policymakers will lose sight of what they can do to promote and improve the process of formative assessment, the component of education reform with the central objective of assessment *for* learning rather than *of* learning.

The purpose of this policy brief is to highlight the formative component of next-generation assessment systems and suggest possible next steps for state and federal policymakers.

What Formative Assessment Is—and a Little About What It Isn't

Formative assessment is defined by its purpose: to help shape a student's learning by eliciting evidence of learning, interpreting this evidence in terms of a student's learning needs, and using this interpretation to adjust instruction to meet those learning needs (Black & Wiliam, 1998).

Formative assessment in the classroom ranges from informal to formal, from responsive to planned, and from brief to extended; it can be focused on an individual student or an entire classroom (Shavelson, 2009). An example of an informal, responsive, and brief assessment focused on one student is when, to address an apparent misconception evidenced by a student comment during class discussion, a teacher asks the student a question designed to identify the source of his or her misunderstanding.

But if formative assessment can be as simple, spontaneous, and seemingly straightforward as a probing question, it can also be as complex, carefully planned, and leading-edge as a computer program that facilitates the use of curriculum-embedded classroom assessments to measure the depths of student knowledge and provide immediate diagnosis and feedback for teachers and students alike. Consider the middle school science classroom evoked below (Quellmalz, 2013):

> As part of an ecosystems unit, students are engaged in an online simulation-based assessment. To find out how well her students understand

[1] Besides formative and summative components, some visions of a balanced assessment system also include interim/benchmark assessments, which are periodic measures of student progress toward achieving performance objectives. If assessment practices and tools are conceived of as being on a continuum from formative to summative, then interim/benchmark assessments fall somewhere in the middle of that continuum, with the exact placement of a given interim/benchmark assessment task dependent on its form and purpose.

[2] While assessment for accountability has clearly been the main focus of CCSS-related assessment work, it should be noted that both states and the consortia are aware of the need to increase and improve the use of formative assessment in the classroom. The consortia have committed to making CCSS-related formative-assessment-support resources available online as part of their next generation assessment systems.

the flow of energy and matter through ecosystems, the teacher has asked them to develop a description of a mountain lake ecosystem for a hypothetical park visitors' center. Working in pairs on classroom laptops, students start by observing an online animation of underwater creatures found in a mountain lake—e.g., algae, shrimp, alewife, and trout. Then, based on prior instruction about the roles of consumers and producers in ecosystems, and based on their observations of the dynamic interactions between the mountain lake organisms, students are to create a food web diagram on-screen, drawing arrows to show how energy and matter are transferred between organisms throughout the ecosystem.

When one student draws an on-screen arrow showing that a predator is eating the wrong prey, a feedback message pops up on the screen indicating that the arrow is incorrect and suggesting that the student take another look at the animation to identify the source of the matter and energy for that predator. Another student gets a similar message after mixing up predator and prey and drawing an arrow that points from the food source to the predator. The teacher watches and listens carefully, providing feedback of her own as students observe and discuss the organisms' interactions.

What Formative Assessment Is and Is Not

Formative assessment is not ...	Rather, it is ...
A specific instrument or a test	A process that may use a variety of possible instruments to elicit actionable data about performance
A follow-up to an instructional unit	Part of the instructional unit
A unidirectional approach, from teacher to student	A shared responsibility
A final score report of success or failure	A process that evokes evidence-based judgment and action

(Heritage. 2010; Sadler, 1989; Centre for Educational Research and Innovation, 2008; Shepard, 2005; Pellegrino, 2010)

At the end of the simulation, the program generates a progress report identifying the concepts and science practices for which each student is "on track," "is making progress," or "needs help." The reports help the teacher decide how to differentiate instruction on the following day when students will work offline in small groups to practice analyzing three new ecosystems.

In this example of next-generation formative assessment, students are absorbed in an engaging, dynamic, collaborative online learning and assessment environment that exposes their thinking to each other and to their teacher. In response to evidence of students' understanding—or misunderstanding—both the simulation itself and the teacher provide immediate, individualized feedback and gradualized coaching. The teacher uses the progress report to plan how to adjust additional instruction for the next activity, and during small-group activities, students receive and offer feedback about their own and each other's learning. Throughout the process, the students and their teacher are assessing their own performance (Heritage, 2010; Trumbull & Gerzon, 2013; Noyce, 2011). Leading educators from the United States and other developed nations around the world have advocated for this kind of process to become a part of everyday practice in the classroom (Centre for Educational Research and Innovation, 2008; Heritage, 2010).

Certain classroom, teacher, and student characteristics are essential for effective formative assessment. There must be a positive culture in the classroom, with students made to feel safe to engage with the teacher and fellow students, and with interactions focused on collaborative learning, not on competition among peers (Centre

Positive outcomes from the use of formative assessment have been documented for specific subgroups, including English language learners, learning-challenged students, and adults.

for Educational Research and Innovation, 2008; Quellmalz, 2013; Black & Wiliam, 1998). A teacher must be able to cultivate this kind of classroom environment. He or she must also possess deep content knowledge (including about best ways to teach the content to students of differing ability levels), be skilled in designing and using learning experiences that are aligned with standards, be a keen observer, be able to provide in-the-moment feedback, be able to adjust instruction based on student performance, and have a clear sense of desired learning outcomes (Bennett, 2011). Effective formative assessment also depends on the teacher's understanding of the likely trajectory of student learning over an extended period of time in a particular subject area (Heritage, 2010). Current efforts by researchers and educators to formally define and disseminate to the field these expected trajectories, or "learning progressions," for particular subjects offer promise for more-effective formative assessment practices in classrooms.

For their part, students must generate ideas, be willing to make mistakes, accept input and feedback from others (teachers and peers), take responsibility for their own learning, and be committed to helping peers progress toward desired learning outcomes (Trumbull & Lash, 2013)

In summary, formative assessment results in data-rich classrooms, as skillful teachers, through a variety of interactions with their students, continually elicit information about what students do and do not understand. Those interactions range from informal observations to "purposefully planned performance items embedded in instruction" (Heritage, 2010, p. 10; Centre for Educational Research and Innovation, 2008). Acting on what they learn from these interactions, teachers provide students with actionable feedback, enabling students to figure out their next steps. By linking the assessment process directly with the process of teaching and learning in classrooms, formative assessment adds the "balance" to a balanced assessment system in which different measures implemented at different times produce data for different participants in the educational system (students, parents, teachers, school-level administrators, district leaders, and state and national policymakers) so that they may all be more effective in their roles (Darling-Hammond & Pecheone, 2010; Heritage, 2010).

Evidence About How Formative Assessment Affects Learning

While research shows a general consensus that formative assessment practices have a positive impact on student learning, estimates of the size of that impact vary due to differences in how formative assessment has been defined in different research investigations, as well as differences in the methodological rigor of various impact studies and the methods used to evaluate and summarize their reported impacts.

Fifteen years ago, a landmark review of studies on the impact of formative assessment for K–12 students reported highly positive results (Black & Wiliams, 1998). More recently, a review of only those formative-assessment studies that employ rigorous methods found more modest, though still usually positive, formative assessment impacts on student learning (Kingston & Nash, 2012). Positive outcomes from the use of formative assessment have been documented for specific subgroups, including English language learners, learning-challenged students, and adults (Bunch, Kibler, & Pimentel, 2012; Nyquist, 2003; Quellmalz, Silberglitt, & Timms, 2011; Centre for Educational Research and Innovation, 2008;

Heritage, 2010). There is some evidence to suggest that, although teachers may generally succeed in using formative assessments effectively to identify specific student learning needs, they have more difficulty designing and implementing the appropriate instructional strategies for addressing those needs (Heritage, Kim, Vendlinski, & Herman, 2009; Herman, 2010).

One limitation of the current research base on formative assessment is that studies and reviews in this area are often about very diverse classroom practices. For example, one study might examine peer assessment, another the impacts of teacher questioning, and a third might examine teachers' use of tests for instructional decision-making. This diversity makes it difficult to draw conclusions about formative assessment as a coherent instructional intervention with a clear theory of action linked to a specific set of key strategies (e.g., student questioning, peer assessment) and associated instructional techniques (Bennett, 2011). To move the field forward, a future research agenda should include theory-based development, testing, and refinement of different combinations of formative assessment strategies and techniques. In addition, it should attempt to understand the effects of, at a minimum, specific teacher and student characteristics, technology and other supports, professional development approaches, and administrator behaviors on formative assessment implementation and impacts.

What Are Next Steps for Policymakers?

Formative assessment is an important and timely means for strengthening teaching and learning in the classroom, as called for by the CCSS-driven and other reform agendas. It complements summative assessments, which can identify students who lack the knowledge and skills expected for their grade level, by providing insights into *why* these students may not be proficient.

This policy brief posits that high-quality formative assessment can serve a critical systemic role in improving instruction and learning. As with so many promising education practices, however, the potential of formative assessment will only be realized with strategic and sustained policy leadership. Most importantly, policymakers need to see formative assessment in the classroom as a part of, rather than apart from, the next-generation state assessment systems being developed for implementation of the CCSS, and they must be mindful of the key differences between formative and summative assessment, so that they do not expect the same assessments to serve both purposes. Formative assessment should be considered an essential part of education improvement agendas in conjunction with other key components, which may include (in addition to the aforementioned balanced assessment systems) more rigorous teacher preparation, support, development, and evaluation systems; enhanced leadership development; and more comprehensive longitudinal data systems.

A growing rhetorical commitment to balanced assessment systems around the CCSS provides an unprecedented opportunity for policymakers to develop an integrated and coherent formative assessment improvement agenda with the goal of catalyzing high-quality formative assessment practices in classrooms nationwide.

Yet many policymakers do not appear to be thinking off formative assessment as an essential part of education; if they think of it at all, they tend to confuse it with interim/benchmark assessments of student growth, which are periodic measures of student progress toward achieving performance objectives. While interim/benchmark assessment may be able to provide some formative feedback to teachers

for student placement decisions and instructional priorities, they are *not* formative assessments because they fail to offer actionable real-time feedback that teachers and students can use during instruction to improve teaching and learning.

Policymakers also need to understand that, to be useful, formative assessment practices need to be of high quality. This requires teachers to be knowledgeable in their content areas and in pedagogy related to that content; to be well trained in the particular instructional techniques that are critical to high-quality formative assessment practices (e.g., setting clear learning targets, eliciting evidence of student learning, fostering student use of assessment information), and to have access to resources that support such practices.

There are currently isolated efforts aimed at supporting the use of formative assessment. For example, several teacher preparation programs have recently begun to develop required courses in formative assessment for their teacher candidates (Greenberg & Walsh, 2012). In addition, numerous professional development programs around the country have been training teachers in formative assessment practices, as well as training administrators to evaluate formative assessment tools and practices (Newman & Gullie, 2010; Wylie, 2008; Palucci, 2010). Charlotte Danielson's teacher observation framework (Danielson, 2011) is being adapted in many sites across the U.S. as a component of new state-mandated teacher evaluation systems, and it includes formative assessment practices as part its rubric for evaluating classroom practice. (The pilot rubric itself can be found online on several districts' websites. See, for example, Ann Arbor Public Schools, n.d.) Resources are currently available for teachers to construct their own performance tasks and classroom assessments for both formative and summative purposes, and both next-generation assessment consortia are committed to creating online libraries of formative assessment resources aligned with their assessment systems and with the CCSS.

These are all noteworthy developments, suggesting a growing recognition, within and beyond the academic community, of the promise of formative assessments to improve the education enterprise, as well as the field's increasing capacity to deliver on this promise. However, these efforts are also highly fragmented, reflecting the absence of a coherent policy agenda that sees assessment *for* learning as an integral part of systemic education reform, one that has comparable standing to relatively well-focused initiatives for assessment *of* learning. Such an agenda would include regulatory changes, such as new licensure and certification standards for teachers and administrators that would require knowledge of, and facility with, formative assessments for both entry and career advancement in the profession. It would also encompass a focused set of technical assistance and professional development offerings to build the capacity needed to engage in high-quality formative assessment practices. These offerings would include materials and services that demonstrate alignment and coherence among the different components of assessment systems—formative, interim/benchmark, and summative—as well as professional development in formative assessment that, rather than being generic, is specific to a teacher's relevant content area. Finally, this agenda would contain a robust research-and-development program. This program would act to validate learning progressions associated with the CCSS and other new standards, improve conceptual clarity about what specific instructional and assessment features are associated with effective formative assessments, and heighten the field's understanding of how these features can best operate in real-world settings to enhance student learning.

A growing rhetorical commitment to balanced assessment systems around the CCSS provides an unprecedented opportunity for policymakers to develop an integrated and coherent formative assessment improvement agenda with the goal of catalyzing high-quality

formative assessment practices in classrooms nationwide. The next decade will reveal whether this opportunity is realized. Policymakers will continue to have choices regarding how best to spend scarce resources and which policies to promote in order to improve student learning. Many currently think of policies and investments in assessment reform only in terms of improving accountability. By also considering how assessments can inform both teachers and their students about whether and how learning is taking place, and what to do to advance it, policymakers can play a critical role in ensuring that new assessment systems are as balanced as the rhetoric suggests.

References

Ann Arbor Public Schools. (n.d.). Danielson Pilot Rubric. Retrieved from http://www.aaps.k12.mi.us/aaps.forstaff/files/

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

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

Bogira, S. (2013, February 6). Chicago parents and teachers unite against standardized testing. *Chicago Reader.* Retrieved from http://www.chicagoreader.com/Bleader/archives/2013/02/06/chicago-parents-and-teachers-unite-against-standardized-testing

Bunch, G. C., Kibler, A., & Pimentel, S. (2012). *Realizing opportunities for English learners in the Common Core English language arts and disciplinary literary standards.* Palo Alto, CA: Understanding Language, Stanford University.

Center on Standards and Assessment Implementation & North Central Comprehensive Center (2013, February). *Formative Assessment Program webinar.* Retrieved from https://cc.readytalk.com/cc/s/meetingArchive?eventId=nm7p4mjpdtow&campaignId=ntg18hf6i6c1

Centre for Educational Research and Innovation. (2008). *Assessment for learning formative assessment*. Paris, France: Author.

Common Core State Standards Initiative. (2012). *In the States.* Retrieved from http://www.corestandards.org/in-the-states

Danielson, C. (2011,). Evaluations that help teachers learn. *The Effective Educator,* 68(4), 35–39.

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

Gordon Commission on the Future of Assessment in Education. (2013). *A public policy statement.* Princeton, NJ: Author. Retrieved from http://www.gordoncommission.org/publications_reports.html

Greenberg, J., & Walsh, K. (2012). *What teacher preparation programs teach about K–12 assessment: A review.* Washington, DC: National Council on Teacher Quality.

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

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 achievement success (*AACC Report). Los Angeles, CA: University of California.

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

Moschkovich, J. (2012). *Mathematics, the Common Core, and language: Recommendations for mathematics instruction for*

ELs aligned with the Common Core. Palo Alto, CA: Understanding Language, Stanford University.

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

Noyce, P. E. (2011). Introduction and overview. In P. E. Noyce & D. T. Hickey (Eds.), *New frontiers in formative assessment* (pp. 1–10). Cambridge, MA: Harvard Education Press.

Nyquist, J. B. (2003). *The benefits of reconstructing feedback as a larger system of formative assessment: A meta-analysis*. Unpublished master's thesis, Vanderbilt University, Nashville, TN.

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

Pellegrino, J. W. (2010). Technology and formative assessment. *International Encyclopedia of Education, 8*, 42–47.

Quellmalz, E. S. (2013). *Using technology to support classroom formative assessment*. San Francisco: Wested

Quellmalz, E. S., Silberglitt, M. D., & Timms, M. J. (2011). *How can simulations be components of balanced state science assessment systems?* San Francisco: WestEd.

Rabinowitz, S. N. (2010). Next-Generation Assessment Systems. *Education Week, 29*(22), 36.

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

Shavelson, R. (2009, June). *Reflections on learning progressions*. Paper presented at the Learning Progressions in Science (LeaPS) Conference, Iowa City, IA. Retrieved from http://education.msu.edu/projects/leaps/proceedings/Shavelson.pdf

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

Stiggins, R. J. (April 2008). *Assessment manifesto: A call for the development of balanced assessment systems*. Portland, OR: ETS Assessment Training Institute.

Trumbull, E., & Gerzon, N. (2013). *Professional development on formative assessment: Insights from research and practice*. San Francisco: WestEd.

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

Wylie, E. C. (2008). *Formative assessment: Examples of practice*. Washington, DC: Council of Chief State School Officers.

Suggested citation: Orland, M., & Anderson, J. (2013) *Assessment for learning: What policymakers should know about formative assessment.* San Francisco: WestEd.