

DISTINGUISHING FORMATIVE ASSESSMENT FROM OTHER EDUCATIONAL ASSESSMENT LABELS



FORMATIVE ASSESSMENT
FOR STUDENTS AND TEACHERS

PREPARED BY THE FORMATIVE ASSESSMENT
FOR STUDENTS AND TEACHERS (FAST) SCASS

MEMBERS | ARKANSAS, CONNECTICUT, HAWAII, ILLINOIS, IOWA,
KANSAS, KENTUCKY, MARYLAND, MICHIGAN, NORTH CAROLINA,
AND WEST VIRGINIA.

The FAST SCASS gratefully acknowledges the invaluable assistance of Jim Popham, Dylan William, and Margaret Heritage in the preparation of this paper.

STATE REPRESENTATIVES

Suzanne Knowles - *Arkansas*
Sherri Thorne - *Arkansas*
Patricia Foley - *Connecticut*
Joseph Di Garbo - *Connecticut*
Monica Mann - *Hawaii*
Rampal Singh - *Hawaii*
Gil Downey - *Illinois*
Rachel Jachino - *Illinois*
Colleen Anderson - *Iowa*
Michelle Hosp - *Iowa*
Jackie Lakin, *Kansas*
Jeannette Nobo - *Kansas*
Joy Barr - *Kentucky*
Sean Elkins - *Kentucky*
Denise Hunt - *Maryland*
Greg Sucre - *Maryland*
Ed Roeber - *Michigan*
Kimberly Young - *Michigan*
Carmella Fair - *North Carolina*
Sarah McManus - *North Carolina*
Stacey Murrell - *West Virginia*
Denise White - *West Virginia*
Sonya White - *West Virginia*

AFFILIATE MEMBERS

Michael Kaspar - *National Education Association*
Stuart Kahl - *Measured Progress*
Kelly Goodrich - *NWEA*
Sally Venezuela - *CTB/McGraw-Hill*

The FAST SCASS also thanks Debra Hawkins, Issaquah School District, WA and JohnHosp, University of Iowa for their support in the preparation of this paper.

Any opinions expressed in this paper are those of the FAST SCASS and not necessarily of CCSSO.

THE COUNCIL OF CHIEF STATE SCHOOL OFFICERS

The Council of Chief State School Officers (CCSSO) is a nonpartisan, nationwide, nonprofit organization of public officials who head departments of elementary and secondary education in the states, the District of Columbia, the Department of Defense Education Activity, and five U.S. extra-state jurisdictions. CCSSO provides leadership, advocacy, and technical assistance on major educational issues. The Council seeks member consensus on major educational issues and expresses their views to civic and professional organizations, federal agencies, Congress, and the public.

DISTINGUISHING FORMATIVE ASSESSMENT FROM OTHER EDUCATIONAL ASSESSMENT LABELS

COUNCIL OF CHIEF STATE SCHOOL OFFICERS

Thomas Luna (Idaho), President
Gene Wilhoit, Executive Director

Council of Chief State School Officers
One Massachusetts Avenue, NW, Suite 700
Washington, DC 20001-1431
Phone (202) 336-7000
Fax (202) 408-8072
www.ccsso.org

INTRODUCTION

In 2006, based on a close review of the literature on formative assessment and in consultation with national and international experts on the topic, the Council of Chief State School Officers' (CCSSO) Formative Assessment for Students and Teachers (FAST) State Collaborative on Assessment and Student Standards (SCASS) developed a now widely cited definition of formative assessment:

“Formative assessment is a process used by teachers and students during instruction that provides feedback to adjust ongoing teaching and learning to improve students’ achievements of intended instructional outcomes.”

Central to our view is that formative assessment takes place during the course of ongoing instruction to support student learning as it develops. This stands in contrast to summative assessment, or interim assessment, that is intended to evaluate or benchmark what students have achieved after a particular phase in their schooling — for example, after a course or a unit of study (NRC, 2001). We also stress student involvement in formative assessment through peer- and self-assessment, and the use of feedback by students to move their learning forward.

While there is general agreement among the FAST SCASS member states and other groups (e.g., the Assessment Reform Group in the U.K. and the International Network for Assessment for Learning) concerning the nature and purpose of formative assessment, the term itself is often used in different ways throughout the field of education. In this document, the FAST SCASS aims to clarify the meaning and uses of the types of assessment most frequently used in education. By so doing, the FAST SCASS intends to clarify what formative assessment is and is not in order to increase both the understanding and implementation of formative assessment practices in classrooms.

We have identified a modest collection of the labels that are currently used to describe various educational assessment types. Assessment, in our view, includes more than traditional paper-and-pencil testing, although paper-and-pencil tests do, indeed, represent one useful way for educators to arrive at inferences about students’ current knowledge and skills. In a context of formative assessment, evidence gathering may range from dialogic conversations that enable teachers to elicit student thinking, to student peer- or self-assessment, to the completion of elaborate, extended-duration tasks.

The FAST SCASS regards formative assessment practices as essential tools for teachers in supporting students to meet the rigorous Common Core State Standards, which emphasize higher levels of thinking for all students.

BACKGROUND ON ASSESSMENT

Any assessment is basically a process for making inferences about individuals or a group of individuals. Sometimes these inferences take the form of measurements — we want to be able to say that this student knows more third grade mathematics than that student. However, measuring the amount of knowledge of third grade mathematics possessed by a student is not as straightforward as measuring the weight of an object on a scale, measuring the length of a table with a ruler, or measuring air temperature by observing the expansion of mercury against a calibrated scale. In an assessment context, measurement is indirect — we cannot directly observe what is going on inside a student’s head (and it probably wouldn’t tell us much if we could!). We can only observe how a student responds to a series of questions, prompts, or tasks. We hypothesize that correct responses to these questions, prompts, and tasks require the possession of certain

knowledge, skills, or capabilities, so when one student does better than another, we infer that this is because they have more of the knowledge, skills, or capabilities we are most interested in.

At other times, however, the inferences may take the form of classifications rather than measurements. We may infer that one student has sufficiently mastered third grade mathematics to proceed to the fourth grade, while another has not. Although this inference could be based on a measurement, it could also be based on a careful comparison of the many things the student can do, and the things the student needs to be able to do to thrive in the fourth grade.

In any assessment, as in all measurement, there is a degree of uncertainty related to the accuracy of the method used for collecting evidence of learning and the way that evidence is used. One set of test items may be much better at eliciting evidence of student mastery of a particular concept than a different, but similar looking set. The outcome of any assessment also depends on a variety of other factors, such as how the student was feeling that day, the reliability of the rater, and so on.

Thus, assessment is about trying to understand what or how much is “in a student’s head.” A central component of formative assessment is helping teachers learn how to elicit such evidence so their insights into student thinking can be used to advance learning. Since we cannot measure directly, we ask questions that attempt to get at knowledge or skills in order to make reasonable inferences.

OUR DEFINITIONS

It is not our intent to provide a complete glossary of assessment terms. Rather, we offer a series of formative assessment terms and include a definition and brief commentary on each. The specific terms to be addressed in this paper are formative assessment, interim benchmark assessment, and summative assessment. In addition, several other educational assessment terms are defined: diagnostic assessment, curriculum-embedded assessment, universal screening assessment, and progress-monitoring assessment.

I. FORMATIVE ASSESSMENT

The FAST SCASS definition of formative assessment developed in 2006 is *“Formative assessment is a process used by teachers and students during instruction that provides feedback to adjust ongoing teaching and learning to improve students’ achievements of intended instructional outcomes.”*

The attributes below have been identified as critical features of effective formative assessment:

- **Learning Progressions.** Learning progressions should clearly articulate the sub-goals of the ultimate learning goal
- **Learning Goals and Criteria for Success.** Learning goals and criteria for success should be clearly identified and communicated to students

- **Evidence of Learning.** Evidence of learning is elicited during instruction
- **Descriptive Feedback.** Students should be provided with evidence-based feedback that is linked to the intended instructional outcomes and criteria for success
- **Self- and Peer-Assessment.** Both self- and peer-assessment are important for providing students an opportunity to think meta-cognitively about their learning
- **Collaboration.** A classroom culture in which teachers and students are partners in learning should be established (McManus, 2009)

Commentary. FAST SCASS adopted this definition of formative assessment as a process because the empirical evidence then available, and not contradicted by subsequent research, stressed the

importance of using ongoing assessment evidence so that teachers could, if necessary, adjust their instructional activities, or students could adjust their learning tactics. Although not all of the attributes need to be present for the practice to be formative, in concert they enhance the process.

II. INTERIM ASSESSMENT

Definition: Interim tests are typically administered periodically throughout the school year (e.g., every few months) to fulfill one or more of the following functions: predictive (identifying students readiness for success on a later high-stakes test), evaluative (to appraise ongoing educational programs), and/or instructional (to supply teachers with individual student performance data).

Commentary. Also called “benchmark,” “interim benchmark,” “common,” or “quarterly assessments,” these tests have become particularly popular in recent years. Many interim tests were originally marketed as though they were, in and of themselves, “formative assessments,” and yet, as Heritage recently reminded us in a CCSSO white paper.

“The thesis of this paper is that, despite the pioneering efforts of CCSSO and other organizations in the U.S., we already risk losing the promise that formative assessment holds for teaching and learning. The core problem lies in the false, but nonetheless widespread, assumption that formative assessment is a particular kind of measurement instrument, rather than a process that is fundamental and indigenous to the practice of teaching and learning” (2010).

Often, interim assessments are used to predict students’ scores on subsequent high-stakes tests such as a state’s annual accountability assessment. The rationale for such predictions is that the test can identify students who are apt to be successful or not likely to be successful on the subsequent test. Assuming such students can be identified by the use of interim tests, it seems sensible to supply tailored instruction for such students. However, after serious scrutiny of research studies on the

effectiveness of interim assessments in promoting students’ learning, it has been found that, at the moment, there is no meaningful evidence that such assessments do, in fact, enhance student learning (Arter, 2010).

Interim tests can provide periodic snapshots of student learning. However, because they are not proximate to student learning as it is developing, they do not serve the purpose of informing ongoing teaching and learning.

III. SUMMATIVE ASSESSMENT

Definition: Assessment referred to as summative is designed to provide information regarding the level of student, school, or program success at an end point in time. Summative tests are administered after the conclusion of instruction. The results are used to fulfill summative functions, such as to (1) reach an evaluative judgment about the effectiveness of a recently concluded educational program; (2) arrive at an inference about a student’s mastery of the curricular aims sought during an in-class instructional sequence; (3) arrive at a grade; or (4) meet local, state, and federal accountability requirements.

Commentary. Assessments referred to as summative can range from large-scale assessment systems, such as the annual assessments administered across states, to district-wide assessment systems or tests, to classroom summative tests created by teachers. In each instance, the assessments are designed to yield interpretations regarding students’ achievement or program success up to that point in time.

Additional Definitions of Assessment Types.

Listed below are several other assessment terms referring to assessments processes/tests that may be viewed as synonymous with one or more of the assessments described above. Our intent is to explain how these both differ and overlap with other types of assessment processes or uses of assessment instruments.

IV. DIAGNOSTIC ASSESSMENT

Definition: Diagnostic assessments are evidence-gathering procedures that provide a sufficiently clear indication regarding which targeted subskills or bodies of enabling knowledge a student possesses or does not possess — thereby supplying the information needed by teachers when they decide how to most appropriately design or modify instructional activities. Because of the time intensive and specific nature of diagnostic assessments, they are only used for the subset of students identified as not making sufficient progress.

Commentary. The function of the diagnostic assessment process is to provide information to teachers about what is not being learned if students are not making progress. Such assessment information is most useful when it indicates to teachers or students what needs to be done to progress learning. Because diagnostic assessments are only of genuine value when they provide a reasonably accurate estimate of students' status with respect to the curricular aims being measured, it is important for these tests to contain a sufficient number of items/tasks—per assessed curricular aim—to permit valid inferences regarding students' current status.

When diagnostic tests provide teachers with immediate, instructionally tractable information, they are a useful resource in the process of formative assessment.

V. CURRICULUM-EMBEDDED ASSESSMENT

Definition: Curriculum-embedded tests are those that have been deliberately incorporated either in the instructional materials being used by students or in the instructional activities routinely taking place.

Commentary. In contrast to a distinct break in an instructional sequence when students are told that they should put aside their materials in order “to

take a test,” curriculum-embedded assessment provides an essentially non-intrusive way of ascertaining students' current status with respect to the curricular aims being promoted by a teacher (or, more frequently, on the “en route” enabling subskills and bodies of knowledge leading to student mastery of those curricular aims).

The function of these assessments is more apt to be formative when non-graded and used by both teachers and students to support further learning. If they are used in the grading process, the only feature that distinguishes this sort of assessment from other summative assessment processes is the manner in which it has been blended into ongoing instructional activities.

VI. UNIVERSAL SCREENING ASSESSMENT

Definition: Universal screening tests are periodically conducted, usually two or three times during a school year, to identify students who may be at risk, monitor student progress, or predict students' likelihood of success on meeting or exceeding curricular benchmarks. Universal screening tests are typically brief and conducted with all students at a particular grade level.

Commentary. Universal screening measures consist of brief tests focused on target skills (e.g., phonological awareness) that are highly predictive of future outcomes (Ikeda, Neessen, & Witt, 2008). Some universal screening assessment systems feature two or more forms of the same tests measuring student mastery of the same curricular targets.

Results of the screening assessment may be used to identify students needing more targeted progress monitoring or more challenging curricular targets. Within the formative assessment process, these assessment systems are used to identify students who subsequently need more frequent or intensive opportunities to reveal their knowledge and skills during an instructional cycle.

VII. PROGRESS MONITORING ASSESSMENT

Definition: Progress-monitoring tests are periodically administered, typically weekly or biweekly, to gauge students' growth toward mastery of (1) a target curricular aim or (2) the en route-subskills and bodies of enabling knowledge contributing to students' mastery of a target curricular aim.

Commentary. If educators wish to ascertain the degree to which their students are making satisfactory progress toward certain curricular

targets, then one way to do so is to administer, at different points during an instructional sequence, tests specifically intended to gauge student progress. This could also be done, of course, by re-administering precisely the same test repeatedly. However, because students' interactions with earlier administrations of the same test form or items may boost their scores on subsequent test administrations (leading to less valid score-based inferences), most progress monitoring assessments feature multiple forms of tests measuring students' mastery of the same curricular targets.

SUMMARY

We have identified, defined, and commented on specific assessment types because we believe in the importance of referring to tests and assessment processes accurately and consistently. Lack of clarity in labeling educational assessments and tests can lead to misunderstanding or even inadequate use of assessments among educators. Formative assessment is particularly vulnerable as it is often misunderstood or misinterpreted as a particular test or product, as opposed to a process used by teachers and their students as an ongoing gauge of the current status of student learning. A clear understanding of formative assessment is vital so that educators receive the assistance they need to learn how this valuable process can be successfully deployed to improve their instruction and ultimately benefit student learning. This professional learning is essential to assuring that effective formative assessment practices are used in the nation's classrooms to improve student achievement. ■

REFERENCES

- Arter, J. (2010, May). *Interim benchmark assessments: Are we getting our eggs in the right basket?* Paper presented at National Council on Measurement in Education conference, Denver, CO.
- Heritage, M. (2010). *Formative assessment and Next-generation assessment systems: Are we losing an opportunity?* Washington, DC: CCSSO. Retrieved July 31, 2012, from http://www.ccsso.org/Resources/Publications/Formative_Assessment_and_Next-Generation_Assessment_Systems.html.
- Ikeda, M.J., Neessen, E., & Witt, J.V. (2008). Best practices in universal screening. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology* (pp. 103-114). Bethesda, MD: National Association of School Psychologists.
- McManus, S. (2009). *The attributes of effective formative assessment*. Washington, DC: CCSSO. Retrieved July 31, 2012, from <http://www.ccsso.org>
- National Research Council. (2001). *Knowing what students know: The science of design and educational assessment*. Washington, DC: National Academy Press.

ADDITIONAL READING

- Black, P.J., & William, D. (1998). *Assessment and classroom learning*. *Assessment in Education: Principles Policy and Practice*, 5(1), 7-73.
- Erickson, F. (2007). Some thoughts on “proximal” formative assessment of student learning. *Yearbook of the National Society for the Study of Education*, 106(1), 186-216.
- Hattie, J., & Timperely, H. (2007). *The power of feedback*. *Review of Educational Research*, 77, 81-112.
- Heritage, M. (2008). *Learning progressions: Supporting instruction and formative assessment*. Washington, DC: CCSSO. Retrieved July 31, 2012, from <http://www.ccsso.org/>
- Heritage, M. (2010). *Formative assessment: Making it happen in the classroom*. Thousand Oaks, CA: Corwin Press.
- Hosp, J.L. (2011). Using assessment data to make decisions about teaching and learning. In *The APA Handbook of Educational Psychology* (pp. 87-110). Washington DC: American Psychological Association.
- National Center on Response to Intervention. Please visit <http://www.rti4success.org/> for more information.
- Organisation for Economic Co-operation and Development (OECD) Centre for Educational Research and Innovation. (2005). *Formative assessment: Improving student learning in secondary classrooms*. Retrieved from <http://www.oecd.org/dataoecd/19/31/35661078.pdf>.
- Popham, W.J. (2008). *Transformative assessment*. Alexandria, VA: ASCD.
- Sadler, D.R. (1989). Formative assessment and the design of instructional strategies. *Instructional Science*, 18, 119-144.

Tunstall, P., & Gipps, C.V. (1996). How does your teacher help you to make your work better? Children's understanding of formative assessment. *The Curriculum Journal*, 7(2), 185-203.

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

Torrance, H., & Pryor, J. (1998). *Investigating formative assessment*. Buckingham: Open University Press.

William, D. (2012). *Embedded formative assessment*. Bloomington, IN: Solution Tree Press.