2011

Research summary: Assessment

Robert Michael Capraro  
*Texas A & M University*

Mary F. Roe  
*Arizona State University*

Micki M. Caskey  
*Portland State University*, caskeym@pdx.edu

David Strahan  
*Western Carolina University*

Penny A. Bishop  
*University of Vermont*

*See next page for additional authors*

Let us know how access to this document benefits you.

Follow this and additional works at: [https://pdxscholar.library.pdx.edu/ci_fac](https://pdxscholar.library.pdx.edu/ci_fac)

Part of the Curriculum and Instruction Commons, and the Junior High, Intermediate, Middle School Education and Teaching Commons

**Recommended Citation**


This Article is brought to you for free and open access. It has been accepted for inclusion in Curriculum and Instruction Faculty Publications and Presentations by an authorized administrator of PDXScholar. For more information, please contact pdxscholar@pdx.edu.
Authors
Robert Michael Capraro, Mary F. Roe, Micki M. Caskey, David Strahan, Penny A. Bishop, and Christopher C. Weiss

This article is available at PDXScholar: https://pdxscholar.library.pdx.edu/ci_fac/7
Assessment is important for middle level teachers and their students. In fact, the National Middle School Association (NMSA) highlighted curriculum, instruction, and assessment in *This We Believe: Keys to Educating Young Adolescents* (NMSA, 2010). The intention of this summary is to establish assessment’s rightful position as one priority for middle grade teachers and their students. When used wisely and well, teachers obtain information about their students’ strengths and needs, and their students remain informed about their achievements.

To begin, educators need an operational definition of assessment. Based on the work of many scholars (e.g., Delclos, Vye, Burns, Bransford, & Hasselbring, 1992; Poehner, 2007), assessment is defined as a process for documenting, in measurable terms, the knowledge, skills, attitudes, and beliefs of the learner. Although this definition of assessment is rather straightforward, the process of assessment in the classroom is complex. At the classroom level, teachers must decide which specific knowledge, skills, attitudes, and beliefs warrant assessment; at what point and for what specific purpose they should be assessed; and which tools might best accomplish these classroom-based assessments. This research summary addresses two forms of assessment, formative and summative.

**Formative Assessment**

Formative assessment occurs throughout the school year. Initially, it identifies baseline information about students’ achievements to inform instruction. As the school year progresses, formative assessments update teachers’ understandings of their students’ needs and accomplishments (Afflerbach, 2008). Formative assessment data include the cognitive components (e.g., skills and strategies) and the affective dimensions (e.g., attitudes, motivation, and experiences) of learning that allow it to occur (Guthrie & Wigfield, 1997). According to Stiggins and Chappuis (2006), formative assessment is assessment for learning. Studies such as one conducted by Kerr, Marsh, Ikeomota, Darilek, and Barney (2006) noted many favorable outcomes attributed to formative assessment. Teachers increase their regard for data, and the alignment between the curriculum and instruction improves. Their initial use of formative assessment provides a window into students’ achievement and indicates strengths and impediments to future learning. As the school year progresses, formative assessments detail students’ learning, growth, and challenges (Afflerbach, 2008). While benefitting teachers, formative assessment also provides advantages to students. They become more closely attuned to learning goals and their progress toward achieving them. As noted by Black and Wiliam (1998), student performance also improves. When taken as a whole, the artifacts used for formative assessments provide progressive indications of student knowledge of strategies and content. They provide a richer and more complete picture of what students know than would otherwise be available for teachers and students.

Within the classroom, and as found by Bryk, Sebring, Allensworth, Luppescu, and Easton (2010), “data streams create the information feedback loops needed to support a continuous improvement regime” (p. 205). This overall benefit justifies the time and attention that using formative assessment entails. To maximize the advantages of formative assessment, several attributes warrant consideration: (a) the composition of the students (i.e., group versus individual), (b) the content, (c) outcome expectations, (d) time frame, and (e) the time students spend on the activity. These attributes point to the important differences in assessment tools that stem from the number of students to assess, the discipline area under consideration, the amount of time available for the assessment, and the extent of the activity that drives the formative assessment product. Black, Harrison, Lee, Marshall, and Wiliam (2004) identified four central types of formative assessment that seemed to matter most for students: (a) questioning, (b) feedback, (c) peer assessment, and (d) self-assessment. As Black and his colleagues concluded, “The overall message is that formative tests should become a positive part of the learning process. Through active involvement in the testing process, students can see that they can be the beneficiaries rather than the victims of testing, because tests can help them improve their learning” (p. 16).

For many scholars, formative assessments must also have a ring of authenticity (e.g., Hall, 2010; Serafini, 2010). This call for authenticity stems from a basic tenet of quality assessment, which confirms the importance of construct validity and matching assessments to key concepts in the discipline (e.g., Niemi, 1996; Phelan et al., 2009). While there is no clearly agreed upon definition of authentic assessment, the major
focus is that the product is relevant to the learner. Authentic
assessment matches the content being learned, is produced
in conjunction with student interests, and is guided by clearly
defined outcomes. Simply stated, authentic and formative
assessments must coincide with the discipline under
consideration by aligning with what experts in the field
(e.g., historians, scientists, mathematicians, or authors) actually
do. Examples include sketching a science report, listing
historical events, and noting the qualities of good writing.
Authentic assessment can take many more forms, but central
is its link to real world applications (e.g., Darling-Hammond,
Ancess, & Falk, 1995) including 21st century skills (Partnership

**Summative Assessment**

Summative assessment attempts to capture the culmination
of students’ achievements within a specified time frame;
summative assessment is assessment of learning (Stiggins &
Chappuis, 2006). This often occurs at the end of an academic
year as schools and districts administer mandated and
standardized tests to determine annual yearly progress. The
purpose of this end-of-the-year testing involves documenting
what students have learned. Summative assessment can also
occur at the end of an academic unit to identify the overall
success of a program of study with students. In contrast to
formative assessment, summative assessments do little or
nothing to shape future instruction. Instead, summative
assessment captures a moment in time that represents
students’ achievements within the parameters of the test and
testing environment. Some scholars (e.g., Afflerbach, 2008)
asserted that the central role of summative assessment in
the lives of teachers and students introduces an “imbalance”
into a school’s assessment program that minimizes the more
pertinent contribution of formative assessment tools.

Empirical work regarding high-stakes summative assessment
points to many unintended consequences including
increases in school drop-out rates, cheating on exams at the
teacher and school level, and teacher departure from the
profession (Amrein & Berliner, 2002). Further, “the extent to
which states with high-stakes tests outperform states without
high-stakes tests is, at best, indeterminable” (Amrein-Beardsley

Because of the high stakes status of many of today’s
summative assessments, teachers often engage in an array of
activities to prepare students for them. This decision diverts
time from other important instructional tasks; and yet, the
positive effects of such test preparation practice have not
been verified (Valli & Chambliss, 2007).

Other scholars document the narrowing of the curriculum
that often becomes a by-product of summative assessment
(Grant, 2004; Huber & Moore, 2000). For example, according to
Wright (2002), some summative and high-stakes assessments
resulted in new district standards and assessments, the
adoption of new curricular materials in mathematics and
language arts, and the de-emphasis or elimination of content
areas such as art-based education, social studies, the sciences,
engineering, and business options.

Wright (2002) also found that high-stakes summative
assessments affected the school environment. For example,
teachers of tested content were perceived to be more
valuable and received a greater proportion of school
resources. Those who taught outside of the tested content
less often collaborated with peers, were less involved in school
decision making, and were less inclined to critically examine
their teaching practices. According to Wright, this dichotomy
can lead to schools where teachers fail to form close
connections to their peers and consequently foster unhealthy
competition and increase discontentedness.

In a final example, Valli, Croninger, Chambliss, Graeber, and
Buese (2008) found that these high-stakes summative tests led
the teachers whom they studied to stray from the qualities of
good teaching. This included setting aside learner-sensitive
responses in favor of moving lessons forward and covering
the identified content, reducing the cognitive challenge of the
lessons they designed, and posing lower level questions.

Nevertheless, summative assessment remains an important
inclusion in an assessment package (American Psychological
Association, 2010). Citing unintended consequences to set
aside summative assessments would be ill advised. Instead,
and as Wiggins (1998) suggested, a willingness to consider
the downside of assessment—its consequential validity
(Madaus, Russell, & Higgins, 2009)—can be used to leverage
the improvement of the tools that teachers use for summative
assessments. For example, teachers can use results of a
summative assessment alongside other relevant and
valid information (e.g., documentation of students’ learning)
to make decisions about students’ achievement and
instructional needs.

**Conclusion**

In the end, good assessment practices include both formative
and summative assessments. In concert, they offer local and
global evidence that teaching and learning are progressing.
Formative assessments direct teachers’ day-to-day decisions
while summative assessments assure a broader base of
educational stakeholders that the attainments of our nation’s
youth meet local, national, and global expectations. Attaining
a strong and coherent relationship between verifying local
gains and confirming national competitiveness demands
careful attention to the data obtained from formative and
summative options to attain a comprehensive and thoughtful
combination of assessments.
REFERENCES


**ANNOTATED REFERENCES**


In this brochure, the American Psychological Association (APA) emphasized that large-scale tests need to be well developed, properly scored, and used appropriately. APA provided a succinct explanation of a critical issue of assessment—measurement validity—whether a test accurately measures the test taker’s knowledge of the subject being tested. Next, the authors highlighted the appropriate use high-stakes testing and identified a set of principles designed to advance “fairness in testing and avoid unintended consequences.” Subsequently, they provided information about gaps between these testing principles and current realities in education. After noting that large-scale testing is only part of a quality assessment system, the authors recommended that future research examine the long-term effect of high-stakes testing on student achievement.


Black and Wiliam conducted an extensive review of the research literature to examine the effects of formative assessment in the classroom. Their study included the review of numerous books and nine year’s worth of more than 160 journals as well as previous reviews of research. Of the approximately 580 articles and chapters, they selected 250 for analysis. (For the full research report, see Black, P., & Wiliam, D. [1998]. *Assessment and classroom learning, Assessment in Education, 5*(1), 7–74.) Black and Wiliam found evidence that improving formative assessment in the classroom led to significant gains in student achievement (effect size 0.4-0.7). They noted that improvement in formative assessment made a greater difference for low achieving students, which narrows the achievement gap and raises overall student achievement. Black and Wiliam also highlighted the need for improvement with regard to assessment practice and specified three issues: (a) self-esteem of student, (b) self-assessment by students, and (c) effective teaching practice. Following specific recommendations for improving formative assessment (e.g., self-esteem of pupils, self-assessment by pupils, and effective teaching), the authors offered their ideas for changing policy and suggested four-point proposal for teacher development that includes (a) learning from development, (b) broadening dissemination efforts, (c) reducing obstacles, and (d) research.


In this brief, Heritage noted that formative assessment could be used as a means to inform effective instruction because it provides information on student needs and progress. She described succinctly what teachers need to know about formative assessment and how they should use formative assessment. After characterizing formative assessment as a way to enlighten teaching and learning in schools, she discussed the accountability environment that relies on summative assessment. She detailed the four core elements of formative assessment: (a) identifying the gap, (b) feedback, (c) student involvement, and (d) learning progressions. Additionally, Heritage highlighted critical elements of teacher knowledge including: (a) domain knowledge, (b) pedagogical content knowledge, (c) knowledge of students’ previous learning, and (d) assessment knowledge; she also identified the skills teachers need including (a) creating classroom conditions for successful assessment, (b) teaching students to self-assess, (c) interpreting the evidence, and (d) matching instruction to the gaps. Heritage concluded by calling for a focus on assessment in preservice and inservice teacher education programs.
In this research report, Wiliam, Lee, Harrison, and Black described their experimental study of teachers’ development of formative assessment and its effect on student achievement. Wiliam and colleagues worked collaboratively with 24 secondary teachers (two math and two science teachers at each of six schools) to develop aspects of formative assessment for use in their classrooms. Specific interventions included (a) in-service sessions for teachers to learn about formative assessment and develop action plans for incorporating formative assessment; and (b) school visits for teaching observations, discussion of teaching, and planning how to use formative assessment more effectively. During the first six months of the project, the teachers experimented with formative assessment strategies (e.g., questioning, sharing criteria with students, self- and peer-assessment). Next, each teacher developed an action plan for implementing formative assessment into his or her own teaching practice. The subsequent fall term, the teachers put their “plans into action.” Wiliam and colleagues found “firm evidence that improving formative assessments does produce tangible benefits in terms of externally mandated assessments” (p. 63). In other words, formative assessment has a positive effect on students’ achievement on standardized tests.

**LIST OF RECOMMENDED RESOURCES**


**AUTHOR**

*Robert M. Capraro* is a professor of mathematics education in the Department of Teaching, Learning, and Culture at Texas A&M University and co-director of the Aggie STEM Center. He also is a member of AMLE’s Research Advisory Committee, the associate editor of *School Science and Mathematics* and *Middle Grades Research Journal*, and was associate editor of *American Educational Research Journal*.

*Mary F. Roe* is professor in the Mary Lou Fulton Teachers College at Arizona State University. She is a president of the Association of Literacy Educators and Researchers (formerly the College Reading Association) and a member of the AMLE’s Research Advisory Committee.

*Micki M. Caskey* is a professor of middle grades education in the Department of Curriculum and Instruction at Portland State University. She is the chair of AMLE’s Research Advisory Committee, past editor of *Research in Middle Level Education Online*, and immediate-past chair of the Middle Level Education Research SIG.

*David Strahan* is the Taft B. Botner Distinguished Professor in the School of Teaching and Learning at Western Carolina University. He is a member of AMLE’s Research Advisory Committee.

*Penny A. Bishop* is a professor of middle level education at the University of Vermont. She directs the Tarrant Institute for Innovative Education. She also is the chair of the Middle Level Education Research SIG and a member of AMLE’s Research Advisory Committee.

*Christopher C. Weiss* is the director of Quantitative Methods in the Social Sciences Seminar, Institute for Social and Economic Research and Policy. He is a member of AMLE’s Research Advisory Committee.
Karen Weller Swanson is associate professor in the Tift College of Education at Mercer University and director of doctoral studies in curriculum and instruction. She is also editor of Research in Middle Level Education Online and a member of AMLE’s Research Advisory Committee.

CITATION