**Student Assessment**

# Formative Assessment Systems: Finding the Right Fit for Your School

**naesp**™    **National Association of Elementary School Principals**

## About NAESP

The mission of the National Association of Elementary School Principals (NAESP) is to lead in the advocacy and support for elementary and middle-level principals and other education leaders in their commitment for all children.

Gail Connelly, *Executive Director*
Michael Schooley, *Deputy Executive Director*

National Association of
Elementary School Principals
1615 Duke Street
Alexandria, Virginia 22314
800-386-2377 or 703-684-3345
www.naesp.org

NAESP members receive access to this white paper as a member benefit. Learn more at www.naesp.org/BPBS

## About BEST PRACTICES FOR BETTER SCHOOLS™

*Best Practices for Better Schools*, an online publications series developed by the National Association of Elementary School Principals (NAESP), is intended to strengthen the effectiveness of elementary and middle-level principals by providing information and insight about research-based practices that strengthen education. This series of publications is intended to inform discussion, strategies, and implementation, not to imply endorsement of any specific approach by NAESP.

## About This White Paper

This white paper is a component of a partnership between NAESP and ASSISTments, a powerful, free web-based assessment tool that supports all subjects, with a very robust repository of math content that teachers can use to write or select specific questions to support classroom instruction, connect data to instruction and student achievement, give students immediate feedback, get instant reports to help inform instruction and delivery, and assign targeted work to each student directly. As such it enables principals to provide ongoing instructional support for teachers, create common assessments across an entire grade level or school, use data-based formative assessments to track student learning, and build benchmarks aligned to common-core standards. ASSISTments was developed with a federal grant from the US Department of Education and the National Science Foundation, among others, and was developed as a collaborative effort between Worchester Polytechnic Institute and Carnegie Mellon University.

NAESP cares about the environment. This white paper is only available from NAESP as an online document. Readers are encouraged to share this document with colleagues.

# Formative Assessment Systems: Finding the Right Fit for Your School

**K-12 EDUCATORS** need access to reliable measures that accurately determine student achievement growth between state-level assessments. Formative assessments systems (FAS) offer educators immediate feedback on whether students have learned what their teacher intended and how the classroom curriculum or learning activities can undergo real-time course corrections.

The demand for formative assessment is clear. Schools are employing a variety of FAS ranging from "home-grown" tests created by teachers to several commercially packaged assessment systems costing $12 or more per student. But how do school administrators find the FAS that meets the unique needs of their student population?

In ["Which One is 'Just Right'? What Every Educator Should Know About Formative Assessment Systems,"](#) authors Matthew Militello and Neil Heffernan provide a framework for school leaders who have to make important decisions about FAS.

### Driving Force(s) Behind Formative Assessment

The accountability requirements in the No Child Left Behind (NCLB) law and the push to better prepare U.S. students to compete globally are driving educators' need to improve student performance. International assessment results remind us of the gap between U.S. students and their international peers. NCLB ushered in a new era of accountability rooted in the collection, analysis, and use of student assessment data for educational improvement.

As a result, school districts and their leaders must raise the stakes on educational assessments. Beyond using state-level data, school leaders are searching for an assessment that can show within-year growth on learning objectives, diagnose within-year learning needs of students, and predict achievement level on the state assessment.

### Fit Matters

Militello and Heffernan use the iconic story of Goldilocks as an illustration of the school administrator's search for the "just right" fit. When contemplating the fit of the FAS for their student population, educators' metrics should include both the purpose of the assessments (e.g., properties of assessment including validity) and the intended uses by school educators (e.g., lesson planning).

Formative assessment companies aren't inherently bad. They have filled a niche. However, schools and educational policymakers need to be better consumers. Schools are searching for an assessment that can show within-year growth on learning objectives, diagnose within-year learning needs of students, and predict achievement level on the state assessment.

## EDUCATIONAL ASSESSMENTS OF EVERY SHAPE AND SIZE

- **Large-scale assessments can be either criterion-referenced (e.g., NCLB state-level assessments) or norm-referenced (e.g., TIMSS, NAEP, SAT, ACT).**
- **Small-scale assessments tend to be conducted and analyzed at the classroom level and may includevaluable qualitative understandings of the teachers.**
- **Formative assessments reside in the middle.**

## Teacher as End User

While most formative assessment systems claim utility for classroom teachers, Militello and Heffernan assert there is little evidence that the right assessment data is being provided to them. They highlighted two frameworks that help school administrators find the right assessment system that meets the needs of teachers.

First, the National Research Council's "assessment triangle" offers anchors of assessments including the ability to:

- Diagnose student cognition within a specific subject area,
- Conduct student observations that elicit responses from students and offer evidence of competencies, and
- Make fair and accurate inferences about student achievement.

The second framework outlines the standard features of meaningful, effective use of teacher-level formative assessments:

- Assessments that are linked to a curriculum that is aligned with the district scope and sequence and state curricular benchmarks,
- Assessments that provide timely, student diagnostic-level data,
- Ability to disaggregate data with other datasets (e.g. other student achievement data, perceptional data, etc.) and to easily access and communicate reports with a variety of audiences, and
- Availability of ongoing professional development and immediate on-site assistance to translate data into instructional knowledge.

## A Tale of Three Formative Assessment Systems

To illustrate why fit matters, Militello and Heffernan describe the characteristics of three FAS currently available as well as analyze the characteristics of each system in relationship to the concept of fit.

*Northwest Evaluation Association (NWEA) Measures of Academic Progress*
NWEA MAP assessments are adaptive, online achievement tests.

- If a student answers a question correct, the test presents a more challenging item.
- If the student misses a question, MAP offers a simpler item. (The test narrows in on a student's learning level, engaging them with content that allows them to succeed.)
- All items within each subject area pool are calibrated onto a common scale.
  - ☐ This scaling allows students to be placed onto this same scale even though they respond to different items.
  - ☐ The common scale also allows for analysis of student growth across time.
  - ☐ State-specific item pools are created from the "universe" of all MAP items so that the pools used for a particular state are best matched to the state's curriculum frameworks.

While most formative assessment systems claim utility for classroom teachers, there is little evidence that the right assessment data is being provided to them.

5

*Assessment Technology Incorporated's (ATI) Galileo*

ATI's Galileo is a system for building benchmark assessments.

- ATI works collaboratively with a district to design an assessment system that is aligned with local instruction and informs curriculum planning.
- ATI has an online Benchmark Planner in which the district defines the assessment goals, specifies the standards to be measured and the number of items per standard, and reviews preliminary versions of the assessments.
- Aggregate level reports can be produced at the class, school, and district level, and many can be created interactively to suit the user's needs.
- A primary report for the benchmark assessments lists and describes all the standards (objectives) measured and provides an achievement level classification for each student for each standard.

*The ASSISTment System*

The ASSISTment System blends assessment and instructional assistance. It allows:

- Teachers to write or select questions,
- Students to get immediate useful tutoring,
- Teachers to get instant reports to help inform their instruction,
- Students to receive individual feedback while they are being assessed, and
- Teachers to differentiate instruction by assigning special problem sets for students who needs extra practice, allowing students to get practice until they reach mastery.

Like a human tutor, the ASSISTment System breaks an item down to the individual steps needed to solve the problem. By tracking where students' understanding ends, the

system can track the specific components of a student's learning and give teachers data on the precise pieces of knowledge students have and have not mastered.

## Assessments in Action or Inaction

Militello, Stephen Sireci, and Jason Schweid studied the actual use of these three FAS.

- The researchers deemed NWEA's MAP most appropriate for use at the district-level.
  - ☐ The data generated by the assessment gives district administrators longitudinal scores; patterns from year-to-year can assist them in their decision-making (e.g., professional development opportunities).
  - ☐ However, because the MAP data does not generate item-level data reports, teachers find little utility in its use.
- ATI's Galileo provides a rich set of interim or benchmark assessments that school-level educators are able to use to monitor student gains on what was recently taught.
  - ☐ This FAS also puts pressure on teachers to teach to the curriculum.
- Of the three systems, teachers used only one, the ASSISTment System, in a real-time, cognitive diagnostic manner.

## Which System Best Suits Your School's Needs

The quick answer to the question, "Which formative assessment should we use?" is: *It depends*. How educators want to use assessment data should guide their decision making. But school administrators need to be better consumers. Importance must be placed on the intended use of FAS and the characteristics of the systems must be assessed.

The answer to the question, "Which formative assessment should we use?" should be: It depends. The utility of a formative assessment system is predicated on the end-use. That is, how educators want to use assessment data should guide their consumerism in the selection of a product.

## Role of School Leaders in the Decision Making

School leaders would be well served if they:

- Understood the concept of assessment fit,
- Build teachers' capacity to use assessments that provide student-level diagnostic data,
- Provide adequate resources and support mechanisms, and
- Monitor the use of assessment data.

Finding the "right fit" between the purpose of an assessment system and the intended uses by local educators is an important issue. Asking teachers to use data to inform their teaching in order to advance student achievement requires careful consideration. The constant press to use "data" may result in the use of any data that is readily available.

Such misfit leads to inappropriate uses and, at worst, to poor pedagogy and student confusion. Appropriate uses of formative assessment data require local educators to develop efficacy toward assessments.

Militello and Heffernan suggest that this is a function of utility (how teachers can actually use the data in their practice) and outcomes (teachers can see student growth as a function of using the data in their practice). As school administrators are bombarded with more formative assessment models, the authors caution educators to resist the urge to rally against all tests. Rather they encourage educators to develop their capacity to discriminate among assessment types and embrace, train, and use those assessments that are "just right" for our students.

The future of FAS looks promising, according to Militello and Heffernan, if the assessments are technology-based, curricula aligned, readily accessible to parents and educators, useful to students, and give teachers information about what students are thinking, how they are learning, and strategies they are employing. Only when educators find the assessment that is "just right" will we be able to feed the practice of teachers and improve the achievement of students.

## Related Links

"Which One Is 'Just Right'? What Every Educator Should Know About Formative Assessment Systems": Review the unabridged content related to this best practice.

Northwest Evaluation Association: Learn more about MAP, or Measures of Academic Progress.

Assessment Technology, Incorporated: Explore Galileo's features.

ASSISTments: Examine ASSISTments' offerings.

Assessing Young Children's Learning and Development: *Principal* magazine article from May/June 2011 by Jacqueline Jones describes how the process of assessment is different from the common perception of testing.

Rethinking Teaching and Learning: *Principal* magazine article from March/April 2011 by Daniel V. Salaz focuses on a Phoenix school's alignment of curriculum and instruction to earn big academic gains.

Formative Assessment & Standards-Based Grading: Robert J. Marzano's 2009 book is available from NAESP.

As more and more assessments bombard schools, we should not embrace a Luddite mentality, railing against all tests. Rather we should develop our capacity to discriminate among assessment types to embrace, train, and use those assessments that are ww"just right" for our students.

## References

Brunner, C., Fasca, C., Heinze, J., Honey, M., Light, D., Mandinach, E., et al. (2005). Linking data and learning: The Grow Network study. *Journal for Students Placed at Risk*, 10(3), 241-267.

Coburn, C., & Talbert, J. (2006). Conceptions of evidence use in school districts: Mapping the terrain. *American Journal of Education*, 112(4), 469-495.

Heritage, M. (2007). Formative assessment: What do teachers need to know and do? Phi Delta Kappan, 89(2), 140-145.

Kerr, K. A., Marsh, J. A., Schuyler Ikemoto, G., Darilek, H., & Barney, H. (2006). Strategies to promote data use for instructional improvement: Actions, outcomes, and lessons from three urban districts. *American Journal of Education*, 112(4), 496-520.

Marshall, K. (2008). Interim assessments: A user's guide. Phi Delta Kappan, 90(1), 64-68.

Militello, M., & Schweid, J. (2009). WPI PIMSE Annual Report. Washington, DC: National Science Foundation, Graduate STEM Fellows in K-12 Education (GK-12).

Militello, M., Schweid, J., & Sireci, S. (2010). Formative assessment systems: Evaluating the fit between intended use and product characteristics. *Educational Assessment, Evaluation, and Accountability*. 22(1), 29-52.

Militello, M., Sireci, S., & Schweid, J. (2008, March). Intent, purpose, and fit: An examination of formative assessment systems in school districts. Paper presented at the American Educational Research Association, New York City, NY.

Militello, M., & Sykes, G. (2006, April). Why schools have trouble using data. Paper presented at the National Council on Measurement in Education, San Francisco, CA.

Murnane, R., Sharkey, N. S., & Boudett, K. P. (2005). Using student-assessment results to improve instruction: Lessons from a workshop. *Journal of Education for Students Placed at Risk*, 10(3), 269-280.

National Research Council. (2001). Knowing what students know: The science and design of educational assessment. Washington, DC: National Academic Press.

Streifer, P. A., & Shumann, J. S. (2005). Using data mining to identify actionable information: Breaking ground in data-driven decision making. *Journal of Education for Students Placed at Risk*, 10(3), 281-293.

Wayman, J., & Stringfield, S. (2006). Data use for school improvement: School practices and research perspectives. *American Journal of Education*, 112(4), 463-468.

Wylie, E. C., & Ciofalo, J. (2008). Supporting teachers' use of individual diagnostic items [Electronic Version]. Teachers College Record. Retrieved October 13, 2008 from http://www.tcrecord.org/PrintContent. asp?ContentID=15363