Abstract Title Page

Title: Teacher Support for Mathematics Formative Assessment Using Student Work Samples—Effects on Teacher Practice and Student Learning

Authors and Affiliations:
Andrea D. Beesley, IMPAQ International
Tedra F. Clark, McREL International
Kathleen Dempsey, McREL International
Anne Tweed, McREL International

Abstract Body

Background / Context:
Assessment has profound effects on student learning and motivation. Well-designed formative assessment is associated with major gains in student achievement across all ages and subjects, and has the greatest positive impact on struggling students (Black & Wiliam, 1998a-b). Formative assessment is an evidence-based process of gathering information on three questions: 1) Where am I going? 2) How am I doing now? 3) Where do I go next? to support a learning cycle (Hattie & Timperley, 2007; Sadler, 1989). Therefore, the most important formative assessment practices involve 1) students' understanding of their learning target, 2) the criteria by which they will know how they are doing on that target, and 3) the feedback they receive to help them understand next steps. Literature supports prioritizing these three practices or dimensions of formative assessment (see Exhibit 1).
(insert Exhibit 1 here)

To understand learning targets, students need to have clear knowledge of where they are going with their learning, to avoid the inefficiency and frustration of trial and error (Sadler, 1989). Lower-achieving students in particular benefit from understanding learning targets, because they do not intuit them on their own (Black & Wiliam, 1998b; James, et al., 2006).

Clear success criteria enable students to understand how they are doing on learning targets, so they can judge and regulate the quality of their work and increase their achievement (e.g., Andrade, Du, & Mycek, 2010; Brookhart, Andolina, Zusa, & Furman, 2004; Sadler, 1989). When teachers involve students in the assessment process, students perceive more control of and more responsibility for their own learning (Rieg, 2007). Allowing students to help determine success criteria gives them a feeling of empowerment and makes assessment of their work seem less punitive and more constructive (Brookhart, 1997). In turn, students feel more competent and are more likely to engage in learning (Stiggins, Arter, Chappuis, & Chappuis, 2006).

Descriptive feedback helps learners understand how they are doing and what they might do in the future to increase their knowledge and performance (Brookhart, 2008; Callingham, 2008; Cowie, 2005). In contrast, evaluative feedback such as letter grades and comments like "good job!" does not help students identify strengths and weaknesses, and may even undermine confidence, motivation, and performance (Black & Wiliam, 1998b; Chappuis & Stiggins, 2002).

Purpose / Objective / Research Question / Focus of Study:
We are conducting a three-year project funded by an IES research grant (Goal 2, development) to create and pilot test an innovative professional development (PD) program, *Learning to Use Formative Assessment in Mathematics with the Assessment Work Sample Method* (AWSM). The program builds middle school mathematics teachers' understanding of the characteristics of high-quality formative assessment and their ability to develop and use high-quality formative assessment in their classrooms. The Assessment Work Sample Method is a way to collect, create, discuss, and learn from authentic samples of formative assessment work, so that discussions of assessment are grounded in student and teacher practice. AWSM is intended to help teachers develop a deep and practical understanding of and ability to use formative assessment in mathematics.

Research questions:

1. To what extent can AWSM be implemented with fidelity in an authentic education delivery setting?
2. To what extent does AWSM show promise for improving teacher practice of mathematics formative assessment?
3. To what extent does AWSM show promise for increasing student achievement in mathematics?

Setting:
The study was conducted in a large urban school district in Colorado.

Population / Participants / Subjects:
In the pilot year (2012-2013), seven middle school math teachers and their students from one school participated. In the field test year (2013-2014), 41 teachers from six schools and their students participated.

Intervention / Program / Practice:
We were inspired to create AWSM following results from previous research (Randel et al., 2011) on a program that did not change teacher practice and had few mathematics examples. By contrast, AWSM provides PD that integrates mathematics pedagogical content knowledge and formative assessment practices and skills. In AWSM, student work samples help teachers shift from thinking of teaching as something teachers do to learning as something students do, because student work makes student thinking visible (e.g., Hattie 2009). The AWSM approach uses the Assessment Work Sample Method to collect, create, discuss, and learn about assessment using student work samples that help ground the learning in daily practice (Matsumura, Patthey-Chavez, Valdés, & Garnier 2002). The work samples include a teacher cover sheet that conveys the goals of the lesson, the type of knowledge/skill being developed, the criteria for meeting learning goals, and general information that will help reviewers understand the "what and why" of the assignment. Each sample includes four pieces of student work, two that achieved the learning objectives and two that did not. We selected work samples for AWSM that best reflected the content of each session.

AWSM is structured around 13 face-to-face meetings which include an extended introductory session and 12 sessions of about 45 minutes each. At the introductory session, participants build their understanding of formative assessment as one component of a larger assessment system with an emphasis on providing descriptive feedback to move the learning forward. Participants discuss the characteristics of positive classroom culture, why it is essential for implementing this process, and consider structures that will help them create such a culture.

Sessions 1-8 target the three dimensions of the AWSM process through the examination of authentic student work. Dimensions 1 and 2 represent the foundation for AWSM as teachers identify the goal for student learning and specific criteria for meeting the goal. This process includes an examination of student tasks to ensure that the learning goal, success criteria, and tasks are all well aligned to both academic content and the inferred cognitive demand in the learning goal. As importantly, these dimensions require teachers to clearly communicate these

expectations so that students can track their own progress and make adjustments as they are learning (see Exhibit 1).

During sessions 5-8 participants investigate Dimension 3 and learn how to provide effective oral and written feedback. Participants learn that cues, questions, and recommendations for next steps help students take more responsibility for learning than simply providing correct solutions or step-by-step actions. Participants also learn that students should be active in the feedback process, and strategies for implementing effective peer- and self-assessment are presented. Sessions 9-12 ask participants to share their instructional practice more directly as they generate their own student work samples, present the work to colleagues, and request feedback on how to improve their own implementation of the process. The process is fully integrated as participants reflect on their own progress for implementing formative assessment and identify next steps. Authentic student work samples are used in each session to illustrate the relevant dimension. Some were chosen to be examples of high-quality assessment practice, so that teachers can see what it looks like in the middle school math classroom, while others were chosen to prompt discussion about possible improvements.

Research Design:
The study of AWSM is a mixed methods iterative design-research process with a small pilot at one school and a larger field test at six schools. The AWSM prototype was implemented in 2012-2013, underwent revisions, and then was implemented in 2013-2014 in the same urban district with a larger sample, and is now undergoing final revisions.

Data Collection and Analysis:
Exhibit 2 identifies the research questions, data sources, and intended outcomes. Observation and questionnaire data were analyzed with descriptive statistics, while pre-post work sample data were analyzed in a multilevel model to account for nesting within schools. Open-ended comments on observations and questionnaires as well as focus group results were analyzed in MAXQDA with qualitative methods to identify emerging themes. Analysis on student data is being conducted with a difference-in-differences model; results will be available for the SREE spring 2015 conference.
(insert Exhibit 2 here)

Findings / Results:
Effect sizes for changes in assessment work sample scores from pre- to post-test in the 2013-2014 study indicated that teachers showed the greatest improvement in two dimensions involving feedback: “feedback integrates student involvement” (d = 2.42) and “feedback type” (d = 1.98). Participants also showed high effect sizes (greater than 1) on the two dimensions involving assessment criteria: “alignment of the learning goals and assessment criteria” (d = 1.04) and “clarity of the student assessment criteria” (d =1.35). Participants showed the least improvement in the two dimensions involving learning goals, “focus of the goals on student learning” (d = .17) and “alignment of learning goals and task” (d = .46). The lower effect sizes for the learning goals dimensions were affected by the relatively high baseline scores for this group.

For the average work sample scores, HLM analysis revealed a significant increase from pre- to post-test (see Exhibit 3). Dimensions 3-6 (assessment criteria and feedback) showed statistically

significant increases from pre- to post-test, while changes in Dimensions 1 and 2 (learning goals) were not statistically significant. Scores on a test of mathematics pedagogical content knowledge did not significantly predict change in work sample scores, indicating that it was not a moderator of these effects.
(insert Exhibit 3 here)

Teachers reported in sessions, focus groups, and questionnaires that they had incorporated formative assessment strategies into their classroom practice. They communicated learning goals more frequently and visually and in student-friendly language, and they made success criteria explicit and shared them with students using a variety of tools. They incorporated student peer- and self-assessment into the classroom, focused on the success criteria. Once they had information about student understanding from formative assessment, they grouped students so they could provide re-teaching, practice, or enrichment, depending on students' needs. Many teachers graded fewer assignments in favor of providing descriptive feedback that helped students understand where to go next and elicited student involvement in learning.

Results from the impact analysis of on student achievement will be presented at the conference, as analyses are still ongoing.

Conclusions:
Overall, the findings suggest that the AWSM professional development program can be implemented in a school setting and improved teachers' formative assessment practice, particularly in the areas of assessment criteria and feedback. These effects were seen regardless of teachers' prior math pedagogical content knowledge.

Some of the key lessons learned were:
1. Mathematics teachers see the value of using student peers as resources while learning is progressing, and of empowering students to be more aware of and accountable for their own learning. In AWSM many teachers reported that, consistent with formative assessment research, struggling students benefited the most from these strategies.
2. Mathematics students, even those who are struggling, will be more invested in their own learning when they truly understand learning goal(s) and success criteria, their own current level of mastery, and the next steps. In AWSM, teachers were pleased to learn that students will care about mathematics and invest energy in assignments when they understand the learning goal and how to get there.
3. Similarly, mathematics students are capable of and willing to engage in cognitively demanding tasks, and justify their answers, when supported by teachers through a formative assessment process. In AWSM, teachers who assigned complex problems and provided support found that students were able to explain and justify how they approached the work.
4. Administrative support positively influenced teacher engagement in formative assessment PD and practice. Sustained administrator support seems necessary for long-term instructional change, because administrators can reinforce teachers' formative assessment strategies and alter policies (such as grading policies) that could otherwise unintentionally undermine formative assessment.

Appendices

Appendix A. References

Andrade, H. L., Du, Y., & Mycek, K. (2010). Rubric-referenced self-assessment and middle school students' writing. *Assessment in Education, 17*(2), 199-214.

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

Black, P., & Wiliam, D. (1998b). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139–144.

Brookhart, S. M. (1997). A theoretical framework for the role of classroom assessment in motivating student effort and achievement. *Applied Measurement in Education, 10,* 161–180.

Brookhart, S.M. (2008). *How to give effective feedback to your students*. Alexandria, VA: ASCD.

Brookhart, S., Andolina, M., Zusa, M., & Furman, R. (2004). Minute Math: An action research study of student self-assessment. *Educational Studies in Mathematics*, 57(2), 213–227.

Callingham, R. (2008). Dialogue and feedback: Assessment in the primary mathematics classroom. *Australian Primary Mathematics Classroom, 13*(3), 18–21.

Chappuis, S., & Stiggins, R. J. (2002). Classroom assessment for learning. *Educational Leadership, 60*(1), 40–43.

Cowie, B. (2005). Student commentary on classroom assessment in science: A sociocultural interpretation. *International Journal of Science Education, 27*(2), 199–214.

Even, R. (2005). Using assessment to inform instructional decisions: How hard can it be? *Mathematics Education Research Journal, 17*(3), 45–61.

Hattie, J. (2009). *Visible learning*. London: Routledge.

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

James, M. Black, P., Carmichael, P. Conner, C., Dudley, P., Fox, A., Frost, D., Honour, L., MacBeath, J., McCormick, R., Marshall, B., Pedder, D., Procter, R., Swaffield, S., & Wiliam, D. (2006). *Learning how to learn: Tools for schools*. London: Routledge.

Matsumura, L. C., Patthey-Chavez, G. G., Valdés, R., & Garnier, H. (2002). Teacher feedback, writing assignment quality, and third-grade students' revision in lower- and higher-achieving urban schools. Elementary School Journal, 103(1), 3-25. Morgan, C. & Watson, A. (2002). The interpretative nature of teachers' assessment of students' mathematics: Issues for equity. *Journal for Research in Mathematics Education, 33*(2), 78-110.

McLaughlin, M. W. & Talbert, J. E. (2006). Building school-based teacher learning communities. New York: Teachers College Press.

Morgan, C. & Watson, A. (2002). The interpretative nature of teachers' assessment of students' mathematics: Issues for equity. *Journal for Research in Mathematics Education, 33,* 78–110.

Randel, B., Beesley, A. D., Apthorp, H., Clark, T.F., Wang, X., Cicchinelli, L. F., & Williams, J. M. (2011). *Classroom Assessment for Student Learning: The impact on elementary school mathematics in the Central Region.* (NCEE 2011-4005). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Rieg, S. A. (2007). Classroom assessment strategies: What do students at-risk and teachers perceive as effective and useful? *Journal of Instructional Psychology*, *34*(4), 214–225.

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

Stiggins, R. J., Arter, J. A., Chappuis, J., & Chappuis, S. (2006). *Classroom assessment for student learning: Doing it right—using it well.* Portland, OR: Assessment Training Institute.

Appendix B. Exhibits

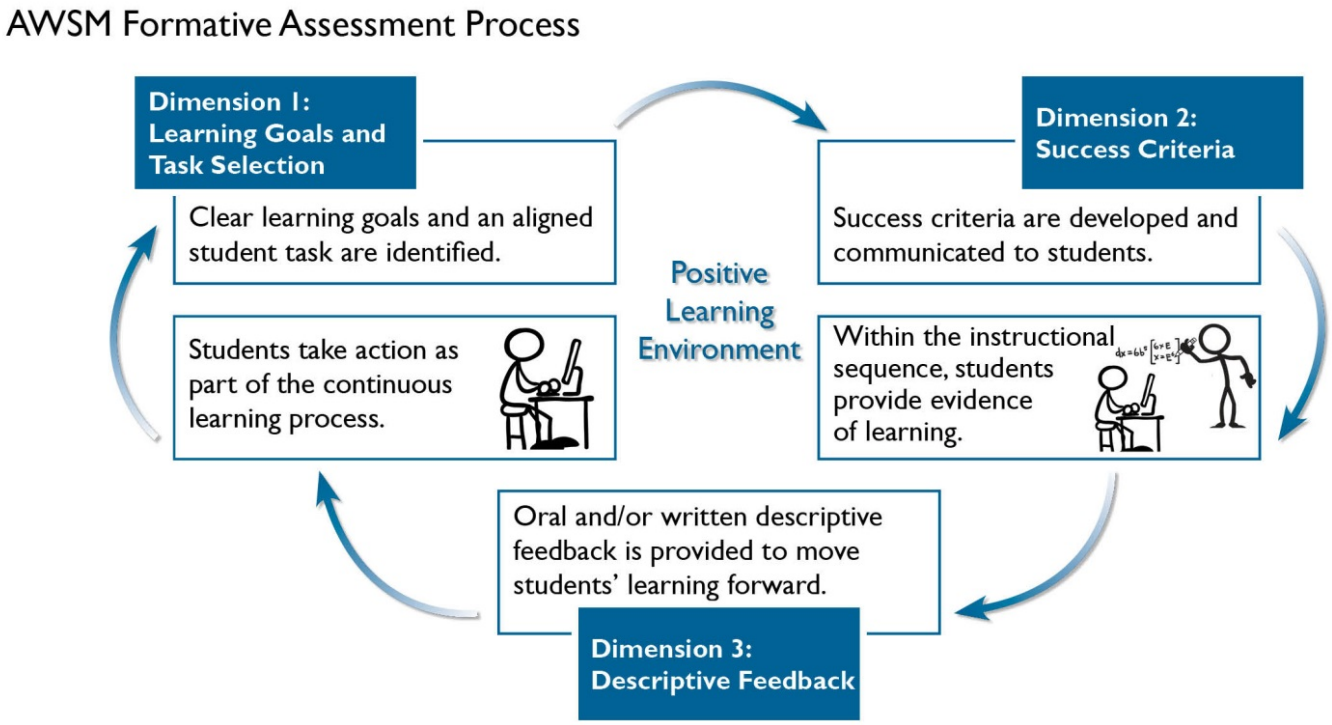

Exhibit 1. Formative assessment process, as used in AWSM

Research Questions	Data Sources	Outcomes
1. To what extent can AWSM be implemented with fidelity in an authentic education delivery setting?	Attendance, observations, questionnaires, focus groups, facilitator debriefs	Developed fidelity measures and data regarding the level of the intervention's implementation in an authentic education delivery setting (feasibility).
2. To what extent does AWSM show promise for improving teacher practice of mathematics formative assessment?	Teacher work sample pre- and posttest scores on rubric	Estimate of the intervention's promise for improving teachers' practice of formative assessment in mathematics.
3. To what extent does AWSM show promise for increasing student achievement in mathematics?	Student pre- and posttest scores on MAP tests	Estimate of the intervention's promise for improving student learning and achievement in mathematics.

Exhibit 2. Alignment of research questions, data sources, and outcomes

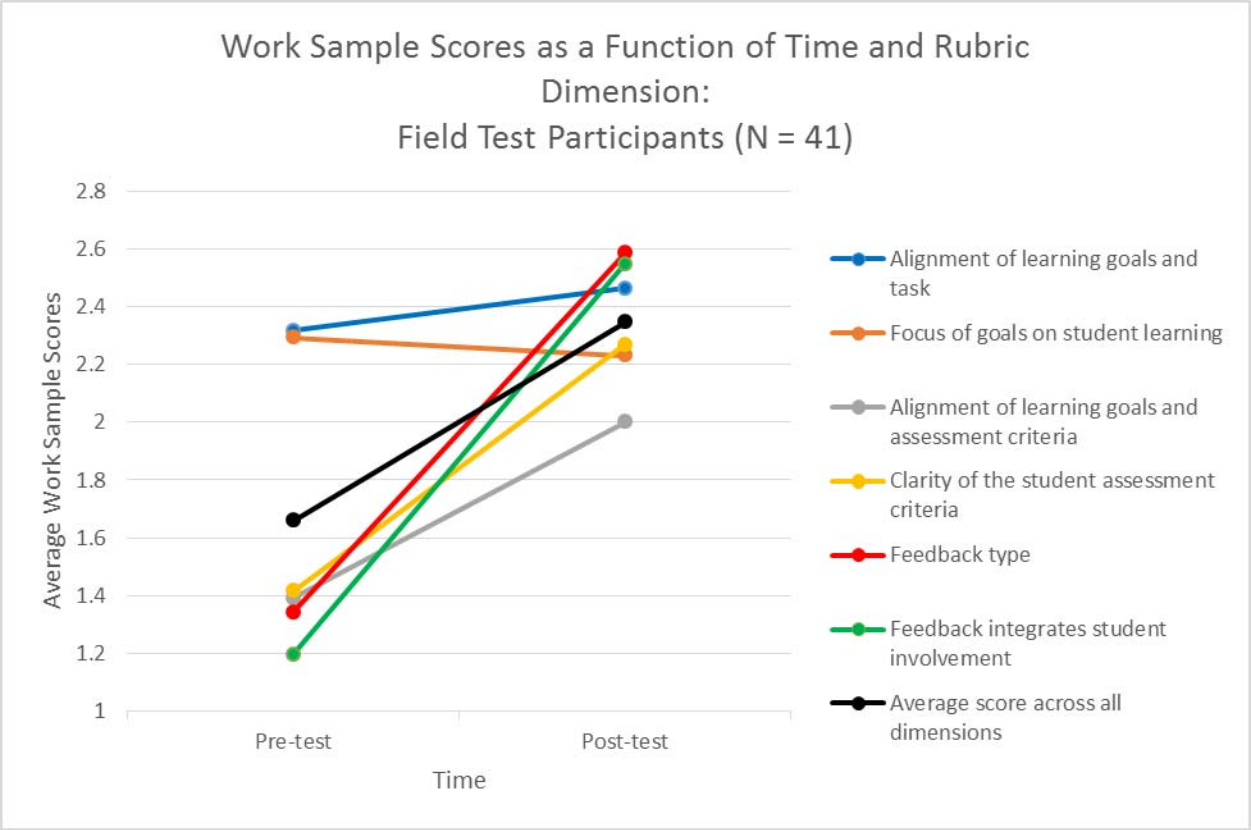

Exhibit 3. Average work sample scores as a function of formative assessment dimension and time (pre-test; post-test)