Putting understanding into practice: Middle-Grades Mathematics Teachers Learn to Implement the Formative Assessment Process

Lynn T. Goldsmith   Sophia S. Mansori

Education Development Center, Inc.

43 Foundry Avenue

Waltham, MA

Abstract

This paper reports on middle-grades teachers' learning about formative assessment in FACETS, a two-year PD program designed to promote an understanding of FA as a comprehensive cycle of instruction and to provide supports for teachers during development of FA classroom practices. Both teacher successes and challenges will be discussed.

Acknowledgements: Cheryl Tobey, Susan Jansen Creighton, Emily Fagan, and Eric Karnowski and have been our FACETS colleagues and sounding boards. In addition, Fred Gross worked with us for the first portion of this project and Cindy Char has provided able project oversight and feedback. Advisors Steve Benson, Mark Dricoll, Margaret Heritage, Karen King, Paola Sztajn, and Holt Wilson have also helped us to shape our work. This research was supported by the National Science Foundation under Grant DRL-0918438. Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Putting understanding into practice: Learning to implement formative assessment

Formative assessment (FA), sometimes called assessment *for* learning, has captured the interest of educators for nearly three decades (Black & Wiliam, 1998; Sadler, 1989), and recent focus on the Common Core has stimulated further interest in the potential for FA to support new standards for student learning. The Council of Chief State School Officers (CCSSO), National Council of Supervisors of Mathematics, and Association of Mathematics Teacher Educators all emphasize the value of FA for creating more effective learning environments (AMTE & NCSM, n.d., CCSSO, 2008).

CCSSO identifies FA as "a process used by teachers and students during instruction that provides feedback to adjust ongoing teaching and learning to improve students' achievement of intended instructional outcomes" (CCSSO, 2008, p. 3). This definition touches on central features of FA—it is an instructional process (rather than a set of discrete assessments), in which feedback is a critical element and students and teachers alike are active participants. As such, FA articulates a stance toward instruction that accords with modern theories of motivation and learning (Bandura et al., 2003; Duckworth & Seligman, 2005; Flavell, 1979) as well as with recent descriptions of ambitious teaching practices that aim to promote deep conceptual understanding for all students (Ball & Forzani, 2009; McDonald et al., 2013).

FA entails a set of interrelated practices that inform teachers' instructional decisions from beginning to end of a lesson. For individual lessons, the aim of FA is to close the gap between students' present understanding and lesson learning goals (Heritage, 2010). As a broader approach to instruction, a salient premise of FA is that students' learning is the joint responsibility of teachers and students, and an underlying goal of FA is the development of students' ownership of their own learning (Andrade, 2010; Butler & Winne, 1995; Creighton, Tobey, Karnowski, & Fagan, 2015; Wylie et al., 2012).

Building on earlier models of FA, *Formative Assessment in the Mathematics' Classroom: Engaging Teachers and Students (FACETS)* developed a model emphasizing the role of students (as well as teachers) in the FA process (Creighton et al., 2015), and designed and investigated the course of a two-year PD experience to support middle-grades mathematics teachers' coherent implementation of critical components of the FA process. Our research questions for the project, and for this paper, are the following:

1. How did the teachers' knowledge and practice of FA change after participating in the FACETS program?
2. How did the course of learning to implement FA unfold, and what are common challenges and/or barriers to implementation?

FACETS model for the FA process: Focus on teachers' and students' roles

The FA process itself is cyclical; individual components of the process come into play over the course of a lesson, and then the cycle begins anew to support the next lesson's instructional goals. Individual elements of the process drive instructional decisions designed to identify and address gaps between students' current understanding and the learning goals for the lesson (Heritage, 2010). Models of the FA process typically emphasize the work of the teacher in identifying and remediating these gaps—an emphasis that makes sense given that PD goals are to support teachers' implementationof FA practices.

The FACETS model of FA builds on earlier articulations of the FA process (Heritage, 2010; Nicol & McFarlane-Dick, 2006), paying specific attention to both teachers' and students' roles in the iterative instructional process. As Figure 1 details, the model is comprehensive and complex, highlighting both teachers' and students' work throughout the cycle (see also http://famath.edc.org/resources/formative-assessment-cycle for a video tour of the model's components and interconnections.) The explicit attention to students is consistent with our view that a primary goal of FA is to support students' becoming self-regulating learners who can answer, and act on, three basic questions: What are my goals for my learning? Where am I currently in relation to my goals? If I haven't yet met those goals, what do I need to move toward them? (Creighton et al., 2015; Nichol & McFarlane-Dick, 2006; Sadler, 1989; Tobey & Goldsmith, 2013).

Insert Figure 1 about here

The FACETS model parses the elements of the FA process into four *critical* aspects of FA process and two more global *supporting* aspects. The four critical aspects reflect the elements of FA identified in the literature.

1. Learning intentions (LI) and success criteria (SC): defining and sharing with students the lesson's intended learning goals the indicators of successful achievement of the intended learning
2. Eliciting and interpreting evidence of students' thinking: providing opportunities for students to make their thinking explicit in order to assess their understanding with respect to lesson [LI and] SC
3. Formative feedback: providing students with descriptive, actionable feedback about students' progress toward SC that scaffolds continued learning
4. Student ownership and involvement: supporting students' active engagement and increasing independence through development of self-regulation skills.

The two supporting aspects address the intellectual and social contexts that support classroom learning and teaching:

1. Learning progressions: articulating typical pathways for the development of content understanding, which are used to inform content-based instructional decision-making within and across lessons
2. Classroom environment: the physical layout of the classroom, norms for social behavior, and expectations for participation

While attention to learning progressions is solely the responsibility of the teacher, the nature of the classroom environment is shaped by both teacher and student participation in classroom activities. Teachers must, of course, be the ones to set the tone and expectations for developing a classroom culture that is respectful to all and conducive to everyone's learning, but they cannot create such an environment on their own. Students must learn and participate in the norms for behavior, engagement in classroom activities, and interactions with others.

The FACETS model also emphasizes that FA is more than a set of strategies for interacting with students *during* classroom lessons, identifying aspects of the cycle that occur as teachers are planning lessons and engaged in post-lesson review and reflection. Finally, the model highlights the importance of making instructional decisions in response to information gathered about students' current understanding in relation to lesson learning goals.

These critical and supporting aspects of FA pertain for instruction across subject area domains. In fact, many authors discuss FA as a general approach to instruction and draw on examples from a range of subject areas (c.f., Black & Wiliam, 1998; Heritage, 2010; Wiliam, 2011; Wylie et al., 2012). FA as a practice is deeply embedded in the acquisition of domain-specific knowledge, as attention to learning progressions attests. Teachers seeking to learn to enact FA skillfully must not only develop general instructional strategies that support the implementation of the critical aspects described above, but must ground the decision-making motivating these moves in understanding of subject area content and how students learn that content. In the case of mathematics instruction, learning to implement FA not only requires building new instructional strategies such as exit slips or peer conferencing, but the application of these new approaches to mathematics learning through articulation of clear and conceptually worthwhile goals for mathematical understanding.

FACETS was designed to address the content-specific nature of FA in practice. FACETS staff developed and studied FA-based professional development for middle-grades mathematics teachers, grounding the FA focus in the context of mathematics teaching and learning. By doing so, we sought to avoid a situation where teachers learned about general FA approaches and strategies but were then left on their own to figure out how to use them effectively for promoting mathematics learning. Our goal in investigating the course of teachers' learning during the FACETS PD was to gain a deeper understanding of how teachers learn to implement FA in their classrooms; as a field we know that the process is a time consuming one (Trumbull & Gergon, 2011), but we know less about how that process unfolds over time. In particular, we were interested in understanding how teachers implemented FA practices over time, and common challenges and/or barriers to their growth as FA practitioners.

The FA model guided the development of a two-year PD experience, focusing both on promoting greater understanding of FA as a framework for instructional decision-making and on supporting coherent, integrated, and fluid implementation of FA practices over the course of the two years. We developed a pilot version of the PD with a cohort of 23 teachers during the 2010-2012 academic years, refined the approach based on data from that pilot group, and offered a second iteration of FACETS from 2012-2014. This paper describes the final (2012-2014) version of the PD, and reports on findings from the second cohort of 40 participating teachers.

FACETS Professional Development

The FACETS program provided more than 90 hours of PD over a 24-month period. Program components are summarized in Table 1.

Insert Table 1 about here

The summer institutes and workshops focused on exploring the FA model and its component parts with the goal of deepening participants' understanding of the conceptual underpinnings of FA and scaffolding opportunities for teachers to explore strategies for implementing the critical and supporting aspects in their own classrooms. The focus of mathematical work during the institutes and workshops was proportional reasoning.

During the school year monthly learning groups provided ongoing opportunities for teachers to explore their implementation efforts, putting their knowledge of FA into practice. Each learning group was facilitated by a FACETS staff member. Mathematical content of learning groups were more variable because teachers often discussed the mathematics they were teaching at that particular time. Meetings during the first year were organized around implementation of individual critical aspects; the focus in the second year shifted to exploring how the individual aspects contributed to a more integrated and coherent practice that emphasized the role of students in the FA process. The overall arc of learning group topics was common across the four groups (jointly developed by all four learning group facilitators), although facilitators had the flexibility to adapt session agendas to fit the specific needs of their group members. In addition, learning group facilitators made classroom visits to group members on a rotating basis, debriefing with the individual teachers. The four facilitators also used their observations in planning subsequent learning group activities.

Project staff used these different PD structures to move between theory and practice in order to deepen both—teachers would understand both the how's and the why's of FA better by implementing its components on a regular basis, and their implementation of FA practices would become more skillful, intentional, and habitual (routine and seamless) the more they understood how FA builds on mathematical content and relies on certain pedagogical approaches.

Methods

The study uses a "convergent-parallel" mixed-methods design (Creswell & Plano Clark, 2007), in which both quantitative and qualitative data help to develop a picture of teachers' learning about the theory and implementation FA and to explore challenges teachers face while learning to implement FA.

A number of data sources, described below, inform our findings. Teachers completed all research measures at the beginning and end of the project; in addition, we collected mid-program data (end of year 1 of PD) for several measures. Table 2 outlines data sources, the construct assessed, and timing of administration.

Insert Table 2 about here

Mathematical Knowledge for Teaching. We assessed teachers' mathematical knowledge for teaching (MKT) (Hill, Schilling, & Ball, 2004) at the beginning and the end of the PD to determine whether participation in FACETS affected MKT. Additionally, we were interested in understanding how teachers' MKT affected teachers' implementation of different critical aspects of FA. Teachers completed an online version of the assessment for number and operation (Schilling & Hill, 2007) through the Teacher Knowledge Assessment System (TKAS) website. TKAS offers the choice of using either a fixed form or computer-assisted version; we elected the computer-assisted version, which presents items to the participant only as long as necessary to estimate a stable IRT score (which effectively functions as a z-score).

Concept maps. Teachers were given instructions to create concept maps that represented their understanding of FA; these were collected from teachers at four points in time: pre-PD, end of first summer institute, end of year 1, end of year 2. Teachers also completed questionnaires that asked them to compare earlier and later maps, writing reflections on how the similarities/differences reflected growth in their understanding of FA. The research team developed a five-point scoring rubric for maps that captured the major content of higher order nodes, as well as the structural connections among notes (Hough, O'Rode, Terman, & Weisglass, 2007). The rubric was anchored at Level 1 by maps organized around the idea of gathering information without reference to its use, and at Level 5 by representation of all the critical aspects, within a structure that reflected an interconnected process. Maps were reviewed by both authors and scored by consensus.

Classroom Practices survey. This survey, developed by the FACETS staff, consists of Likert items about teachers' frequency of implementation and their satisfaction with their FA practice, as well as open response items that included a question asking for a definition of FA as well as an example and a non-examples, and questions about teachers' perceived success and challenges. The survey was administered online prior to the beginning of the PD, midway through the program (the end of year 1), and at FACETS's end. Pilot work with the first cohort of teachers indicated that it would not be possible, prior to the PD itself, to ask in a meaningful way about some of the specific practices targeted in the PD. These items were therefore not included in the pre-program version of the survey, but were added for mid- (end of year 1) and post-program administrations.

Scaled items were analyzed both in the aggregate, and by pairing individual teacher's responses pre and post to measure change. The FA definitions were coded and scored based on the inclusion of each of the critical aspects along with other concepts such as reference to FA as an instructional process. Other open response items were also coded using the critical aspects, along with emergent coding of challenges and barriers.

Student Work Analysis. This measure, developed by the FACETS staff, was modeled after the POWERSOURCE measure (Heritage et al., 2008) and served as a proxy for classroom FA practice. Teachers read a brief description of a teacher's upcoming lesson, including a description of one of the math problems students will work on during class. Teachers then identify key mathematical ideas of the problem, analyze accompanying student work (selected to highlight common student errors), and indicate how they would respond to these students if they had been teaching the class. This was a pencil-and-paper measure, administered pre- and post-program. Responses were scored according to a scoring schema that assessed implementation of FA practices: (1) identification and use of LISC (maximum of 7 points), (2) eliciting and interpreting students' thinking (maximum of 5 points), (3) providing formative feedback to students (maximum of 9 points).

Classroom observations, video, and interviews. For the case study teachers, the research staff interviewed and observed these teachers in the spring of year 1, conducted a focus group interview during the second summer institute, video taped two consecutive days of lessons in the fall and the spring of year 2, and debriefed these observations afterwards. Observation and interview data were described and annotated by multiple researchers, and analyzed as descriptive case studies (Stake, 1995 Yin, 2003).

Artifacts from PD assignments and activities. Artifacts included PD assignments (lesson plans highlighting implementation of critical FA aspects and periodic blog entries), field notes of summer institutes, workshops, selected learning group meetings, and facilitators' classroom visits, as well as occasional email correspondence between participants and project staff. These artifacts were used to inform the researchers' understanding of the PD and to provide context for the analysis conducted using the other data sources.

Participants

Fifty-four teachers from four districts in the New England area entered the FACETS program in the summer of 2012. All but three teachers attended the PD with at least one other colleague from their school. Over the course of the two years, 14 teachers left the program (26% of the original sample).[1] In the winter of year 1, we used facilitator recommendations to select six case study teachers to follow more closely..

Results and Discussion

Tables 3 and 4 present background information of teachers completing the PD. Overall, the group of incoming FACETS participants were experienced classroom teachers, with an average of 10 years of math teaching. About a third of the participants had majored or minored in

[1] Since reporting on attrition is not commonplace in studies of PD, there are few benchmarks against which to compare the losses we experienced in this program. Attrition rates for the two cohorts were, however, similar (6 leavers of 29 original cohort 1 participants, or 21% attrition). Half of the teachers who left FACETS were upper elementary teachers, [which may not be so surprising, given that the mathematical focus of the PD was explicitly middle-grades mathematics) We found no significant differences between pre-program measures of mathematical knowledge for teaching (MKT) or reported frequencies of using nine of 10 FA practices in class, suggesting that teachers who left the program were not systematically different than those who stayed in terms of mathematical understanding or pedagogical approaches. (The one statistically significant difference was that stayers reported more frequent sharing of LI with students at baseline than did leavers, a difference in practice that we would not expect to have a negative impact leavers' sense of the possibility of becoming skillful FA practitioners: $t = 4.19$; $\alpha = .005$, Bonferroni correction for multiple tests; $p = .000$.)

mathematics during college; nearly a quarter of the group had received a graduate degree in mathematics education. FACETS teachers were also involved in their own professional learning. All those responding to the TKAS Teacher Survey indicated that they had participated in some form of PD in the past year, and 92% of the group reported having taken math PD within the year prior to FACETS. Most of the participating teachers (80%) had their own classes; the remaining 20% (6 teachers) were supporting teachers (math coaches or special educators) who worked with lead teachers in in multiple grades. The group was predominately female (34 of the 40 teachers, 85%).

Insert Tables 3 and 4 about here

Growth in Teachers' Understanding of FA

Analyses of teachers' responses to the open ended items in Classroom Practices survey and their concept maps provide convergent evidence for the ways in which teachers developed deeper and more coherent understanding of the FA process over time. Teachers' definitions of FA became more complete over the course of the PD, increasing in the number of relevant characteristics mentioned (pre-program mean = 2.49; post-program mean = 3.89; t=5.70, p=.000). Of the four critical aspects emphasized in the FACETS model, only reference to gathering evidence was part of most teachers' definitions at the beginning of the PD (Table 5).

Insert Table 5 about here

Many participants did include reference to overarching characteristics of FA as an ongoing instructional process and as informing instruction. Given that an alternative view of FA as episodic, test-based collection of student performance is also an active part of the national discourse (Brown, Afflerbach, & Croninger, 2014; Marchand & Furrer, 2014; Popham, 2006), we were pleasantly surprised that so many teachers began FACETS with a perspective on FA that was consistent with the CCSSO position. At the end of the program, substantially more teachers included references to LISC, feedback, and/or students' role in FA in their definitions.

Teachers' progression of cognitive maps tells the same overall story of greater specification of critical aspects and of their interconnections (Figure 2). Three-quarters of teachers' pre-program maps were scored as Levels 1 or 2, levels that reflect map structures that identify some of the elements of FA practice, but lack connections among them so as to reflect a coherent system for using information about students' understanding to make instructional decisions. The changes in content and structure observed in the first two maps (one drawn at

Insert Figure 2 about here

the start of the first summer institute, which kicked off the PD and one drawn at the end of the institute) indicate that teachers grasped the overall structure of the FA process quickly. At the end of summer institute teachers' maps already began to reflect the more coherent view of FA that FACETS promoted— 63% of the maps scored level 3 or above, as compared to 24% of the pre-program maps, all of which were scored at level 3.

Concept maps completed mid-program (end of year 1) and at the end of FACETS continue reflect a trend toward greater completeness with respect to referencing the critical aspects of FA. Over the two years of PD, teachers' maps contained more critical aspects, as well as representations of connections among them. While teachers' pre-program maps elaborated upon their understanding of FA by providing examples of strategies for assessing students' thinking (e.g., exit tickets, questioning, thumbs up/thumbs down, tests/quizzes), examples at the end of the program were more often focused on further justifying the purpose of different critical aspects (e.g., success criteria should be both conceptual (why) as well as procedural (how to);

formative feedback can be clues, hints, or models; formative feedback is also actionable; responsive actions include decisions to gather more evidence, providing feedback, providing more instruction, or moving on to the next lesson).

As a group, teachers' MKT scores also improved, even though the PD did not focus directly on developing teachers' own mathematical understanding (pre-program mean =.156, post-program mean=.438, n=35, t=2.44, , *p*=.02).[2] We speculate that these increases were due to FACETS's ongoing attention to articulating the important mathematical ideas of lessons in order to write conceptual learning goals and to interpreting students' thinking with respect to the mathematics. If this perspective is correct, then FA that is content-focused is likely to provide an effective context for promoting teachers' deeper disciplinary knowledge needed for teaching across a variety of subject areas.

Growth in Teachers' Implementation of FA

The Student Work Analysis and Classroom Practices survey provide information about how the teachers, as a group, implemented FA practices over time

Teachers enacted FA practices associated with the critical aspects with increasing frequency. Both teacher self-report and classroom observations indicated that teachers took up practices related to the critical aspects over the two years of the study: paired t-tests of the items on the Classroom Practices survey indicated significant pre- to post-program changes for all items (Figure 2). While frequencies reported mid-program were all larger than those reported before FACETS began, only two items were statistically significant—an overall statement about the use of FA in instruction, and an item about determining the success criteria for a lesson. This finding reflects the PD focus on LI and SC as the driving point of the FA process and the first

[2] A mean gain of .28 is consistent with average gains in MKT observed across a large number of mathematics PD programs (Phelps, 2015).

critical aspect they were encouraged to explore intensively in their classrooms. Even when facilitators began to discuss implementation of other critical aspects and offer teachers supporting structures and strategies during the first year of the PD, they emphasized the importance of framing the new critical aspects in terms of how they contributed to students achieving the lesson LISC. Figure 3 represents changes in frequency of items teachers rated only mid- and post-program. Mean frequencies for all but one of these practices increased significantly for during the second year of PD.

Insert Figure 2 about here

Teachers shared their work and questions in the learning groups, where many noted that despite a sense of progress toward fuller and more skillful implementation of FA, practices related to student ownership and self-regulation had been the most unfamiliar to them at the outset of FACETS and remained the most challenging throughout the PD. This was the dimension least frequently enacted at the end of the PD, even though teachers came to value the importance of supporting students to become more self-regulating and to wish they were more able to do so. At project's end nearly 50% of the teachers identified some aspect of promoting student ownership and self-regulation as their greatest challenge in implementing FA.

Insert Figure 3 about here

Teachers developed their ability to analyze and respond to student work. The Student Work Analysis instrument complements the teachers' self-report of classroom practices by providing insight into how teachers interpreted evidence of students' thinking and how they made decisions about next instructional steps, based on such evidence. The measure, therefore, served as a proxy for conducting and debriefing classroom observations with the whole group of 37 teachers (Table 6). Paired t-tests of teachers' pre- and post-program responses to students'

Insert Table 6 about here

work indicated significant gains in post-program scores, as well as gains on two of the four sub-scales—interpreting evidence of student thinking and responding to that evidence by providing formative feedback to students. These results appear to be at odds with Heritage and colleagues' finding that teachers completing the POWERSOURCE measure were more successful at interpreting evidence of students' thinking than they were at responding to that evidence by identifying what they might do next instructionally (Heritage et al., 2009)

Discussion

It takes a long time to learn to implement FA in a coherent and consistent fashion. Teachers spent the first year exploring practices relating to the individual critical aspects with their students. In many cases, this work involved teachers' rethinking how practices they already regularly engaged in might be reimagined to be better aligned with FA practices. For example, at the beginning of FACETS, when teachers shared LIs with their students, they used goals from their textbooks or the district/state standards. These were framed in language that students generally could not understand, and therefore were of little use to students as guides for monitoring their understanding as they worked on math tasks in class. Most of the time, teachers who did share LIs did so in a perfunctory way, announcing them at the beginning of the lesson and then moving on to the math tasks for the day. Over the course of the PD, teachers not only shared LIs more often, but they spent more time at the beginning of the lesson helping students understand the lesson goal, connecting it to earlier lessons and articulating the ways that they, and the teacher, would know that they've reached their goals. Teachers then returned to the LIs and SCs at strategic points in the lesson to keep the class focused on the purpose of their work and to ground feedback about students' work.

During the second year of the PD, teachers worked to weave the individual aspects they had been exploring into a more integrated and coherent whole. While they understood early on in the PD that the FA model articulated a cyclic process in which individual critical aspects followed one another, it was only as they tried to engage with these aspects in the classroom that they developed a deeper understanding of how the various critical aspects informed and motivated each other. This process was not without some frustration as teachers worked toward an image of FA that they could not yet achieve; though the *idea* of FA is easy to grasp, putting it into practice is a more challenging task. Through their efforts to bring their goals and their instruction into greater alignment, teachers developed both a more refined understanding of the conceptual underpinnings of FA as well as more FA-rich instruction. Teachers tell us that their classrooms are very different now than when they started FACETS; while teachers vary in the sophistication with which they implement different practices, all embed more FA in their teaching.

Even so, we, as well as the teachers themselves, would claim that they still have work to do to regularly and skillfully implement FA "full bore." The largest challenge for teachers is the development of student ownership and self-regulation. In the first year of PD, most teachers saw the role of teacher and the role of student as largely disconnected—the teacher should provide learning experiences and guidance for students and students should apply themselves to the work teachers assigned. While teachers talked about how much they wished their students would "own" their own work, most saw this ownership as a matter of individual interest and motivation. In the second year of PD teachers continued to focus on how the three other critical aspects could contribute to students taking increasing responsibility for monitoring and managing their learning. Most continued to articulate their expectations that students reflect on their own progress, but had not yet learned how to promote and scaffold students' ability to do so.

References

Association of Mathematics Teacher Educators, & National Council of Supervisors of Mathematics (n.d.). *Improving student achievement in mathematics through formative assessment in instruction.* Joint position paper, http://amte.net/sites/default/files/overview_amte_ncsm_position_paper_formative_assessment.pdf

Ball, D. L., & Forzani, F. M. (2009). The work of teaching and the challenge for teacher education. *Journal of Teacher Education, 60,* 497-510.

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology, 52*, 1-26.

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*. Phi Delta Kappan [Online Journal]. Retrieved September 12, 2007, from http://www.pdkintl.org/kappan/kbla9810.htm

Blackwell, L., Trzesniewski, K., & Dweck, C.S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development, 78*(1), 246-263.

Brown, N.J.S., Afflerback, P.P., & Croninger, R.G. (2014). Assessment of critical-analytic thinking. *Educational Psychology Review, 26*, 543-560. DOI 10.1007/s10648-014-9280-4

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research,* 65(3), 245-281.

Creighton, S.J., Rose Tobey, C., Karnowski, E., Fagan, E. (2015). *Engaging students in mathematics learning: Implementing formative assessment practices in the middle grades* (working title). Thousand Oaks, CA: Corwin.Creswell, J.W., & Plano Clark, V.L. (2007). *Designing and conducting mixed methods research*. Thousand Oaks, CA: Sage.

Duckworth, A. & Seligman, M., (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science, 16*(12), 939-944.

Heritage, M. (2008.) Learning progressions: Supporting instruction and formative assessment. Washington, DC: Council of Chief State School Officers. Retrieved from http://www.ccsso.org/publications/details.cfm?PublicationID=366. January 2, 2015.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-development inquiry. *American Psychologist, 34*(10), 906-911.

Heritage, H. M., (2010). *Formative assessment: Making it happen in the classroom*. Thousand Oaks, CA: Corwin.

Heritage, H., Kim, J., Vendlinksi, T., & Herman, J. (2009). From evidence to action: A seamless process in formative assessment? *Educational Measurement: Issues and Practice. 28*(3), 24-31.

Hill, H.C., Schilling, S.G., & Ball, D.L. (2004). Developing measures of teachers' mathematics knowledge for teaching. *Elementary School Journal, 105*(1), 11-30.

Hough, S., O'Rode, N., Terman, N., & Weisglass, J. (2007). Using concept maps to assess changes in teachers' understandings of algebra: A respectful approach. *Journal of Mathematics Teacher Education*, 10, 23-41.

McDonald, M., Kazemi, E., & Kavanagh, S.S. (2013). Core practices and pedagogies of teacher education: A call for a common language and collective activity. *Journal of Teacher Education*, *64*(5), 378-386.

Marchand, G.C., & Furrer, C.J. (2014). Formative, informative, and summative assessment: The relationship among curriculum-based measurement of reading, classroom engagement, and reading performance. *Psychology in the Schools, 51*(7), 659-676. DOI: 10.1002/pits.21779

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199-218.

Phelps, G. M., (2015, March 9). Personal email communication.

Popham, W.J., (2006). Phony formative assessments: Buyer beware! *Educational Leadership. 64*(3), 86-87.

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

Stake, R. E. (1995). *The art of case study research.* Thousand Oaks, CA: Sage.

Tobey, C. R., & Goldsmith, L. T. (2013). Developing self-regulating learners: The critical role of feedback. *ASCD Express, 8*(11).

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

Wiliam, D. (2011). *Embedded formative assessment.* Bloomington, IN: Solution Tree.

Wylie, E. C., Gullickson, A. R., Cummings, K. E., Egelson, P. E., Noakes, L. A., Norman, K. M., & Veeder, S. A. (2012). *Improving formative assessment practice to empower student learning.* Thousand Oaks, CA: Corwin.

Yin, R. K. (2003). *Case study research design and methods, 3rd ed.* Thousand Oaks, CA: Sage.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice, 41*(2), 64-70. DOI: 10.1207/s15430421tip4102_2

Figures and Tables

Figure 1: FACETS formative assessment cycle (Creighton et al., 2015)

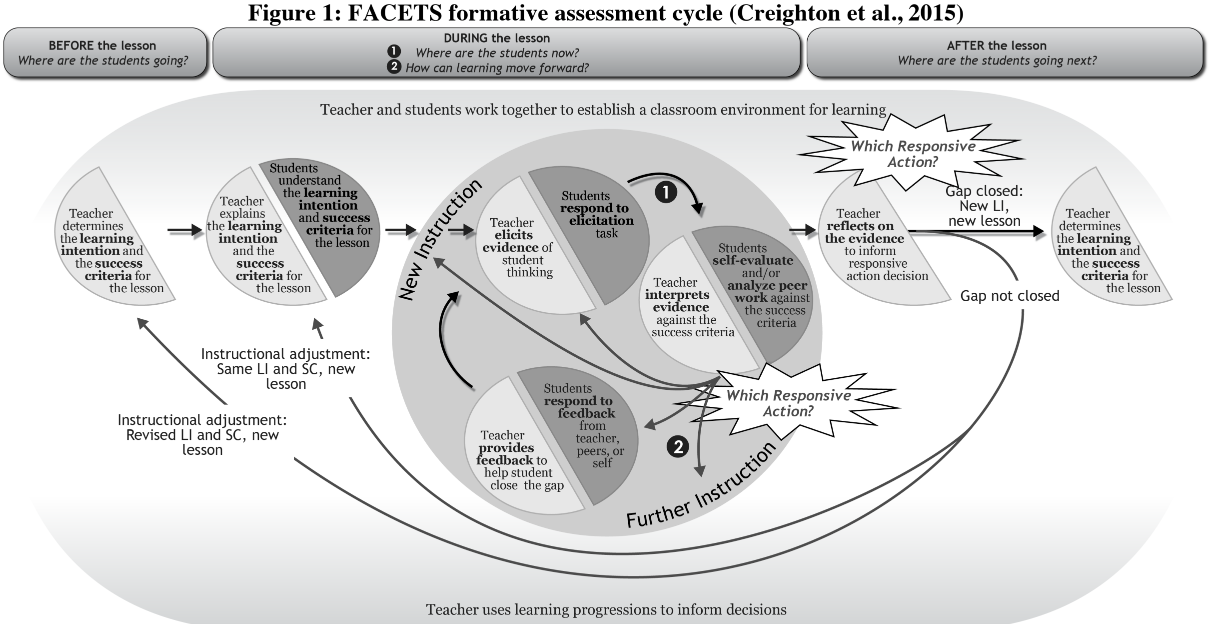


Figure 2: Map Scores over time

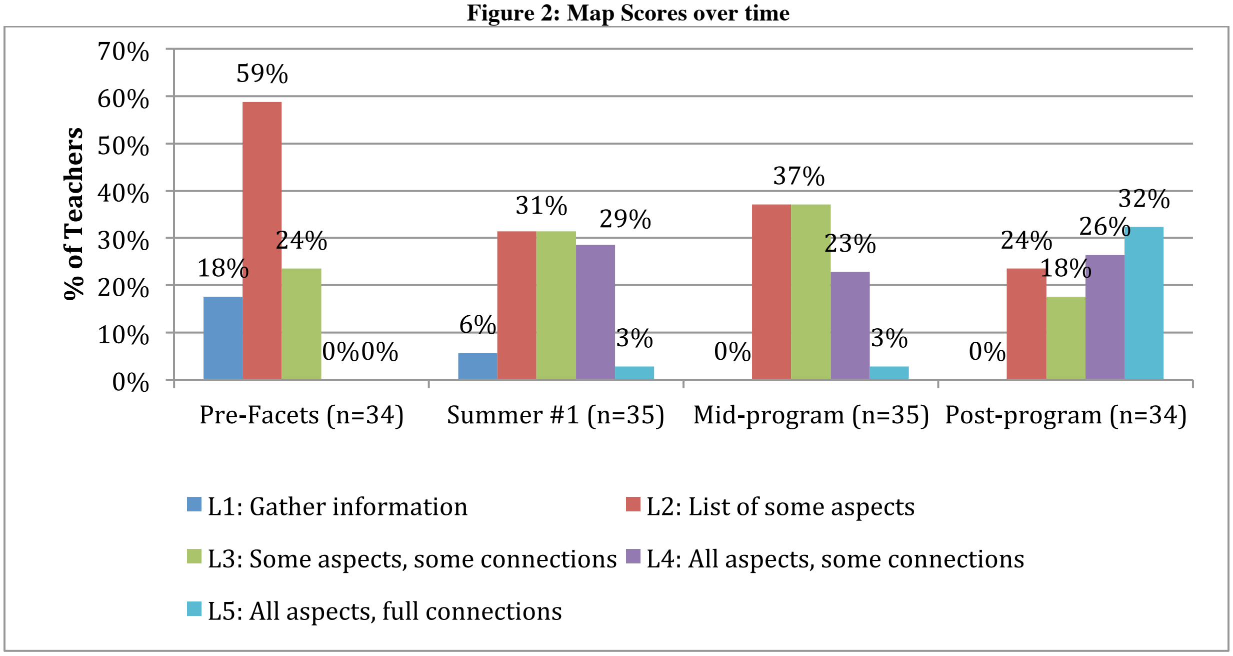
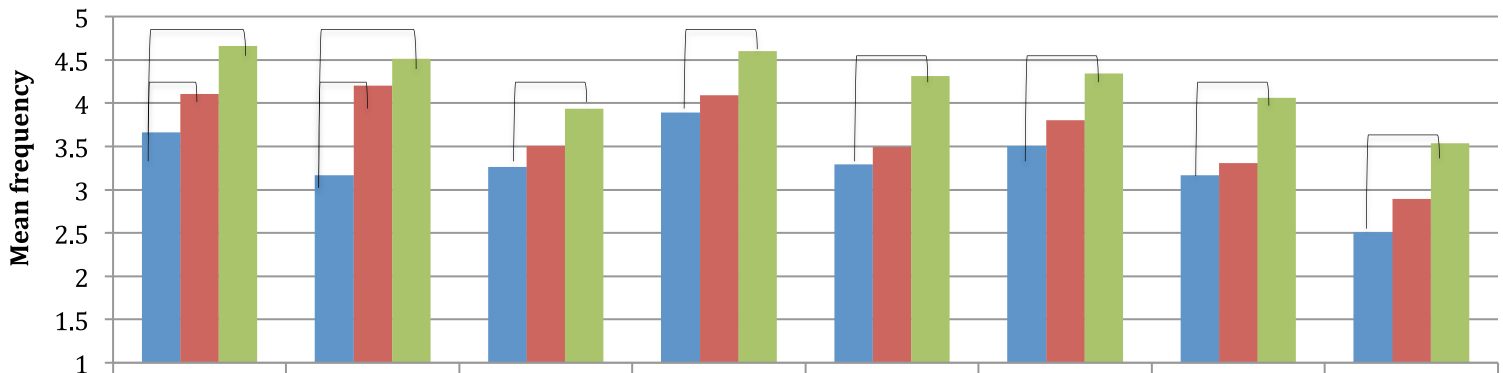


Figure 3. Mean frequency of implementing FA practices pre-, mid-, and post-program



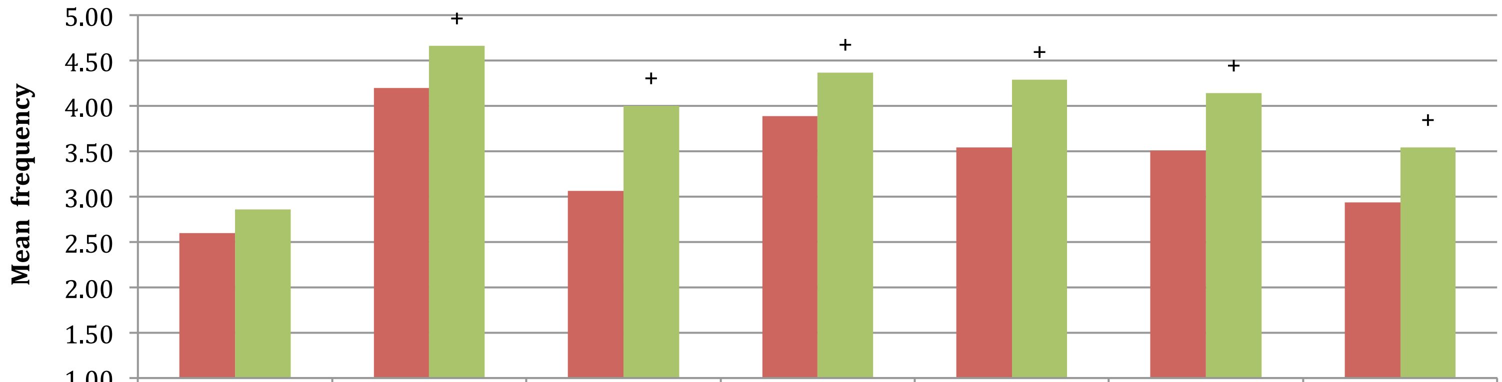
	Regularly implement FA practices	Write SC for an LI	ID where in lesson to get evidence of stu thinking re: LISC	Share LI/SC with stu at start of lesson	Get evidence of stu thinking re: LISC during class	Use evidence for instructional decisions	Give ffb to indiv students re: their work toward LISC	Encourage stus to monitor/ assess thinking re: LISC
Mean Pre	3.66	3.17	3.26	3.89	3.29	3.51	3.17	2.51
Mean Mid	4.11	4.20	3.51	4.09	3.49	3.80	3.31	2.89
Mean Post	4.66	4.51	3.94	4.60	4.31	4.34	4.06	3.54

Likert scale: 5= nearly every lesson, 4= about once a week, 3 =1-2 times per month, 2 =less than once per month, 1= do not currently do this

brackets from pre- to mid-program reflect significant differences (*p*<.006, Bonferroni correction for multiple comparisons)

brackets from pre- to post-program reflect significant differences (*p*<.003, Bonferroni correction for multiple comparisons)

Figure 3. Mean frequency of implementing FA practices, mid- and post-program only



	Use unit LPs to plan LI	Write LIs	ID where in lesson to get evidence of stu confidence re: LISC	Check that stus understand the lang/meaning of LISC	Refer stu to LISC at key points throughout lesson	Provide whole class ffb re: their work toward LISC	Give students chances to respond to ffb
Mean Mid	2.60	4.20	3.06	3.89	3.54	3.51	2.94
Mean Post	2.86	4.66	4.00	4.37	4.29	4.14	3.54

Likert scale: 5= nearly every lesson, 4= about once a week, 3 =1-2 times per month, 2 =less than once per month, 1= do not currently do this
+ : significant change from mid- to post-program (*p*<003, Bonferroni correction for multiple comparisons)

Table 1. PD components

	Year 1	Year 2
Summer institute	3 days (18 hours)	2 days (12 hours)
Academic Year Workshop	3 days (18 hours)	2 days (12 hours)
Learning Group meetings	6 meetings (15 hours)	7 meetings (17.5 hours)
Total number of hours	51 hours	41.5 hours

Table 2. Data sources

Source	**Construct**	**Administration**
Mathematical Knowledge for Teaching (MKT)	Math knowledge for teaching (implementation of FA)	Pre-program Post-program
Concept Maps & Comparisons	Knowledge of FA	Pre-program End of 1st Summer Institute Mid-PD (End of year 1) Post-program
Student Work Survey	Implementation of FA	Pre-program Post-program
Classroom Practices Survey	Knowledge of FA Implementation of FA	Pre-program Mid-program (end of year 1) Post-program
PD Artifacts (Blog posts, FA plans)	Knowledge of FA Implementation of FA	Ongoing
Observations	Implementation of FA	Spring Year 1 Fall Year 2 Spring Year 2

Table 3. Teacher background information (n=37*)

Undergrad math major/ minor	Grad degree math ed	K–5 teaching credential	6–8 teaching credential	Math teaching credential	Years teaching math
13 (35%)	9 (24%)	28 (76%)	33 (89%)	16 (45%)	10 yrs range: 0-29 yrs

*Technical issues with logging web-based responses resulted in missing records for three teachers

Table 4. Grade levels taught

5th	6th	7th	8th	Multiple grades
6 (15%)	15 (31%)	2 (9%)	11 (28%)	6 (15%)

Table 5: Number of Teachers Identifying Characteristics of FA in Pre and Post-program Definitions (n=37)

Gather evidence (Critical Aspect)	31 (84%)	30 (81%)
Ongoing process	27 (73%)	24 (65%)
Inform teaching	20 (54%)	26 (70%)
Role of LI & SC (Critical Aspect)	8 (22%)	26 (70%)
Involves feedback (Critical Aspect)	4 (11%)	19 (51%)
Student role (Critical Aspect)	2 (5%)	19 (51%)

Table 6: Mean pre- and post-program Student Work Analysis scores

		Mean (SD)	*t*	*p*
TOTAL SCORE	Post Pre	23.03 (3.06) 18.76 (5.23)	4.34	.000*
Identify LISC	Post Pre	3.79 (1.27) 3.46 (1.06)	1.28	.209
Elicit evidence of student thinking	Post Pre	4.18 (.92) 3.27 (1.31)	3.17	.002*
Interpret evidence of student thinking	Post Pre	9.06 (1.03) 8.15 (2.03)	2.27	.030
Provide formative feedback	Post Pre	6.00 (1.77) 3.90 (2.04)	4.86	.000*

α=.01, Bonferroni correction for multiple tests