

April 2018

Evaluating and Supporting Teacher Practice of Formative Assessment: Assessing Posing, Pausing, and Probing Moves

Carrie Holmberg
San Jose State University

Brent Duckor
San Jose State University

Follow this and additional works at: https://scholarworks.sjsu.edu/second_ed_pub



Part of the [Educational Assessment, Evaluation, and Research Commons](#), and the [Teacher Education and Professional Development Commons](#)

Recommended Citation

Carrie Holmberg and Brent Duckor. "Evaluating and Supporting Teacher Practice of Formative Assessment: Assessing Posing, Pausing, and Probing Moves" *American Educational Research Association (AERA) Annual Meeting* (2018). <https://doi.org/10.302/1315545>

This Presentation is brought to you for free and open access by SJSU ScholarWorks. It has been accepted for inclusion in Faculty Publications by an authorized administrator of SJSU ScholarWorks. For more information, please contact scholarworks@sjsu.edu.

Evaluating and Supporting Teacher Practice of Formative Assessment: Assessing Posing, Pausing, and Probing Moves

PURPOSES

The purpose of the study was to design and pilot a performance-based assessment of teachers' formative assessment practices related to questioning "moves" in the context of multilingual middle school mathematics classrooms with high percentages of socioeconomically disadvantaged students. Research has long supported that teachers' engagement in formative assessment (FA) can powerfully impact student outcomes (Black & Wiliam, 1998; Hattie, 2009, 2012; Wiliam, Lee, Harrison, & Black, 2004) and can decrease achievement gaps between high- and low-achieving students. Scholars consider *teacher questioning* a high-leverage practice (Alexander, 2006; Ball, Sleep, Boerst, & Bass, 2009; Palinscar & Brown, 1984; Schoenfeld, 2013) fundamental to eliciting data about student learning.

Heritage & Heritage (2013) suggest that to make productive pedagogical moves within students' zones of proximal development teachers must be adept at posing, pausing, and probing practices to better elicit and develop learners' emergent understanding. Yet how teachers develop skill in questioning within FA practice is still largely anecdotal and impressionistic. This is despite the recognized potential of the inherently dialogic process of formative assessment between teacher and students to provide linguistic-minority and high-poverty students in particular with multiple, authentic opportunities to develop academic language while engaging meaningfully in discipline-specific practices (Abedi, 2010; Linqianti, 2104; Ruiz-Primo, Solano-Flores, & Li, 2014).

We know teachers' mathematical teaching practices can foster more equitable learning spaces (Boaler, 2002; Boaler & Sengupta-Irving, 2016) and outcomes (Boaler & Staples, 2008). Given this, and the role students' achievement in middle school mathematics can play in college-going trajectories (Balfanz, 2009), the importance of examining (and articulating) teachers' developing competencies in questioning as a formative assessment practice *when the goal is improving feedback to teachers in professional development efforts*--and doing so within the context of multilingual, middle school mathematics classes of underprivileged students--becomes clear.

THEORETICAL FRAMEWORK & RESEARCH QUESTIONS

The study employs the "FA moves" framework (Author B, 2014), which conceptualizes formative assessment as a dynamic, pedagogical process of moves between teacher and students. It helps teachers to learn more about students' understandings and to productively respond to those understandings (not merely "misconceptions" or "wrong" answers) during instruction. It requires acts of planning, instructing, and reflecting on soft data to make better decisions. It places a premium on feedback loops in classroom talk, building up of repertoires of auditory and verbal skills, and providing instructional space for students to use academic language and register during lessons.

Using a teacher learning progressions (Author B & A, 2017) framework, we identify seven moves accessible to novices and useful to more expert teachers across content areas. These FA

moves include (Fig. 1): *priming, posing, pausing, probing, bouncing, tagging, and binning*. Each move ties together instruction and assessment practices by requiring evidence of student engagement and visible routines for eliciting thinking. At the core of the FA moves framework are deep, sustained routines related to questioning: posing questions, pausing for think time, and probing on initial student responses to invite elaboration.

This moves-based framing of formative assessment practice is rooted in other theoretical frameworks. It links to Sadler's (1989) focus on feedback loops in classroom discourse and builds on Wiliam's (2007) framework of "key strategies" of assessment for learning to provide "'window[s] into [student] thinking'" (p. 1069). In mathematics, it aligns with Ball and associates' (2009) concept of high-leverage practices which enhance opportunities for mathematical reasoning and discussion.

The purpose of this empirical study was to design and pilot a performance-based assessment of three *facets* of teachers' practice of formative assessment (FA)--*planning* for, *enactment* of, and *reflection* on practice of FA--along three *dimensions* of FA practice hypothesized: teachers' (a) *posing* of questions, (b) *pausing* to foster equity of participation and quality of response, and (c) *probing* student thinking. The study asked:

RQ1. Is there a continuum of practice from novice to more expert knowledge and skills in teachers' posing questions, pausing, and probing student thinking in middle school mathematics?

RQ2. How can levels of practice be represented along a generalized continuum in mathematics teaching and learning at the classroom level?

RQ3. Can teacher practices of posing, pausing, and probing along any or all of these proposed dimensions of practice be reliably and validly evaluated?

To address them, we designed an evidence-based performance assessment suite (See, e.g., Pellegrino, Chudowsky, & Glaser, 2001) that included learning progression-oriented construct maps, performance task and interview protocols for video-based lesson analysis, and scoring guides to evaluate teachers on the proposed continua of *posing-pausing-probing* practice. Using the construct mapping approach (Wilson, 2005), the goal was to also meet minimal standards for weighing evidence of accomplished questioning practice with respect to validity and reliability (AERA, APA, NCME, 2014). The construct mapping (CM) approach was initially designed for supporting K-12 content assessment, but can be generalized to other populations including teachers and areas related to their learning and professional practice (see, e.g., Wilson & Sloane, 2000).

METHODS

The CM approach to assessment design requires qualitative--not merely quantitative--investigation (Wilson, 2005). Empirical inquiry through a comparative, multiple-case study approach was employed since a goal was to investigate a contemporary phenomenon in depth in its real-world context (Yin, 2014, p. 16). Case boundaries were defined by one "cycle of inquiry" encompassing teachers' planning, enacting, and reflecting on their instruction through the lens of the FA moves. Each participant planned for, enacted, and reflected on a single lesson.

Subjects were six seventh and eighth grade mathematics teachers (n=6) drawn from three schools in a high-needs Pre-K-8 school district in California. Class size averaged 30 students. On state tests taken two years prior, 18, 41, and 42 percent of students at the schools met or exceeded standards in mathematics. Percentages for English Language Arts were 25, 43 and 27. See table 1 for school demographics.

Subjects volunteered to participate. To minimize self-selection bias, we employed district-wide selection criteria. See table 2 for sample characteristics. Four had taught more than 15 years, two fewer than three. One taught mathematics at community college. All had 2-4 years experience teaching middle school students with the curriculum.

DATA SOURCES AND ANALYSIS

Data were collected for each subject over 4-6 weeks December 2016-March 2017 (participation was staggered) from each phase of the cycle of inquiry (planning, enacting, and reflecting). Primary data sources included responses from pre- and post-study surveys, transcriptions of audio recordings of participant responses to two in-person interviews each, lesson plans and planning artifacts, video recordings (and transcriptions) of enacted lessons, lesson artifacts and field notes. Analytic memos augmented analysis.

Evidence of planning was prompted through participants' use of a common lesson planning template and 19 interview and survey items targeting planning for posing, pausing, and probing. To collect evidence of enactment, each lesson was video recorded with two cameras: a free-standing, stationary camera captured a wide-angle perspective; another captured teacher-student interactions during group work.

To facilitate collecting data on teacher reflection, a video-stimulated recall (VSR) protocol was conducted 1-10 days post lesson. The timeline allowed participants to review the recording of their lesson and (optionally) choose a clip for reflection. The VSR protocol prompted teachers to "unpack" two or three 3-5-minute video clips. The protocol encouraged participants to press "pause" and provide context, intentions, explanations, and ideas for "next steps" (i.e., the reflection facet of study).

Analysis of data served three goals:

1. to explore accuracy of--and improve--hypotheses of teacher progression in formative assessment practice articulated at the study's outset via *construct maps/dimensions* of teacher posing, pausing, and probing
2. to locate on the continua of practice articulated each participant's *level* of posing, pausing, and probing based on evidence from *performance tasks* such as FA lesson planning; and
3. to determine, according to a *scoring guide/outcome space*, the level of skill based on three facets (participant's planning, enacting, and reflecting on a particular dimension).

Deductive and emergent coding schemes (Miles, Huberman, & Saldaña, 2014) were used for establishing performance levels and analyzing episodes of instruction. Triangulating classroom

transcriptions, interview responses, responses to tasks (e.g., lesson planning template), and survey responses served to develop a more holistic sense of trajectories of questioning practices.

RESULTS

The results reported in this study primarily address RQ1 and RQ2 and data analysis goals.

The study began with research-based hypotheses of teacher proficiency in posing, pausing, and probing represented via three construct maps, one for each “move” (construct) under investigation (See Figure 2) that described five qualitatively distinct levels of teacher practice, ordered in a continuum of development (see, e.g., Biggs & Collis, 1982). Each construct map identified practices characterized as (1) prestructural, (2) unistructural, (3) multistructural, (4) relational, and (5) extended abstract. Comparison of empirical data to these hypotheses demonstrated that the three middle levels (2, 3, and 4) accurately described evidence of teacher practice elicited by the pilot of the performance-based assessment. The only improvements to the hypotheses the empirical data compelled were to incorporate teacher actions to foster student-to-student probing into the probing construct map.

No responses elicited were consistent with descriptions of level 5. Little empirical evidence elicited matched descriptions of level 1 performances: two teachers’ planning for probing, one teacher’s reflecting on probing, and one teacher’s reflecting on pausing matched descriptions for level 1 proficiency in probing and pausing practice. Given subjects’ years of teaching experience this corresponded with our expectations about *pre-structural* practice, which is more likely to be observed in pre-service populations.

The study also found levels of teacher questioning practice--operationalized as posing, pausing, and probing moves--could be represented along a generalized continuum in the context of middle school mathematics classroom instruction. Based on evidence from three performance tasks in which teachers planned, taught and reflected on a lesson, participant responses enabled us to locate their demonstrated levels of proficiency in posing, pausing, and probing. Subjects’ proficiencies were located for each dimension generally, and by facet (e.g., *planning* for posing, *enacting* pausing, *reflecting* on probing). Subjects’ practices by dimension generally were located in the mid-range of the continua, at the unistructural and multistructural levels.

Locations of subjects’ questioning practice by facet ranged from level 1 (prestructural) proficiency (exhibited in one teacher’s planning for probing) to level 4 (relational) proficiency in posing (demonstrated in one teacher’s posing). Subjects’ locations on the continua of practice for posing, pausing, and probing generally and by facet (planning, enacting, reflecting) demonstrated good coverage of the construct maps (See table 3). Generalized scoring guides aligned with each of the construct maps were employed with an intra-rater agreement protocol.

Based on video evidence and direct observation, we found that teacher enactment of posing, pausing, and probing moves were located at higher levels of proficiency during contexts of small group instruction than during whole class instruction.

SIGNIFICANCE

This work advances the development of an analytic framework and common language that researchers, teacher educators and practitioners can use to describe an “underlying grammar of practice” of formative assessment-driven teaching. Grossman, McDonald and others (2008) have asserted that researchers, as well as novice teachers, “suffer the consequences” of a field that “still lacks a framework for teaching with well-defined common terms for describing and analyzing teaching” (p. 186).

This study adds to the existing knowledge base on the use of teachers’ reflective practices in guided contexts (e.g., lesson study and video clubs) to advance the skills required to bring about more powerful mathematics question and elaboration-focused discussions during instructional segments. By emphasizing research-based assessment “moves” it puts the focus of teacher and student growth in the context of multilingual low-SES classrooms that have struggled to plan, enact and reflect on equity of voice and participation (Author, B., Author, A., & Rossi-Becker, 2017).

Our empirical study also sharpens the discussion on finer-grained observation tools available to administrators, instructional coaches, teachers in professional learning communities, and others interested in the development of teachers’ practices of assessment for learning. The FA moves framework was purposefully constructed to focus strongly on how teachers progress along various pathways to becoming formative assessors, and this study offers pictures of individual trajectories of teacher practice in the field of formative assessment for high-needs schools.

FIGURES AND TABLES

List of Figures and Tables

Figure 1: FA Moves Wheel

Figure 2: Posing Construct Map

Table 1: School Demographics: Student Enrollment by Group

Table 2: Selected Sample Demographic Characteristics

Table 3: Scores for Planning, Enactment, and Reflection Facets of Posing, Pausing, and Probing Dimensions of Formative Assessment Practice

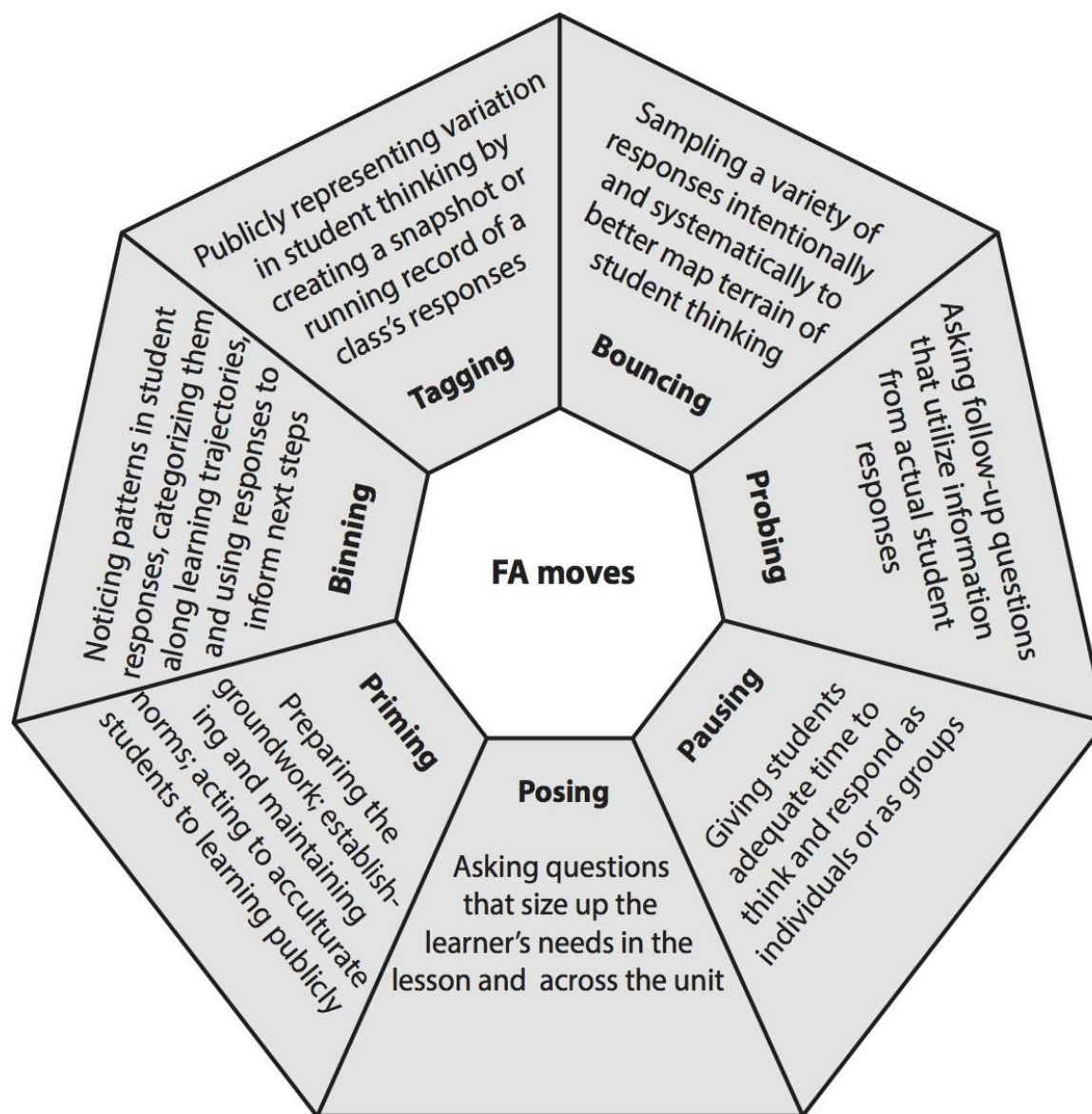


Figure 1: FA Moves Wheel (Author B. et al., 2016)

Construct Map: Teacher Posing of Questions

High	<i>Respondents who integrate relevant features of the context for learning with multiple important purposes for questions (e.g., promoting meta-cognition). They pose questions that size up the context for learning in ways that reflect knowledge of students' development, interests, needs re: learning target(s), and present understandings. They pose questions that relate to the lesson and</i>	Integrative posing (Extended Abstract) 5	<i>Responses to items/tasks indicate flexibility in posing to adjust to students' learning edges in real-time in relation to learning goals. Questions posed leverage a range of student responses (including student questions) in ways that elicit evidence of having furthered students' present understandings</i>
------	---	--	--

	<p>the unit plan and larger essential questions/big ideas of the discipline.</p> <p><u>They plan questions</u> that reveal explicit anticipation of where students may/are likely to get stuck or have misconceptions. They plan questions that serve to provide evidence in helping teachers decide which of a few to several specific (and expressed) decisions they might make that are contingent upon students' responses to these questions (e.g., they plan "hinge" questions and post-hinge question pathways for instruction). They plan in ways that encourage student questions to be springboards for discussion. They plan questions that reflect a balance between content-centered instruction with student-centered instruction.</p> <p><u>They tend to enact lessons</u> that display several ways student responses can be used to further students' own and other students' learning regarding the lesson target. They tend to enact lessons that feature questions that reflect a sensible balance in addressing a variety of learners' needs.</p> <p><u>They are able to reflect</u> on how questions posed functioned to elicit evidence of student understanding in relation to lesson objectives/target(s) of instruction.</p>		<p>in relation to the lesson target and/or essential question/big idea of the discipline. Responses to items/tasks show that respondent has anticipated student pit stops and bottlenecks typical of learning progression of concept/skill/understanding.</p> <p>Observation of teaching shows student responses being used in a variety of ways, including changing the direction of the lesson and/or pausing an activity.</p>
	<p><i>Respondents who</i> demonstrate flexibility in their questioning. They demonstrate an awareness of the variety of purposes of their questions and the need to match kinds of questions to specific purposes.</p> <p><u>They plan</u> a variety of questions designed to elicit a wide range of responses, including misconceptions and "unorthodox" responses. They plan carefully sequenced repetition of key questions. They plan supports/scaffolds for questions.</p> <p><u>Then tend to enact</u> lessons in which activities and pacing clearly reflect teacher decisions that are contingent upon student responses to questions posed about the learning target.</p> <p><u>They are able to reflect</u> on perceived effects of changing questions and/or questioning strategies. They are able to suggest several "next steps" likely to support improved posing. They do so from many perspectives</p>	<p>Flexible posing (Relational) 4</p>	<p><i>Responses to items/tasks</i> indicate posing of <i>how</i> and <i>why</i> questions and questions from a mix of Webb's DOK or other taxonomic levels (e.g. Bloom's, Costa's).</p> <p>Observation of teaching will likely show changing questioning strategies in response to student(s) response(s). Observation of teaching may show students playing significant roles in posing questions. Observation of teaching will show many questions that serve to highlight connecting students' prior knowledge and experiences with present efforts to engage with and "reach" the learning target.</p>

Low	and with specificity.		
	<p><i>Respondents whose purposes for questioning seem to be to get students to say what the respondent-as-teacher is thinking (rather than eliciting from students a range of responses, including unknown responses, responses surprising to the teacher).</i></p> <p><u>They plan questions</u> they consider checks for understanding of the lesson's objective.</p> <p><u>Then</u> tend to enact lessons with high percentages of close-ended questions. They tend to enact lessons in which scenarios arise where students are expected to guess what the teacher is thinking, even when doing so appears more a hindrance to than a help regarding students' advancement toward the learning target.</p> <p><u>They are able to reflect</u> on several aims of improving posing. Their reflection includes specific suggestions for alternate poses to try.</p>	<p>Constrained posing (Multistructural) 3</p>	<p><i>Responses to item/tasks indicate posing a high percentage of, or posing only, what/when/where, fact recall, and lower-level questions (on Webb's DOK, Bloom's taxonomies etc.). Questions planned connect to learning target.</i></p> <p>Observation of teaching shows questions posed as checks for understanding procedures and concepts tied to the learning target. Observation of teaching shows questions that seek to elicit students' prior knowledge. Observation of teaching reveals questions posed seldom elicit a wide range of responses.</p>
	<p><i>Respondents who demonstrate through their questioning a primary focus on orchestrating student behavior, not necessarily learning (activity-based posing). They may not be able to make student thinking visible through questions they pose.</i></p> <p><u>They plan questions</u> that do not reveal clear priorities in the purposes of posing questions. As they plan, they experience challenges in deciding what content is most important to ask about and when.</p> <p><u>They tend to enact</u> teacher-centered lessons that do not reflect an underlying pedagogical structure dependent upon student responses to curricular content.</p> <p><u>They are able to reflect</u> on benefits that might accrue from using a questioning scheme.</p>	<p>Posing to manage (Unistructural) 2</p>	<p><i>Responses to items/tasks indicate posing to manage/control students, e.g., "Do you have a pencil? Are your books open to page 39?" Planned questions do not express recognizable coherence or organizing principle.</i></p> <p>Observation of teaching and questions reveal imbalance of focus between activity/behavior and learning target.</p>
	<p><i>Respondents who give directions to students and whose actions can be interpreted as attempting to pour content into students' minds without eliciting from the students where their current understandings are.</i></p> <p><u>They may plan</u> questions not well-crafted to elicit evidence of student understanding in</p>	<p>Pre-posing (Prestructural) 1</p>	<p><i>Responses to items/tasks indicate no questions aligned with lesson target are posed by the teacher. Planned questions may or may not align with lesson target(s).</i></p> <p>Observation of teaching may show random or arbitrary</p>

	<p>relation to instructional goal(s).</p> <p><u>They tend to enact</u> lessons that do not invite or incorporate students' prior knowledge.</p> <p><u>They are able to "reflect"</u> through descriptions of their instruction that do not push to analysis.</p>		questions.
--	--	--	------------

Figure 2: Posing Construct Map

Table 1

School Demographics: Student Enrollment by Group

	Chavez (T.E. = 516)	Kimm (T.E. = 723)	Sierra (T.E. = 776)
Group	% of T.E.	% of T.E.	% of T.E.
Racial identification			
Black or African American	1.4	1.2	2.4
American Indian or Alaska Native	0.2	0.1	0.1
Asian	10.3	42.7	36.3
Filipino	1.2	5.2	1.9
Hispanic or Latino	85.9	45.9	56.3
Native Hawaiian or Pacific Islander	0.2	0.1	0.4
White	0.8	1.5	0.9
Two or More Races	0.2	3.2	1.5
Socioeconomically disadvantaged	95.9	81.7	82.9
English learners	43.4	46.7	20.9
Students with disabilities	11	8.6	8.4
Foster youth	1	0.3	0.4

Note. T.E. = total enrollment.

Table 2

Selected Sample Demographic Characteristics

	Gr.	Yr. Tchg	Curric.	Yr. tchg w/curric.	Cred.	Bachelor's	Native Eng.
LH	8	20 ^a	CPM	2	MS	dance	Y
LY	7	3 ^b	CPM	2	MS	electrical engineering	N
IM	7	16	CPM	4	MS	liberal studies	Y
LO	8	2	CPM	2	MS	child develop.	N
MM	8	18	MVP	2	MS & SS	mathematics	N
SR	8	30	MVP	4	SS	mathematics	N

Note. MS = multiple subject; SS = single subject; CPM = College Preparatory Mathematics; MVP = Mathematics Vision Project.

^aAll ELA or elementary (MS). ^bSecond career.

Table 3

Scores for Planning, Enactment, and Reflection Facets of Posing, Pausing, and Probing Dimensions of Formative Assessment Practice

Participant	Dimensions and Facets of Formative Assessment Practice								
“Leila”									
Dimension	Posing			Pausing			Probing		
Overall	3			2			3		
Facet	P	E	R	P	E	R	P	E	R
Score each facet	3+	3, 3	3	3-	2, 2	1+	3-	3,3+,4	3+

“Aaron”

Dimension	Posing			Pausing			Probing		
Overall	3			3			3		
Facet	P	E	R	P	E	R	P	E	R
Score each facet	3-	3, 4	3	2-	3-, 4-	3	1+	3-, 3	3+

“Jessica”

Dimension	Posing			Pausing			Probing		
Overall	3			3-				2-	
Facet	P	E	R	P	E	R	P	E	R
Score each facet	3	3, 3+/4-	3	2-	3, 3	2+	1	2,2+	1+

“Lavinia”

Dimension	Posing			Pausing			Probing		
Overall	3			3				3-	
Facet	P	E	R	P	E	R	P	E	R
Score each facet	3	3+, 3+	2+	2	3,3	3+	2	2, 3, 3+	3+

“Eliza”

Dimension	Posing			Pausing			Probing		
Overall	3 -			2				2	
Facet	P	E	R	P	E	R	P	E	R
Score each facet	3	2+, 3-	2+	2-	2, 2	2+	2-	2-,3-,2-	2+

“Selena”

Dimension	Posing			Pausing			Probing		
Overall	3+			3				3-	
Facet	P	E	R	P	E	R	P	E	R
Score each facet	3	4, 4	3+	2	3,3+	3+	3/3-	2+, 2+,3-	3

Note. P = planning; E = enactment; R = reflecting.

REFERENCES

- Author, B. (2014a). Formative assessment in seven good moves. *Educational Leadership*, 71(6), 28-32.
- Author, B., & Author, A. (2017). *Mastering formative assessment Moves: 7 high-leverage practices to advance student learning*. Alexandria, VA: ASCD.
- Author, B., Author, A., & Rossi Becker, J. (April, 2016). *Binning for equity and access: Formative assessment-focused teacher professional development for middle school mathematics classrooms*. Poster presented at 100th Annual Meeting of the American Education Research Association, Washington, D.C.
- Author, B., Author, A., & Rossi-Becker, J. (February, 2017). Making moves: Formative assessment in mathematics. *Mathematics Teaching in the Middle School*, 22(6), 334-342.
- Abedi, J. (2010). Research and recommendations for formative assessment with ELLs. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 181–197). New York: Routledge.
- Alexander, R. J. (2006). *Towards dialogic teaching: Rethinking classroom talk* (3rd ed.) Cambridge: Dialogos.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: Author.
- Balfanz, R. (2009). “Putting Middle Grades Students on the Graduation Path.” *Policy and Practice Brief*. Westerville, OH: National Middle School Association. Retrieved from http://www.amle.org/portals/0/pdf/articles/Policy_Brief_Balfanz.pdf
- Ball, D. L., Sleep, L., Boerst, T. A., & Bass, H. (2009). Combining the development of practice and the practice of development in teacher education. *The Elementary School Journal*, 109(5), 458-474.
- Biggs, J. B., & Collis, K. F. (1982). *Evaluation the quality of learning: The SOLO taxonomy*. New York: Academic Press.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi delta kappan*, 86(1), 8-21.
- Black, P. & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 81(2), 139-148.

- Boaler, J. (2002). Learning from teaching: Exploring the relationship between reform curriculum and equity. *Journal For Research In Mathematics Education*, 33(4), 239-258.
- Boaler, J., & Sengupta-Irving, T. (2016). The many colors of algebra: The impact of equity focused teaching upon student learning and engagement. *Journal Of Mathematical Behavior*, 41, 179-190. doi:10.1016/j.jmathb.2015.10.007
- Boaler, J., & Staples, M. (2008). Creating mathematical futures through an equitable teaching approach: The case of Railside School. *Teachers College Record*, 110(3), 608–645.
- Grossman, P., & McDonald, M. (2008). Back to the Future: Directions for Research in Teaching and Teacher Education. *American Educational Research Journal*, 45(1), 184-205.
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York: Routledge.
- Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. New York: Routledge.
- Heritage, M. and Heritage, J., 2013. Teacher questioning: The epicenter of instruction and assessment. *Applied Measurement in Education*, 26(3), pp.176-190.
- Heritage, M., Walqui, A., & Linqunti, R. (2013). *Formative assessment as contingent teaching and learning: Perspectives on assessment as and for language learning in the content areas*. Paper presented at the Annual Meeting of American Education Research Association, San Francisco, CA.
- Linqunti, R. (2014). Supporting Formative Assessment for Deeper Learning: A Primer for Policymakers. Washington, DC: CCSSO. Retrieved from <http://www.ccsso.org/Documents/Supporting%20Formative%20Assessment%20for%20Deeper%20Learning.pdf>
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook*. Thousand Oaks, CA: Sage.
- Palinscar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, 1(2), 117-175.
- Pellegrino, J. W., Chudowsky, N., & Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.
- Ruiz-Primo, M. A., Solano-Flores, G., & Li, M. (2014). Formative assessment as a process of interaction through language: A framework for the inclusion of English language learners In C. Wyatt-Smith (Ed.) *Designing assessment for quality learning* (pp. 265-282). Springer Netherlands.

- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2), 119-144.
- Schoenfeld, A. H. (2013). Classroom observations in theory and practice. *ZDM*, 45(4), 607-621.
- William, D. (2007). Keeping learning on track: Classroom assessment and the regulation of learning. In F.K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning: A project of the national council of teachers of mathematics* (pp. 1053–98). Charlotte, NC: Information Age Publishing.
- William, D., Lee, C., Harrison, C., & Black, P. (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education: Principles, Policy & Practice*, 11(1), 49-65.
- Wilson, M. (2005). *Constructing measures: An item response theory approach*. New York: Psychology Press.
- Wilson, M., & Sloane, K. (2000). From principles to practice: An embedded assessment system. *Applied Measurement in Education*, 13(2), 181-208.
- Yin, R. K. (2014). *Case study research: Design and methods*. (Fifth ed.). Thousand Oaks, CA: Sage.