A Design-Based Process in Characterizing Experienced Teachers' Formative Assessment Enactment in Science Classrooms

Hannah Sevian and Vesal Dini

Introduction

The initiation-response-evaluation (IRE) discourse pattern in classrooms is pervasive and well documented (Mehan 1979). In this pattern, the teacher initiates a question, students respond, and then the teacher evaluates this response. In spite of extensive science education reform efforts over decades, and the implementation of a wide variety of highly developed hands-on science curricula, studies have repeatedly shown that the IRE pattern persists and classroom science learning remains largely procedural without challenging students to make sense of what they are learning (Banilower et al. 2013; Roth and Garnier 2007). Compounding this problem is that many experienced science teachers consider that their teaching is well aligned with high levels of inquiry in their classrooms. Teachers report engaging their students in questioning, modeling, and communicating evidence several times per month, yet observations reveal that the teachers' definitions of inquiry vary and they often map their classroom practices onto vague notions of inquiry activity (Capps et al. 2016). In fact, it is rare that sense-making activities for students get connected to laboratory-based activities in the classroom and discourse in classrooms that promotes such sense-making is even more rare (Weiss et al. 2003). These persistent problems suggest that teachers could benefit from practical tools to help them attend to sense-making in science classrooms, with specific guidance for experienced teachers to reflect on their own discourse practices.

The aim of this chapter is to share our design-based research approach to addressing the following problem of practice: *There is a lack of practical guidance for science teachers in enacting formative assessment (FA) to support students' sense-making*. In an urban partnership that includes middle and high school science

H. Sevian (✉) · V. Dini
College of Science and Mathematics, University of Massachusetts Boston, Boston, MA, USA
e-mail: hannah.sevian@umb.edu

E. McLoughlin et al. (eds.), *Bridging Research and Practice in Science Education*, Contributions from Science Education Research 6,
https://doi.org/10.1007/978-3-030-17219-0_20

teacher leaders, school district administrators, and science education researchers (Szteinberg et al. 2014), we converged on this problem of practice that combines the perspectives of multiple stakeholders. In the statement of this problem of practice, we intend *practical* to mean based on the practices of experienced science teachers, *guidance* to mean that it is clearly defined and easy for classroom teachers to use, *enacting* to mean that it is about options that teachers can take in their instructional decisions, *support* to honor the teacher as agent in achieving the goals of learning for students, and *sense-making* to recognize that student learning benefits when students are protagonists in their learning. In the work reported here, we focus on the process among teachers and researchers that led to a resource to address our problem of practice. The resource provides guidance for experienced teachers in examining their own classroom FA practice as well as an instrument for researchers to study science teachers' FA enactment.

Design-Based Research Approach

Through this problem of practice, we seek to address important problems facing teachers working in the complex environments of science classrooms. We therefore followed the approach of design-based research (DBR), which was developed "to address theoretical questions about the nature of learning" situated in real-world contexts and "derive research findings from formative evaluation" (Collins et al. 2004: 16).

DBR assumes the entanglement of the design of a learning environment—which may take the form of an instructional approach, type of assessment, or learning activity (Anderson and Shattuck 2012)—with the development of a related learning theory (Brown 1992). DBR stipulates that such design and development should take place in naturalistic settings and be carried out in an iterative process of designing, enacting, analyzing, and redesigning (The Design-Based Research Collective 2003). In a DBR process, not only does this iterative process of design evolve, but there is also a major product goal of communicating how enactments are connected to outcomes of interest in the particular context under study. Our explanations here are crafted for the purpose of guiding fellow practitioners.

We extended our DBR approach to create what Bereiter (2014) called principled practical knowledge (PPK), which is systematic, coherent, and explanatory, but its main purpose is practical guidance. Our aim in this was to "increase the generalizability of knowledge produced through design work and provide a ladder leading to sometimes radical design improvement" (Bereiter 2014: 1). From analyzing researchers' DBR processes, Bereiter identified three stages that help put the production of PPK into context: (1) a practical observation emerging from the DBR experience; (2) a reasonably coherent and generalizable explanation of what has been observed, which may still be limited; and (3) basic research to form results from the second stage into theory. Bereiter points to the second stage as representing PPK, "which is both a foundation for further design advances and a stimulus for theoretical research" (p. 11). We describe our DBR process in terms of Bereiter's stages. After an initial description of Stage 1 (practical observation), we concentrate on the second cycle that took place in Stage 2 (generalizable explanation), because it points toward the integration of results into theory.

Practical Observation

The current team includes five grades 6–12 science teacher leaders from different schools in a large urban school district, two school district administrators (the director and associate director of the district's science, technology, and engineering department), and three science education researchers at a public university in the same city (a chemistry professor, a physics education postdoc, and a doctoral student studying chemistry education). Some members of this team have collaborated for up to 14 years, while others more recently joined the team 2–3 years ago. Members of the team observe in each other's classrooms, both in person and through video, and design and lead professional development (PD) for science teachers and administrators in the school district as well as nationally. A 6-year history is condensed into a story of the process, aided by field notes collaboratively recorded by one of the researchers and one of the teachers.

Development of our practical observation (the product of Stage 1) took shape through several half- and full-day meetings over 1 year. The design team determined that four groups of stakeholders were necessary to forming a practical observation of *the FA practices of science teachers*: teachers, students, school district personnel, and science education researchers. Initially, we followed a process of identifying questions that the team considered to be important, and then observing critically in our own and each other's classrooms, considering how our students experienced FA, and opening discussions with colleagues about PD and resources for FA that they wished for. Our questions evolved over time as we read and discussed current research (Coffey et al. 2011; Colestock and Sherin 2015; Furtak et al. 2014; Talanquer et al. 2015) and reported on our informal investigations.

At the conclusion of this practical observation, we developed a characterization of the problem of practice at the intersection of four stakeholders' priorities. We also agreed upon a definition of FA based on our review of literature: "the process used by teachers and students to recognize and respond to student learning in order to enhance that learning, during the learning" (Bell and Cowie 2001). With reference to Talanquer et al. (2015), we identified three aspects of teachers' approaches to FA that we wanted to better understand: noticing, interpreting, and acting.

Generalizable Explanation

The design team turned next to developing cycles of validation to study the problem of practice, with the goal of creating a coherent explanation that would accomplish two aims: to strengthen theory on attending to students' sense-making and to offer specific guidance for experienced teachers to help them assess their own discourse practices in support of students' sense-making.

Focus Groups Analyzing Student Work

In our first cycle, we aimed to characterize how experienced science teachers notice and interpret students' ideas, how they propose to act on their interpretations, and how they consider FAs to make possible the enhancement of learning during the learning. Following a review of literature, we designed an approach to collecting data via focus groups of experienced chemistry teachers who analyzed students' written artifacts from an open-ended FA designed to uncover students' thinking about how to control chemical reactions. Five focus groups (23 teachers in total) discussed what the teachers paid attention to in the student work, how they interpreted it, what actions they would take based on this, and how the FA could be improved to better capture students' thinking. Analysis of the focus group data resulted in an initial model (Fig. 1) of FA enactment that characterized how teachers evaluate student thinking and plan actions based upon their evaluation (Clinchot et al. 2017). In this model, the teacher initially notices student thinking, either in a descriptive or inferential manner. Next, the teacher interprets what is noticed, either with an evaluative or interpretive approach. Finally, the teacher acts upon what has been interpreted, either by remediating to correct errors or by responding to the disciplinary content in students' thinking.

Outcomes of this first cycle included that it is productive to define scales within noticing, interpreting, and acting, as others have done (Lineback 2015; Talanquer et al. 2015). We found that the teachers' positions on these scales tended to occur in clusters, and we formed four composite "FA personalities" of the most prevalent clusters (Clinchot et al. 2017). These analyses led us to recognize that noticing and interpreting are closely linked and difficult to analyze separately.

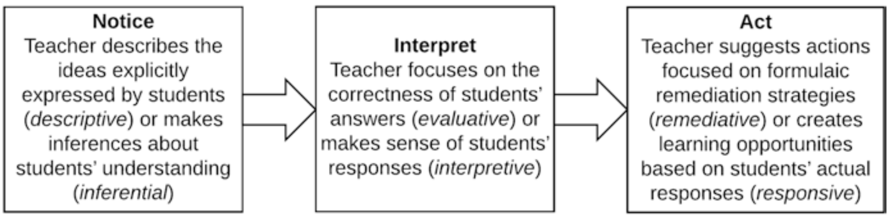

Fig. 1 Initial model of formative assessment enactment

Development of an FA Enactment Resource

In our second cycle, we focused on how FA is enacted by teachers in their classrooms. We expanded the process to include more informants, including the design team, 9 teachers who participated in a 6-month series of PD workshops, and 42 science teacher leaders who participated in a 2-day retreat. We carried out this cycle in a back-and-forth process that oscillated among using the emerging resource to develop and lead PD, collecting data in teachers' classrooms, asking teachers to use the developing resource to analyze their own videos, and further developing the resource through analysis of classroom videos and analysis of field notes from PD. We describe the product of this cycle and how it emerged via four phases. The phases are not design iterations, i.e., they were not marked by articulations of findings in relation to the problem of practice. Rather, the phases are marked by advancements in critical elements of the design-based process, especially the complexities, challenges, and major learnings in each phase (Collins et al. 2004).

Phase 1: The Design Team's Beginning FA Enactment Model

During the first and longest phase, the design team focused on forming an initial FA enactment resource and planning for a district-level PD that would serve to test its usefulness. This work took place at monthly entire-group meetings, as well as more frequent meetings of the researchers in between.

At the beginning, we clarified our understanding around the purpose of FA to ensure common grounding. We built from Bell and Cowie (2001) who specify that FA enhances learning during the learning process. This pushed us to examine classroom discourse (Lemke 1990). As a way to gather experience with this approach, the teacher leaders on the team video recorded FA in their own classrooms. Using these videos, the team explored the rhythm of FA as it unfolds in discourse among a teacher and students. As a way to organize the interactions, the team considered FA as moving in cycles where the teacher first elicits students' ideas, notices something about them, interprets some kind of meaning, and then acts, after which this cycle repeats (Ruiz-Primo and Furtak 2007; Windschitl et al. 2018).

The design team recognized that the PD would require attention to both domain-general (e.g., promoting claims-evidence reasoning) and domain-specific teaching (e.g., exploring the difference between melting and dissolving in chemistry). We recognized that the dimensions of our FA enactment model (noticing/interpreting and acting) could address both of these, because the way we were framing noticing/interpreting requires attention to the substance of students' thinking (Coffey et al. 2011). We first considered teacher noticing/interpreting to exist on a spectrum from evaluative (i.e., seeing student responses through a lens of correct or incorrect) to inferential (i.e., treating students' ideas as having sensible origins) (Talanquer et al. 2015) and acting to occur on a spectrum from what we called at that time prescriptive (i.e., guiding to particular ideas through directive discourse) to responsive (i.e., creating opportunities for proactive student thinking). As a way to investigate these in the context of classroom videos and determine parts of videos to use during the PD, researchers on the team brought several videos with transcripts from different design team teachers' classrooms. The team tried to understand the students' thinking evident in the videos and then characterize teachers' acts as prescriptive or responsive. Together, we learned it is important to begin video analysis by unpacking the sense in students' ideas so as to mitigate the urge to comment on what a teacher should have done. Starting with a scheme based on Lineback's (2015) idea of focus and activity redirections, team members proposed different teaching acts to be included under the prescriptive or responsive categories and gradually refined that list over time.

During one of our meetings involving the review of a classroom video, the team ended up in extensive discussion with the teacher about his moves and what motivated his choices. Because the teachers on the design team found this conversation valuable, we identified this approach as a useful way to learn more about PD participants' intentions and noticings. Planning ahead to Phase 3, we asked teachers who were going to participate in a 6-month PD series to video record FA in their own classrooms and conduct self-interviews about the learning goals of their FAs.

The teachers on the design team also found it very helpful to think about different types of moves a teacher could choose to make; thus we wanted to present the resource as a toolkit of choices. The teachers also appreciated that a choice depends on understanding not only the students' thinking but also the context of the thinking that they worked to unpack. For example, the teachers pointed out that there are affective aspects to supporting students' meaning making, such as affirming progress or empathizing with struggle. They also brought up contextual influences in the form of dilemmas teachers face (Windschitl 2002), such as the pedagogical dilemma of time pressure imposed by preparing students for standardized tests.

The team, however, grappled with the grain size of coding. Although coding of individual teaching moves was beneficial because it helped teachers focus on their choices in those moments, the team questioned whether line-by-line coding of noticing/interpreting would be productive in the PD. To mitigate the concern, we decided to focus on teaching moves at the extremes of the prescriptive-responsive spectrum. We chose contrasting cases of videos from the classrooms of two design team teachers. Each video demonstrated a well-executed FA activity in which students discussed open-ended chemistry problems. One teacher (Kitty) used mostly prescriptive teaching moves, and the other (Thomas) used mostly responsive teaching moves. Kitty asked students a series of leading questions in quick succession to move them to a specific idea. The students responded to the questions in short utterances, either agreeing or introducing new questions that spurred Kitty to respond in the same ways. Thomas facilitated a discussion by listening carefully to students' ideas, rephrasing them as necessary, highlighting inconsistencies, and challenging students to resolve differences.

Phase 1 was marked by the following insights: (1) it is necessary to unpack the complexities of teachers' choices behind their teaching moves beginning with openly discussing what could be the sense behind students' ideas, and (2) focusing on clear-cut contrasting cases of the extremes of the teaching spectrum helps teachers articulate the logic behind their choices.

Phase 2: Testing the Initial FA Enactment Resource

Having worked out a preliminary FA enactment resource, the design team tested it with 42 preK-12 science teacher leaders throughout the district at a day-long retreat. The retreat was led by the design team's teacher leaders. This arrangement for facilitation was intentional because it foregrounds the agency of teachers in making decisions about their FA practices (Stroupe 2017).

The FA enactment resource was used in two main activities at this workshop to probe its usefulness in teachers first examining other teachers' FA practice, and then practicing decisions about how FA practices could be different. In the first activity, teachers experienced two engaging lessons about electrochemistry while taking the role of learners (the topic was chosen because few teachers knew it well). These lessons had deliberately designed teaching moves at prescriptive or responsive extremes. The teachers recounted their experiences as learners and then compared the teaching moves in each lesson. Considering their role as students, the prescriptive teaching moves engendered feelings of passivity and comfort in the way information was presented in a scaffolded manner. In contrast, when in the role of student, they experienced the responsive moves as animated argumentation, requiring self-reliance and peer input to figure things out, and feeling frustrated as they lingered in confusion. Imagining what it would be like to teach in each way, they likened prescriptive teaching to a teacher's ship carrying its passenger students to a destination and spoke of responsive teaching as facilitating discussion through questioning, repeating, and seeking clarification of student ideas. Teachers' descriptions of the student and teacher perspectives established that the resource would be effective in helping teachers account for both perspectives.

In the following activity, teachers worked with the two videos (of Thomas and Kitty) previously chosen. Teachers at the workshop *first* analyzed the students' thinking (saying what they noticed and interpreted) and then examined each teacher's actions using the resource that specifies different kinds of prescriptive and responsive teaching moves collected in Phase 1. Many teachers struggled to focus on identifying the sense in the students' thinking, instead gravitating toward commenting on teaching moves, the coding of which teachers found easier and more natural. This prompted the design team to consider how to better support teachers to notice/interpret in ways that attend to sense in students' thinking.

Phase 2 was marked by this insight: there is benefit in connecting learners' experiences in prescriptive vs. responsive extremes with deliberately orchestrated teaching moves.

Phase 3: Further Development of the Model in PD

The FA enactment resource, specifically the dimensions of noticing/interpreting and acting, was further tested by teachers in a full-day PD workshop during a 6-month PD program with nine K-12 science teachers. Participants used the FA enactment resource to analyze *their own* videos. The teachers were asked to video record an FA activity in their classroom from their own vantage point (using chest harnesses to which their smartphones are attached). Each 10–20-minute recording included the teacher's launch of the FA, interactions with students, and a wrap-up of the activity. Before examining their own teaching moves and those of a colleague, teachers extensively analyzed their videos for students' thinking. As a lead-in to analyzing their teaching moves, participants were first oriented to the acting portion of the FA enactment resource while looking at some of Thomas's and Kitty's moves.

We gleaned four insights from Phase 3. First, the teachers continued to be challenged to focus on the substance of student thinking in considering teaching moves and imagining other possibilities. During the first part of the workshop, which had the exclusive purpose to make sense of students' thinking, teachers primarily evaluated students' ideas as correct or incorrect. They also appeared to be much more comfortable discussing domain-general (e.g., is the student making claims and supporting with evidence) than domain-specific matters (e.g., how is the student thinking about hydrogen bonding). The design team recognized a need to better support noticing/interpreting, particularly the disciplinary substance of students' thinking (Richards and Robertson 2016). We decided for a future phase to create short video segments showing student discussion up to the point of, but not including, the teacher's move, to open space for teachers to discuss multiple possible moves based exclusively on their interpretations of student thinking. This takes advantage of teachers' inclinations to focus on teaching moves, but places emphasis on identifying evidence of students' thinking and the teacher's interpretations of it.

Second, the design team found that care must be taken to prevent dichotomous thinking with respect to teaching acts. That is, teachers perceived prescriptive teaching to be bad and responsive teaching to be good. In a concluding discussion with the teachers at the end of the workshop, participants and the design team came to the idea that the difference between the two is *who is doing the sense-making* (in prescriptive moves, sense-making is at the teacher's initiative, while in responsive moves the student is the protagonist) and that there are appropriate times to use one or the other. A benefit of first watching the videos from Thomas and Kitty was that participants engaged in examining their own videos more productively because both teachers taught effectively in the two extremes. We also recognized that there were negative connotations associated with some of the vocabulary that contributed to teachers' interpretations of the types of teaching acts. We revised wording in the model to reflect teachers' intentions for why they may intentionally choose particular actions. We changed prescriptive to directive, since the latter is more descriptive of the intention behind this type of advancing act, i.e., the teacher intends to direct students toward a particular science view.

Third, the participants were universally appreciative of the opportunity to systematically reflect and comment on their teaching moves using the FA enactment resource and of the opportunity to contribute to this resource by suggesting additions to the resources as they examined their videos. In this sense, a dual practitioner/researcher lens helped reduce the vulnerability threat in examining their classroom practice with peers. Examining each other's videos with common codes helped to lift teachers from the uniqueness of their classrooms into discussing generalized experiences across classrooms.

Fourth, during the discussions teachers had about videos that they had separately analyzed prior, most teachers elaborated on the context behind their moves. We noticed that, particularly when there were emotions (e.g., frustration, surprise, concern, joy) in the teachers' written comments about their own teaching moves, they would give contextual explanations that expanded upon ongoing issues spanning multiple lessons with particular students. We recognized an important synergy between teachers' explanations of their moves and our interpretations of them. For the research, we built in a mechanism for interviewing teachers.

Phase 4: Researchers' Refinement of the Coding Framework

The experiences of Phases 1–3 allowed researchers on the design team to analyze complementary data sources that would provide valuable perspectives on teachers' FA enactment. These sources included teachers' self-interviews about the FA activity, classroom video recordings of the activity, their analysis of specific videos within the activity, field notes taken by researchers at all PD meetings, and anonymous evaluations administered by external evaluators after each workshop. Using a well-defined process, each of the researchers on the design team systematically analyzed these data sources by documenting aspects of teachers' purposes for their FA from their self-interviews, using a coding scheme to characterize teaching moves from their classroom videos, and assessing teachers' in-the-moment purposes and noticings from their comments on videos.

The details of this analysis and findings are presented elsewhere (Dini et al. 2019); here we describe the outcomes, which included three main developments in improving the FA enactment resource. The first development related to a challenge of differentiating eliciting and advancing actions. When analyzing teaching moves, we sometimes found it difficult to discern whether the teacher's intention was to find out more about students thinking (*eliciting*) or to advance it toward canonical understanding (*advancing*).

The following exchange illustrates this issue. The teacher (codenamed Terra) is discussing with a student (codenamed D1910) differences between parallel and series circuits. The student is comparing three circuits and refers to one in which two resistors are connected in parallel with a voltage source:

1. ***D1910:*** *Doesn't the current equal each other when it's parallel (referring to the two branches after the ammeter in the circuit)?*

2. ***Terra:*** *So you're saying the current here equals the current here (pointing to two points in the circuit).*
3. ***D1910:*** *Yeah.*
4. ***Terra:*** *Okay. And then if you wanted to compare that current (in the circuit of interest) to this current (in another circuit), what would you get? How would you do that?*

In this moment, Terra may have intended to learn about D1910's understanding of current, but she also appears to have advanced it by implicitly requesting that D1910 provide reasoning to justify her thinking. In this sense, such a question can play a dual role of informing the teacher about a student's ideas and advancing the student's understanding of scientific practice. Terra is acting to uncover more about the student's thinking (i.e., eliciting); however, without knowing more about Terra's explicit intentions in asking the question in turn 4, it is difficult to say whether she is also trying to advance D1910's thinking.

The second development was the recognition that a FA enactment does not take place in a linear manner (Fig. 1). Rather, it takes place as complex nonlinear sequences of teacher noticing/interpreting followed by eliciting or advancing acts. The resource was modified accordingly, to guide teachers to understand the centrality of noticing/interpreting student thinking in an FA enactment, and the two kinds of acts that follow from it: eliciting or advancing (Fig. 2).

The third development was that teachers can have multiple and often simultaneous purposes while enacting a FA (e.g., developing students' content understanding, attending to students' learning processes, cultivating students' agency). Overarching purposes are often filtered by contextual influences that shift and shape teachers' in-the-moment purposes. And in-the-moment purposes can also be individualized to particular students and can grow out of specific teacher-student relationships. For instance, the same teacher (Terra) communicated an overarching aim for her FA to learn whether her students could remember and apply circuit rules to an open-ended

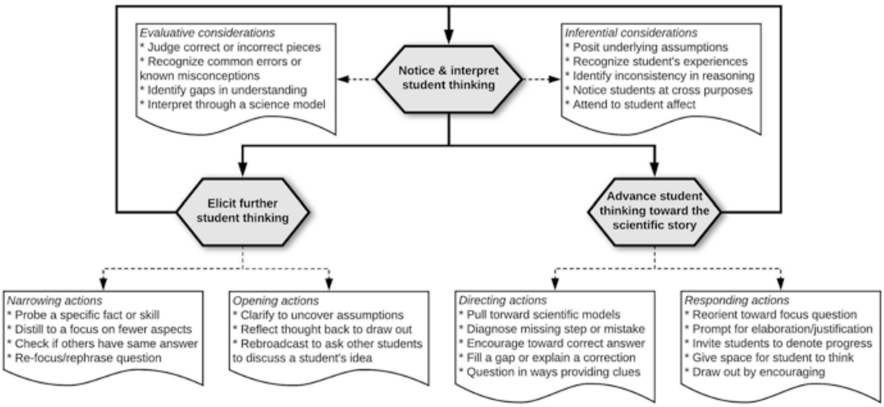

Fig. 2 A summarized version of the FA enactment resource emerging from Cycle 2

conceptual problem. In the course of interacting with one group of struggling students, Terra focused on supporting the students' reasoning. Immediately after this, she moved to another struggling group and simply explained the rule and had them move ahead on the problem from that point. In discussing this with Terra, we learned that contextual influences clearly shifted and shaped her in-the-moment purpose with each group. Terra and other teachers found the opportunity to reflect on these purposes and associated influences—often implicitly operating—very useful in learning to become more intentional about their FA practice.

Phase 4 was marked by these insights: (1) a single move can sometimes be both eliciting and advancing, (2) in-the-moment purposes shift and also shape teaching moves, and (3) a teacher's in-the-moment purposes often are specific to individual students, incorporating disciplinary content-, general-, and domain-specific processes and affect-related goals that the teacher has for the student.

Values of the Design-Based Process

The FA enactment resource was a concrete product that emerged from the PD and represents the PPK described by Bereiter (2014). Bereiter describes PPK as being both procedural and declarative. It is knowledge that is able to be "communicated symbolically, argued about, combined with other propositions to form larger structures, and so on" (p. 5). Rather than being a codification of practice, it is for the purpose of solving problems. The PD guided and opened space for teachers to focus on the substance of student thinking and reflect on their teaching acts in relation to this. The teachers imagined different possible ways of supporting their students, including in enacting different kinds of eliciting and advancing (Fig. 2). Teachers valued the FA enactment resource for the lens it provided to see and characterize their classroom discourse, and how that discourse supported or hindered student learning.

Having teachers contemplate the combination of their in-the-moment purposes and the larger purposes of their lessons supported them in understanding the decision-making around teaching acts that were often taken on a subconscious level. They appreciated that the FA enactment resource characterized the different types of actions that teachers have in their own repertoires and can employ intentionally and strategically in order to support student outcomes. In subsequent design team meetings with the district science administrators, we learned that what teachers value the most in the FA enactment resource, as well as its use in PD, is the capacity it develops in teachers to lead from the classroom.

The team's design-based process also contributed to theory on attending to students' sense-making. Rather than starting with the design of elicitation questions, as many current models of FA suggest (e.g., Ruiz-Primo and Furtak 2007; Windschitl et al. 2018), we learned that the teacher's noticing and interpreting is central to FA enactment. Honoring the teacher as the agent in achieving the goals of learning for students emerged as the most important commitment in the articulation of the problem of practice addressed by our process. Based on this, we advance a further hypothesis that teachers can enact more intentional teaching moves when they have the power to recognize when it is beneficial for students or the teacher to be doing the sense-making in a given learning situation.

Acknowledgments This research was supported by US NSF award DRL-1621228. The authors recognize the contributions of the other members of the design team whose work is described here: Scott Balicki, Gregory Banks, Michael Clinchot, Robert Huie, Rebecca Lewis, Pamela Pelletier, Raúl Orduña Picón, and Holly Rosa.

References

Anderson, T., & Shattuck, J. (2012). Design-based research: A decade of progress in education research? *Educational Researcher, 41*(1), 16–25. https://doi.org/10.3102/0013189X11428813.

Banilower, E. R., Smith, P. S., Weiss, I. R., Malzahn, K. A., Campbell, K. M., & Weis, A. M. (2013). *Report of the 2012 national survey of science and mathematics education*. Horizon Research, Inc. (NJ1).

Bell, B., & Cowie, B. (2001). The characteristics of formative assessment in science education. *Science Education, 85*(5), 536–553. https://doi.org/10.1002/sce.1022.

Bereiter, C. (2014). Principled practical knowledge: Not a bridge but a ladder. *Journal of the Learning Sciences, 23*(1), 4–17. https://doi.org/10.1080/10508406.2013.812533.

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions. *Journal of the Learning Sciences, 2*, 141–178. https://doi.org/10.1207/s15327809jls0202_2.

Capps, D. K., Shemwell, J. T., & Young, A. M. (2016). Over reported and misunderstood? A study of teachers' reported enactment and knowledge of inquiry-based science teaching. *International Journal of Science Education, 38*(6), 934–959. https://doi.org/10.1080/09500693.2016.1173261.

Clinchot, M., Ngai, C., Huie, R., Talanquer, V., Lambertz, J., Banks, G., … & Sevian, H. (2017). Better formative assessment. *The Science Teacher, 84*(3), 69–75.

Coffey, J. E., Hammer, D., Levin, D. M., & Grant, T. (2011). The missing disciplinary substance of formative assessment. *Journal of Research in Science Teaching, 48*(10), 1109–1136. https://doi.org/10.1002/tea.20440.

Colestock, A., & Sherin, M. G. (2015). What teachers notice when they notice student thinking: Teacher identified purposes for attending to student thinking. In A. Robertson, R. Scherr, & D. Hammer (Eds.), *Responsive science teaching* (pp. 126–144). New York: Routledge.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *Journal of the Learning Sciences, 13*(1), 15–42. https://doi.org/10.1207/s15327809jls1301_2.

Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher, 32*(1), 5–8. https://doi.org/10.3102/0013189X032001005.

Dini, V., Sevian, H., Caushi, K., & Orduña Picón, R. (2019). *Characterizing the formative assessment enactment of experienced science teachers* (Manuscript submitted for publication).

Furtak, E., Morrison, D., & Kroog, H. (2014). Investigating the link between learning progressions and classroom assessment. *Science Education, 98*(4), 640–673. https://doi.org/10.1002/sce.21122.

Lemke, J. L. (1990). *Talking science: Language, learning, and values*. Norwood: Ablex Publishing.

Lineback, J. E. (2015). The redirection: An indicator of how teachers respond to student thinking. *Journal of the Learning Sciences, 24*(3), 419–460. https://doi.org/10.1080/10508406.2014.930707.

Mehan, H. (1979). *Learning lessons*. Cambridge, MA: Harvard University Press.

Richards, J., & Robertson, A. D. (2016). A review of the research on responsive teaching in science and mathematics. In A. Robertson, R. Scherr, & D. Hammer (Eds.), *Responsive science teaching* (pp. 36–155). New York: Routledge.

Roth, K., & Garnier, H. (2007). What science teaching looks like: An international perspective. *Educational Leadership, 64*(4), 16–23.

Ruiz-Primo, M. A., & Furtak, E. M. (2007). Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry. *Journal of Research in Science Teaching, 44*(1), 57–84. https://doi.org/10.1002/tea.20163.

Stroupe, D. (2017). Ambitious teachers' design and use of classrooms as a place of science. *Science Education, 101*(3), 458–485. https://doi.org/10.1002/sce.21273.

Szteinberg, G., Balicki, S., Banks, G., Clinchot, M., Cullipher, S., Huie, R., … & Sevian, H. (2014). Collaborative professional development in chemistry education research: Bridging the gap between research and practice. *Journal of Chemical Education, 91*(9), 1401–1408. https://doi.org/10.1021/ed5003042.

Talanquer, V., Bolger, M., & Tomanek, D. (2015). Exploring prospective teachers' assessment practices: Noticing and interpreting student understanding in the assessment of written work. *Journal of Research in Science Teaching, 52*(5), 585–609. https://doi.org/10.1002/tea.21209.

Weiss, I. R., Pasley, J. D., Smith, P. S., Banilower, E. R., & Heck, D. J. (2003). *Looking inside the classroom*. Chapel Hill: Horizon Research Inc.

Windschitl, M. (2002). Framing constructivism in practice as the negotiation of dilemmas: An analysis of the conceptual, pedagogical, cultural, and political challenges facing teachers. *Review of Educational Research, 72*(2), 131–175. https://doi.org/10.3102/00346543072002131.

Windschitl, M., Thompson, J., & Braaten, M. (2018). *Ambitious science teaching*. Cambridge, MA: Harvard Education Pub Group.