

Supporting Teachers' Concept Development Through Aggregation

Grace A. Chen, Elizabeth Metts, Ilana Horn, and Katherine Schneeberger McGugan
grace.a.chen@vanderbilt.edu, elizabeth.c.metts@vanderbilt.edu, ilana.horn@vanderbilt.edu,
k.schneeberger.mcgugan@vanderbilt.edu
Vanderbilt University

Abstract: This study analyzes transcripts of conversations in which mathematics teachers and researchers debrief videotaped lessons by, in part, examining aggregated classroom data from the videotaped lesson. We conclude that aggregating data in debrief conversations can support teachers' concept development when the aggregation a) demonstrates internal contrasts and b) is underscored by participants' discursive moves. Consequently, we recommend that facilitators seeking to prompt teacher learning use lesson-level aggregations to identify and press on comparisons and distinctions in teaching practice. This study can inform research on teacher learning by unpacking how a common practice—aggregating data—contributes to teachers' concept development and has implications both for practitioners and for the emerging field of classroom data visualization.

Introduction

To summarize information about how teaching unfolds throughout a lesson, practitioners—administrators, evaluators, and teachers themselves—often aggregate data from observing teachers' classrooms, such as tallying teacher or student behaviors or quantifying time-on-task (e.g., Mason et al., 2019; Simonsen et al., 2013). A growing body of work on data visualization aims to make such aggregations accessible to teachers, with researchers developing software that illustrates which students are asked which types of questions or which solicitation methods teachers use to ask questions (Herbel-Eisenmann & Shah, 2019) or how class time is spent on different activities or in different configurations (Nagro et al., 2020). Such research, whether rooted in coaching, teacher evaluation, or data visualization, often claims that aggregating data can lead to teacher improvement, but rarely reports how teachers perceive or use such data. In a notable exception, Chen and colleagues (2020) demonstrated that seeing and discussing aggregated evidence of academically productive talk increased teachers' use of academically productive talk but did not investigate how or why it did so. Without understanding how or why aggregating classroom data contributes to teacher learning, aggregators risk the co-optation of powerful tools for problematic purposes, such as teacher or student surveillance under the guise of accountability (Webb, 2005). At best, introducing new data collection practices may increase demands on teachers' time and resources with little benefit for teachers and students. Therefore, this study investigates *how the aggregation of classroom data can support teachers' concept development in lesson debrief conversations*.

Conceptual framework

To do so, we draw on situative theories of teacher learning that take knowledge, and therefore concepts, to be both shared and distributed among actors and artifacts (Hall & Horn, 2012; Hall & Jurow, 2015). Rather than being individually held and carried from situation to situation, the concepts in circulation in any given situation depend on context. This context may include teachers' values, beliefs, and goals (Aguirre & Speer, 1999); the available representational infrastructure (Hall & Jurow, 2015); and the conversational routines, frames, and stories that are common in a particular community (Bannister, 2015; Yoon, 2016); among other things. Concepts serve explanatory purposes in that they help teachers coordinate and “ge[t] information from the world” (diSessa & Sherin, 1998, p. 1171; see also, Philip, 2011) and in doing so, enable pedagogical judgment (Horn, 2020). For this paper, we operationalize a concept as something—a topic, word, phrase, or idea—that 1) carries local meaning; 2) is related to beliefs, values, explanations, definitions, and other concepts in articulable ways; and 3) make interpretation, decision-making, and action possible.

To understand concept *development*, then, we draw on the conceptual change literature. Instead of treating concepts as purely additive (see diSessa [2006] for the limitations to additive conceptualizations of concept development), we take concepts to be pre-existing and gradually transformed through social interaction. In other words, the set of associated definitions, explanations, values, and beliefs that are attached to a concept shift as teachers engage in—among other social practices—conversations about teaching. Shifts in these associations become evident through shifts in teachers' activity patterns (Hall & Jurow, 2015) or, sometimes and more immediately, shifting narratives and mobilizations for future work (Horn et al., 2017; Kane, 2020).

Research methods

Context and data collection

Data for this study was collected as part of a larger research-practice partnership where the research team and the leaders of a professional development organization (PDO) collaborated to understand and improve their support of veteran mathematics teachers' growth (Horn & Garner, 2021). In this partnership, we engaged in video-based formative feedback (VFF) cycles with school-based teams of teachers. In each VFF cycle, teachers identified inquiry questions about their own practice and selected student groups to record; researchers filmed the teachers' classes using a point-of-view camera worn by the teacher, a whole-class camera mounted on a rotating tripod, and four audio recorders, one for each selected student group; researchers reviewed the collected video and audio to identify clips that would support co-inquiry around teachers' inquiry questions; and the teachers, their school-based colleagues, and researchers engaged in a lesson debrief conversation centered on the selected clips. For timely feedback, each VFF cycle took place within two to four days. We collected the aforementioned classroom video and audio recordings, classroom observation fieldnotes, video recordings and transcriptions of the VFF debriefs, researcher memos, and artifacts from the classroom observations and debrief conversations.

Over two years of data collection, we worked with 13 teachers across six school teams. These teachers had between 5-25 years of teaching experience, and most had obtained or were nearing completion of National Board Certification. We conducted 32 VFF cycles and aggregated lesson-level patterns in 15 of them; in three of those VFF cycles, researchers did not share the aggregated data with teachers during the debriefs. In three of the remaining 12 VFF cycles, multiple types of data were aggregated and shared (e.g., teachers' use of class time and strategies used in students' boardwork). From the 12 VFF cycles, then, there were 15 total instances where lesson-level patterns were aggregated and discussed with teachers during the debrief.

Phase 1 data analysis: Identifying instances of concept development

To select focal and comparative cases, we created analytic memos and tables (Miles et al., 2020) to summarize each instance where lesson-level patterns were aggregated, classified the data aggregated, and identified why the data was aggregated. The most common aggregated data included instances of a particular pedagogical move in a lesson (e.g., list of times teacher tried to get the whole class's attention), teachers' uses of class time (e.g., minutes per activity), and students' boardwork (e.g., list of strategies represented). Aggregation was roughly equally prompted by researchers' hunches (eight instances) and teachers' requests (seven instances). Next, three researchers independently coded each instance for the concepts evident in the debrief and how associated concepts shifted over the course of the debrief. Through discussion, we reached consensus on two instances where aggregation clearly contributed to concept development. Then, we wrote narrative memos tracing the evidence for concept development across each instance. Discussing these memos across the broader research team, in conjunction with reviewing relevant literature, led us to two hypotheses about the mechanisms that contributed to concept development: H1) concept development depends on the aggregated data presenting a clear contrast between what teachers expected to see in their classrooms and what actually happened and H2) concept development depends on participants' use of strong discursive moves to introduce and shift associated concepts.

Phase 2 data analysis: Comparative case analyses

To explore these hypotheses, we conducted comparative case analyses (Yin, 2014). Through open, concept-driven coding, we identified three relevant comparative cases that had similar surface characteristics as the two focal concept development cases but that did not result in clear concept development (see Table 1). Next, we used MAXQDA qualitative analysis software to inductively code the focal and comparative case debrief transcripts (Miles et al., 2020) and visualize the coded data over time.

Table 1
Cases coded in Phase 2 analysis

Case (* denotes focal case)	What was aggregated	Participating teachers
* Including Students	How teachers included students during small group work	Abigail, Clark, Veronica, Franck
* Goes Into Questions	Students' uses of "goes into" Teacher questions during small group work	Lizette, Julie Lee, Doha
Boardwork Creativity	Students' solutions to a particular problem	Clark, Franck
Class Time	Minutes per activity	Clark, Franck

We operationalized H1 by identifying contrasts in aggregated data and then coding debrief transcripts for associated concepts. Visualizing these codes revealed clusters of associated concepts, shifts in associated concepts, when these shifts happened, and what prompted the shifts, which we summarized in analytic memos. For H2, we coded researchers' and teachers' discursive moves with deductive codes drawn from research on activities that support teachers' conceptual change (e.g., proposing solutions [Hall & Jurow, 2015]; asking clarifying questions, offering different interpretations [Horn et al., 2017]) and inductive codes developed for additional moves that were common (e.g., elaboration, affirmation). We examined instances of the most frequent codes using techniques from discourse analysis to determine the meanings, accomplishments, and significance of these moves (Gee, 2014). For example, if a researcher offered a differing interpretation of a problem, we considered whether and how this interpretation was taken up, and by whom, in subsequent turns of talk; if a teacher asked a clarifying question, we considered whether the question served to fact-check what had happened in the videotaped lesson and/or to establish common ground among participants, etc. Finally, we articulated connections between contrasts in the aggregated data, shifts in concepts, and participants' discursive moves.

Findings

We found that although a contrast between teachers' expectations and the aggregated data was neither necessary nor sufficient for concept development (H1), internal contrasts within the aggregated data did seem to be necessary, if also not sufficient, for concept development. For example, in the Including Students Case, teachers did not demonstrate any surprise after seeing an aggregation of their strategies for including an excluded student in small group work. However, the aggregated strategies highlighted a contrast between the intention of supporting students' equal participation as a proxy for equity and supporting individual students' participation as a way of meeting their emotional, social, and agentic needs. Conversely, the Boardwork Creativity Case did not show concept development despite internal contrasts within the aggregated data. Teachers continued to rue the homogeneity of students' solutions even though researchers—and in some cases, the teachers themselves—pointed out instances where students demonstrated unexpected and creative solutions, perhaps because researchers accepted teachers' initial dismissals of the surprising solutions without following up.

Findings from H1 confirmed the importance of participants' discursive moves in prompting concept development, but contrary to H2, we found that researchers' strong moves were neither necessary nor sufficient for concept development. Instead, concept development occurred when either researchers or teachers *offered different interpretations* and then *elaborated* on concepts in ways that diverged from what had been taken as shared. For example, in the Goes Into Case, the teacher initially attributed students' interchangeable uses of "goes into" for exponentiation and multiplication to latency and to her leniency in allowing imprecise language. The lead researcher repeatedly proposed students' unclear conceptual understanding as a factor instead, but the teacher did not take up this interpretation and associate different concepts with it until a second researcher elaborated on how to assess students' conceptual understanding. In the Lee Questions Case, the researcher also repeatedly queried the teacher's interpretations of his questioning strategies, asking about his intentions and how else he could have accomplished them. However, neither the researcher nor teacher elaborated on, say, why his questions did not accomplish his intentions. Although the teacher acknowledged the researcher's queries, even indicating agreement with her concerns, he elaborated extensively on his own pre-existing interpretations and brainstormed how he might revise his future questioning strategies without taking up the researcher's ideas.

In brief, our findings suggest that aggregating patterns supports teachers' concept development when multiple possible interpretations of internal contrasts within the aggregated data are unspooled by participants' discursive moves. Next, we illustrate this finding by describing the aggregated contrasts, shifts in associated concepts, and how both researchers and teachers elaborated on differing interpretations in our focal cases.

Case 1: Including Students

Context and aggregation

In the Including Students case, four Algebra 1 teachers at Banneker High School—Abigail, Clark, Franck, and Veronica—were filmed teaching similar lessons in which students worked at whiteboards in randomly-assigned groups of 2-3 students to write and solve series to find the number of blocks in various towers. Abigail had asked to see examples of how each teacher had promoted discussion or equity, so the research team, led in this VFF by Grace, aggregated a series of six videoclips in which one student in a small group appeared to be left out and a teacher had done something to bring that student into the small group conversation. For example, in one clip, Franck lightheartedly asked a student to talk to another student "because he looks lonely." In another, Veronica directed a follow-up question at a student who had not been participating in the conversation. Taken together, these clips highlighted contrasting strategies for supporting students' participation in small groups.

Shifts in associated concepts

In the first third of the debrief, Grace showed the video clips in rapid succession and teachers engaged in mostly binary assessments of students' participation and of the efficiency and effectiveness of intervention strategies for supporting student participation (e.g., "that's it, quick;" "she looks like she's thinking;" "success"). Then, Grace prompted teachers to share their pedagogical reasoning around the aggregated clips. In the second third of the debrief, teachers named forms of student participation other than discussing solutions and analyzed the aggregated intervention strategies in more detail. For example, Clark wondered whether he should create more opportunities for students to ask their group members for help when they do not do so "naturally:"

- Franck: A problem I have with that is that one kid will definitely always be the asker.
 Veronica: The asker?
 Franck: Yeah.
 Veronica: Or the person being asked.
 Franck: Yes, all the time. Especially in a certain makeup of that group, and I don't know how comfortable that would feel to a kid, being the asker or the askee or whatever, you know? So I just wonder, I don't know.
 Abigail: [...] I tell them when who's talking. So that's part of the equity thing that I think is important is that like [students assigned to] reds talk first. There's red and yellow, reds talk first. All reds have to talk and so then all yellows have to talk...

In this portion of the VFF debrief, teachers discussed the associated concepts of *equal participation as a proxy for equity*, *mitigating status differentials across students*, and *participation as an opportunity for academic learning*. These associated concepts represented a shift from the binary assessments demonstrated in the first third of the debrief and contrasted participation for participation's sake with participation for learning's sake and were made salient by teachers' further discussion of the strategies shown in the aggregated videoclips.

Differing interpretations and elaborations generate further contrasts

Then, Grace invited Franck to introduce Roberto, a student who Franck had described as having a substantially different skillset compared to other students in the class. Teachers viewed a clip in which Franck tried to include Roberto in his small group's conversation, and the final third of the VFF debrief became substantially more complex. Continuing to frame *participation as an opportunity for student learning*, Clark initially interpreted the dilemma as one of preserving Roberto's learning opportunities while maintaining the status-equalizing advantage of randomly-assigned and thus heterogeneous groups. He proposed strategically assigning Roberto and a partner to just work on the first problem while other students moved on, and then added:

It's kind of sad to me that [the first problem is] still sitting there probably unsolved for him and very solvable, and [it would be] a nice accomplishment for him. And he doesn't get to do it because we're already moving on to this other thing. So I'm just wondering how we could make it without, honestly, tracking your grouping...

Franck asked how he could assign Roberto and a partner to work just on the first problem "without singling [Roberto] out." In response, Clark and Veronica brainstormed another strategy: still using random groups but then assigning Roberto to write his groupmates' ideas on the whiteboard. They surmised that Roberto could still have learning opportunities from writing mathematical solutions, even if he did not contribute to figuring them out.

Franck offered a differing interpretation: "But maybe that's like you forcing him to do something he doesn't want to. I mean, is that agency?" Teachers did not immediately take up the concept of *student participation as a reflection of agency*; Veronica and Clark elaborated further on Clark's original interpretation, generating and troubleshooting additional strategies that could elicit Roberto's participation while preserving his learning opportunities. After several minutes, Franck re-introduced his question more directly.

- Franck: Do you think he's getting something from standing up there with those two?
 Clark: I personally would say no.
 Veronica: You say no?
 Clark: Very little. Not—
 Franck: More than if he sat by himself.
 Veronica: There's more to school though.
 Clark: Abigail?

- Abigail: Yeah, I do think he's probably getting some, I mean, I don't know. I don't know the student, but I think that he is.
- Franck: [turns to his aide] Do you think he's getting something?
- Aide: Roberto? I don't think so.
- Veronica: But you have to think about the social context, too. It looks like he's participating, his body language is "hey, I'm doing this thing that I'm supposed to be doing" and so maybe we're meeting a different need, which is allowing him to be near and participate in a low-risk way.
- Abigail: Participate in school.

In this exchange, Franck prompted his colleagues to reconsider his differing interpretation: that participating in small groups might be meeting students' needs beyond learning mathematics. This time, Veronica picked up and elaborated on this interpretation, suggesting that Roberto's current form of participation may be serving social needs. Abigail and Clark followed up with additional ways that students' social and emotional needs might be met by listening quietly to their groupmates rather than actively figuring out solutions. Finally, Franck asked what would happen if he simply told Roberto's group that "when I come back, I want Roberto to be explaining number one. Now how does that make Roberto feel? That's my question." In contrast to teachers' earlier discussions of strategies for supporting student participation, Franck identified a proposal with the potential to address both Roberto's mathematical learning and his social need to perform participation in a low-risk way, and solicited his colleagues' feedback on how this strategy would make Roberto feel, as opposed to simply whether it might "work" in the sense of compelling Roberto to talk.

In the final third of this debrief, the concepts associated with supporting student participation shifted from *participation as a proxy for equity* and *participation as an opportunity for academic learning*—both of which suggest that teachers' primary goal is to increase student talk and thereby student learning—to *participation as meeting students' social, emotional, and agentic needs*. These shifts occurred when teachers shared differing interpretations of pedagogical dilemmas illustrated by the aggregated videoclips—specifically, questioning how intervention strategies might affect students' agency and their needs other than academic learning—and then elaborated on these differing interpretations. Doing so led to concept development and mobilization for future work by expanding what teachers considered when assessing intervention strategies.

Case 2: Goes Into

Context and aggregation

In the Goes Into case, Lizette had taught an Algebra 2 lesson in which students engaged in small group work to practice evaluating logarithms. Lizette asked researchers (Brette and Lani) to attend to how students were thinking about logarithms. While reviewing video, Brette and Lani noticed that students were using the phrase "goes into" to describe both multiplication (e.g., "two goes into 128 sixty-four times") and exponentiation (e.g., "three goes into nine twice"). This pattern appeared across multiple student groups and throughout the lesson. To highlight the contrast, Brette and Lani selected two videoclips of the same student group using "goes into" in different ways. Brette introduced the data by saying, "there's also the issue of 'goes into' and the way that students were using that. This group, which was the blue group, they were using 'goes into' in both ways." Lizette initially interpreted Brette's observation as an issue of *imprecise language risking miscommunication*:

So then, if someone is talking to a person who's using it both ways, they don't necessarily know—maybe the kid who's saying it has it in their head straight, but the people who are listening might not. And also maybe the one who's saying it might not.

Brette, however, suggested that "it was interesting that some students were using it both ways to mean both things at different points." By highlighting internal inconsistencies within the same students, rather than inconsistencies between how different students were using the phrase, she gestured towards *imprecise language as an indicator of students' conceptual understanding*.

Shifts in associated concepts

Lizette then explained that she had originally taught logarithms using a number line representation to conceptually illustrate the difference between multiplication and exponentiation, but she had done a long time ago and students may have forgotten in the intervening months. In doing so, she attributed students' *imprecise language to be a reflection of latency* rather than an indicator of their (potentially unclear) conceptual understanding.

- Lizette: But when it's so far separated from when we did that—'cuz that was in the beginning of the semester or the end of last semester, I can't even remember. It was like months ago either way, right? [...] It's been so long that then, especially when the vocabulary stuff comes up, makes sense that it's all over the place.
- Brette: But there were also things like the $\log_b b = 1$. A lot of kids were like, "Oh that is clearly one." And [student] back here was like, "Yeah, y'all, that's one."

In identifying a similarly-aged concept that students appeared to remember clearly, Brette contradicted Lizette's interpretation that latency was the primary cause of students' imprecise language. Lizette, however, continued to elaborate on the pedagogical dilemmas that latency presented. She noted that her Calculus students still struggled with logarithms even though some had also learned the number line representation and wondered whether she ought to be reteaching logarithms to them too. In doing so, Lizette recognized that the number line representation alone could not ensure students' conceptual understanding but continued to attribute student confusion to latency.

Later in the VFF debrief, Brette took a different approach to probing students' uses of "goes into:"

- Brette: I got the sense that maybe [students] came up with the 'goes into' language and you were rolling with it?
- Lizette: [...] I just wasn't intentionally trying to go with it but I also wasn't as strict about—like when we did the warm up, I probably wasn't very strict about it. [...] I should have said something. [...] Sometimes in my head maybe I'm like, "Yeah okay now we've got it revoiced the way I wanted it." But I didn't articulate why that was the way I wanted us to go. And so then they don't see that one way is going to be better for them to use than another when it comes to later stuff.

Lizette's elaborations suggested that she interpreted Brette's prompt as associating *imprecise language with teacher leniency*. She said she should have been stricter about monitoring students' language and more explicit about explaining why precise language mattered. Although she shifted her associated concepts, attributing latency to teacher leniency instead of latency, this was not the shift that Brette had hoped for.

The importance of elaboration

Brette made another bid for her interpretation of students' understanding by positing that students might not have understood that there *were* multiple meanings of "goes into:"

- Brette: I didn't get a sense that kids who were using the "goes into" phrasing knew—
- Lizette: —not to
- Brette: Or were like clearly aware that there were two different uses of "goes into."

This time, Lani jumped in to elaborate on Brette's interpretation.

- Lani: We listened to a bunch of these last night, the "goes into" examples, and we were just listening to a lot of their talk. And I was trying to see if I thought there was a kind of shared understanding of that inverse relationship between exponentiation and logarithms. Like if they saw it as a process or if they saw it as a procedure.
- Brette: Yeah, it felt like sort of a translation thing?
- Lizette: Yeah, that makes sense.
- Brette: Like log whatever of whatever is a language that I'm not familiar with, and so I'm going to rewrite it into words that I know.
- Lizette: Yeah, instead of understanding conceptually that it undoes.

Lani's elaboration, and Brette's further elaboration, seemed to finally cement for Lizette the differing interpretation that Brette had been trying to present throughout the VFF debrief, and she articulated in her own words that students' imprecise language likely indicated the extent of their conceptual understanding. Next, Brette and Lizette reviewed some of the specific examples that Lizette had used in her activity and how they may have contributed to students' confusion. As Lizette began to identify changes she wanted to make to this lesson for next year, Lani pressed Lizette to continue thinking about her current class of students and the state of their conceptual understanding of logarithms:

With how they're talking about it, if they forgot the procedure a month from now, could they recreate it with the words they're saying about what's going on? And I didn't think so. I think that they were kind of like going through the number line and doing things with it and they were kind of powering through that one calculation.

Lani provided a concrete model for how Lizette could treat students' talk as evidence of their conceptual understanding; this elaboration framed students' imprecise language as a useful source of formative assessment rather than merely an inevitability due to latency or a consequence of teacher leniency.

After researchers' persistence in offering an interpretation of students' imprecise language that differed from Lizette's initial interpretations, and more importantly, elaborating on it, Lizette incorporated their interpretations by planning how she would revisit students' conceptual understanding of logarithms on Monday:

So I think I want to just do just generally solving equations, and kind of like scaffold it and end with logs maybe? [...] So looking at a lot of those inverses. That's kind of where my head is at currently so we can kind of go through inverses again and then throw in an exponential. [...] So reintroduce it again as "this is the tool that we need to be able to undo."

The solution that Lizette proposed, reintroducing logarithms as a type of inverse, took *imprecise language as an indicator of students' conceptual understanding* instead of reflecting her previous interpretations of imprecise language as a reflection of *latency* or *teacher leniency*. This demonstrated a shift in the concepts that she associated with students' imprecise language. Although the aggregated videoclips made clear the contrast between students' different uses of "goes into," and although Brette had repeatedly offered an interpretation of the contrast focused on conceptual understanding, Lizette did not fully take up this interpretation until Lani (and then Brette) elaborated on this interpretation, suggesting that *elaboration* as a discursive move is critical to concept development.

Discussion

Examining the highlighted contrasts, associated concepts, and key discursive moves in the Including Students and Goes Into cases suggests that teachers' concept development can be supported by 1) the aggregation of patterns in classroom teaching that reveal internal contrasts and 2) elaboration, whether by teacher colleagues or researchers, on different interpretations of those patterns.

Aggregating patterns across student groups, across the minutes of a lesson, across instances of a particular routine or pedagogical move, or other things, collapses an otherwise-disperse set of data points to a more manageable scale, making comparisons and distinctions more obvious and easier for teachers to make sense of. Lampert writes that zooming in and out between scales—what happens in an interaction compared to what happens over multiple interactions, for example—"enrich[es] our understanding not only of different levels of phenomena but also their relationships to one another" (2001, p. 44). In other words, aggregating data coordinates processes that occur at different scales, highlighting what is typical and where there are possibilities for change (Lemke, 2000; Taylor & Hall, 2013). By doing so, aggregation supports shifts in the beliefs, values, and ideas that teachers associate with particular concepts about teaching, thereby developing their intuitions and ability to make in-the-moment decisions (Horn, 2020; Lampert, 2001).

This study brings research on teacher learning into conversation with the growing research on classroom data visualizations, identifying mechanisms by which aggregating data can support teacher learning. We suggest that researchers designing data visualization tools consider how their tools illuminate contrasts at different levels of scale and also attend to the types of conversations about teaching that might be sparked by their tools. Our findings recommend that facilitators who aim to support teacher learning (and/or instructional improvement by way of teacher learning) may find it useful to aggregate lesson-level data and share the aggregations with teachers not because the aggregations themselves necessarily prompt conceptual change but because aggregations can often make visible internal contrasts, whether between teachers' intentions or expectations and observed classroom practice or between their initial assumptions or interpretations and alternate possibilities. This internal contrast, in turn, can provide fodder for the types of discursive moves that lead to concept development, especially if facilitators are able to attend to and even introduce differing interpretations and to prompt elaborations on these different interpretations. Furthermore, classroom observations should be extended beyond the limited "pop ins" that administrators often employ; our research shows the promise of aggregating data and identifying patterns across full lessons instead of focusing on short episodes. Finally, we hope that future research can examine whether aggregating data across lessons and/or across teachers supports learning in similar ways.

References

- Aguirre, J., & Speer, N. M. (1999). Examining the relationship between beliefs and goals in teacher practice. *The Journal of Mathematical Behavior*, 18(3), 327–356.
- Hall, R., & Horn, I. S. (2012). Talk and conceptual change at work: Adequate representation and epistemic stance in a comparative analysis of statistical consulting and teacher workgroups. *Mind, Culture, and Activity*, 19(3), 240–258.
- Horn, I. S. (2020). Supporting the development of pedagogical judgment: Connecting instruction to contexts through classroom video with experienced mathematics teachers. In G. M. Lloyd & O. Chapman (Eds.), *International Handbook of Mathematics Teacher Education: Participants in Mathematics Teacher Education* (2nd ed., pp. 321–342). Brill Sense.
- Horn, I. S., & Garner, B. (2022). *Teacher learning of ambitious and equitable mathematics instruction: A sociocultural approach*. Routledge.
- Horn, I. S., Garner, B., Kane, B. D., & Brasel, J. (2017). A taxonomy of instructional learning opportunities in teachers' workgroup conversations. *Journal of Teacher Education*, 68(1), 41–54.
- Bannister, N. A. (2015). Reframing practice: Teacher learning through interactions in a collaborative group. *Journal of the Learning Sciences*, 24(3), 347–372.
- Chen, G., Chan, C. K. K., Chan, K. K. H., Clarke, S. N., & Resnick, L. B. (2020). Efficacy of video-based teacher professional development for increasing classroom discourse and student learning. *Journal of the Learning Sciences*, 29(4–5), 642–680.
- diSessa, A. A. (2006). A history of conceptual change research: Threads and fault lines. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences*. Cambridge University Press.
- diSessa, A. A., & Sherin, B. L. (1998). What changes in conceptual change? *International Journal of Science Education*, 20(10), 1155–1191.
- Gee, J. P. (2014). *An introduction to discourse analysis: Theory and method* (4th Ed.). London: Routledge.
- Hall, R., & Jurow, S. (2015). Changing concepts in activity: Descriptive and design studies of consequential learning in conceptual practices. *Educational Psychologist*, 50(3), 173–189.
- Herbel-Eisenmann, B., & Shah, N. (2019). Detecting and reducing bias in questioning patterns. *Mathematics Teaching in the Middle School*, 24(5), 282–289.
- Kane, B. D. (2020). Equitable teaching and the core practice movement: Preservice teachers' professional reasoning. *Teachers College Record*, 122(11), 1–40.
- Lampert, M. (2001). *Teaching problems and the problems of teaching*. Yale University Press.
- Lemke, J. L. (2000). Across the scales of time: Artifacts, activities, and meanings in ecosocial systems. *Mind, Culture, and Activity*, 7(4), 273–290.
- Mason, R. A., Schnitz, A. G., Gerow, S., An, Z. G., & Wills, H. P. (2019). Effects of teacher-implemented coaching to increase the accuracy of data collected by paraeducators. *Journal of Behavioral Education*, 28(2), 204–226.
- Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative data analysis* (4th ed.). Sage Publications.
- Nagro, S. A., Hirsch, S. E., & Kennedy, M. J. (2020). A self-led approach to improving classroom management practices using video analysis. *TEACHING Exceptional Children*, 53(1), 24–32.
- Philip, T. M. (2011). An “ideology in pieces” approach to studying change in teachers' sensemaking about race, racism, and racial justice. *Cognition and Instruction*, 29(3), 297–329.
- Simonsen, B., MacSuga, A. S., Fallon, L. M., & Sugai, G. (2013). The effects of self-monitoring on teachers' use of specific praise. *Journal of Positive Behavior Interventions*, 15(1), 5–15.
- Taylor, K. H., & Hall, R. (2013). Counter-mapping the neighborhood on bicycles: Mobilizing youth to reimagine the city. *Technology, Knowledge and Learning*, 18(1–2), 65–93.
- Webb, P. T. (2005). The anatomy of accountability. *Journal of Education Policy*, 20(2), 189–208.
- Yin, R. K. (2014). *Case study research: Design and methods* (4th ed.). Sage Publications.
- Yoon, I. H. (2016). Trading stories: Middle-class white women teachers and the creation of collective narratives about students and families in a diverse elementary school. *Teachers College Record*, 118(2), 1–54.

Acknowledgments

We would like to acknowledge Patricia Buenrostro, Nadav Ehrenfeld, Brette Garner, and Samantha Marshall for their assistance with data collection, and Katherine Forsthoff for her assistance with data analysis. This material is based upon work supported by the National Science Foundation under Grant No. 1620920. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.